



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English

Franklin Chang *

Dept. of English Language and Literature, Hanyang University, Seoul, 133-791, Republic of Korea
 NTT Communication Sciences Laboratories, NTT Corp., Kyoto, Japan

ARTICLE INFO

Article history:

Received 11 June 2008

revision received 26 June 2009

Available online 11 August 2009

Keywords:

Connectionist model

Cross-linguistic

Sentence production

Syntax acquisition

Learning

Word order

English

Japanese

ABSTRACT

Languages differ from one another and must therefore be learned. Processing biases in word order can also differ across languages. For example, heavy noun phrases tend to be shifted to late sentence positions in English, but to early positions in Japanese. Although these language differences suggest a role for learning, most accounts of these biases have focused on processing factors. This paper presents a learning-based account of these word order biases in the form of a connectionist model of syntax acquisition that can learn the distinct grammatical properties of English and Japanese while, at the same time, accounting for the cross-linguistic variability in processing biases in sentence production. This account demonstrates that the incremental nature of sentence processing can have an important effect on the representations that are learned in different languages.

© 2009 Elsevier Inc. All rights reserved.

Introduction

An adequate theory of sentence production must be able to explain the fact that particular languages have particular biases for ordering words and phrases (Branigan, Pickering, & Tanaka, 2008; Sridhar, 1988). For example, in English, speakers have a tendency to place highly modified “heavy” noun phrases later in the sentence than would ordinarily be expected (*heavy NP shift*, e.g., Ross, 1967). These word order biases are optional, relatively weak, and do not change the grammaticality of the sentence. Therefore, they are not considered to be a part of the core grammatical knowledge of a language and the learning of this behavior is not explained in syntax acquisition theories. Rather, these tendencies have been argued to arise from processing considerations, where one word order is preferred because it is easier to remember, comprehend,

or produce. But since these word order biases can differ across languages – heavy NP shift goes in different directions in English and Japanese – language learning might play an important role in the creation of these biases.

The interaction between learning and processing in these word order biases suggests that an integrated model of syntax acquisition and sentence production might be needed to explain these language-specific biases. This is challenging, because word order knowledge can differ greatly across languages (Baker, 2005; Croft, 2001) and it is not clear how the learner discovers the appropriate settings for these constraints during syntax acquisition (Chang, Lieven, & Tomasello, 2008; Mazuka, 1998). Furthermore, research on sentence processing suggest that speakers make use of representations that change incrementally as they process each word in a sentence (Altmann & Kamide, 1999; Ferreira, 1996), but theories of syntax acquisition are not designed to learn such representations (e.g., Pinker, 1989). To address these issues, a connectionist model of syntax acquisition (Chang, Dell, & Bock, 2006), which can learn representations that support incremental sentence production, was taught to produce

* Address: Dept. of English, Room #215, College of Humanities, Hanyang University, Haengdan-dong 17, Seongdong-gu, Seoul, 133-791, Republic of Korea.

E-mail address: chang.franklin@gmail.com

sentences in English and Japanese. It was then shown to exhibit several word-order phenomena in these languages, crucially phenomena that have puzzled researchers because they differ between these two languages.

The first such phenomenon is heavy NP shift. For example in English, the standard order for a prepositional dative is shown in 1, while the order in 2 is strongly dispreferred or ungrammatical (marked with a *). Example 3, however, shows that the order in 2 is acceptable when the phrase “the book” is lengthened or made heavy by adding a relative clause (“that I bought last week”). Example 4 shows that the order used in 1 is still grammatical, which demonstrates that the shift does not affect grammaticality.

- (1) I gave the book to Mary.
- (2) * I gave to Mary the book.
- (3) I gave to Mary the book that I bought last week.
- (4) I gave the book that I bought last week to Mary.

The preference for placing short phrases before long phrases in English (short-before-long preference) has been attributed to a processing preference that tries to minimize the distance between verbs and their arguments (Gibson, 1998; Hawkins, 1994). This distance minimization account predicts that in languages where verbs appear later in sentences than they do in English, speakers will shift in the opposite direction (long-before-short preference). This has been confirmed for Japanese, a verb-final language (Hawkins, 1994; Yamashita & Chang, 2001). While there is much cross-linguistic evidence supporting this processing account, the exact implementation of the mechanism for minimizing verb–argument distances is still not well understood. In particular, this preference draws its motivation from the dynamics of sentence comprehension and it is not clear that this mechanism would be able to explain heavy NP shift in sentence production.

Another word-order phenomenon is the tendency to place words that are easier to access earlier in sentences (Bock, 1982). Hereafter, I call such effects, lexical/conceptual accessibility effects, or just accessibility effects. For example, speakers are more likely to place the words related to animate entities earlier than inanimate entities. This preference can even induce speakers to use a less frequent syntactic structure, such as the passive, to obtain early placement of accessible words (McDonald, Bock, & Kelly, 1993; e.g., “The students were frightened by the sound” was preferred over “The sound frightened the students”). A wide variety of conceptual and lexical factors related to the accessibility of words have been found to modulate structure selection (e.g., Bock, 1986; Bock, 1987; Bock & Irwin, 1980; Bock & Warren, 1985; Onishi, Murphy, & Bock, 2008; Prat-Sala & Branigan, 2000). However, these conceptual and lexical accessibility phenomena depend on the syntactic context that they occur in. For example, McDonald et al. (1993) found that animacy did influence the order of nouns in transitive sentences, but not the order of nouns in a noun conjunction (there was no preference for either “the crew and the camera suffered minor injuries” or “the camera and the crew suffered minor injuries”). These results have been used to support a particular architecture for sentence production (Bock & Levelt, 1994), where conceptual factors

such as animacy can influence subject and object selection, but not the surface word order. However, this architectural account has been challenged by data from languages with a freer word order like Japanese, where it has been found that conceptual factors can influence the surface position of phrases (Branigan et al., 2008; Tanaka, Branigan, & Pickering, 2008). At present, the implications of these results for a general cross-linguistic theory of sentence production have not been fully worked out.

Although heavy NP shift and lexical/conceptual accessibility effects are evident in sentence production, there is no unified theory that combines the distance minimization account (which focuses on heavy NP shift) with the architectural assumptions of theories of sentence production (which have concerned themselves with the accessibility effects). In this article, an integrated theory of these phenomena is provided in the form of a computational model which can learn syntactic representations for different languages and exhibit the appropriate word order behaviors for each language. This article has six sections. In Section ‘Word ordering phenomena in English and Japanese’, the syntactic features of English and Japanese and the manifestations of the particular word-order phenomena are described. Section ‘The input language of the model’ describes the input languages that the model learns. Section ‘The operation of the Dual-path model’ describes the model’s architecture and how it is trained and tested. Section ‘Heavy NP shift in English and Japanese’ describes the model’s heavy NP shift behavior and the reasons for the shift. Section ‘Accessibility in English and Japanese’ describes the accessibility phenomenon and the factors that influence it. Finally in Section ‘Conclusion’, the implications of the model for theories of sentence processing and acquisition are discussed.

Word ordering phenomena in English and Japanese

English and Japanese syntax differ in several respects. One important difference is the position of the verb, where verbs in English occur early in clauses and verbs in Japanese occur at the end of clauses. Another related difference is the position of the relative clause (heavy NP shift is often tested with utterances that have a relative clause). In English, relative clauses occur after the head noun phrase, while in Japanese, the relative clause appears before the head noun phrase. The English sentence “the boy [that the man saw] ran” would have an order in Japanese like “[man-ga saw] boy-ga ran”. (In this paper, Japanese sentences will be written with their content words replaced with their English equivalents). Japanese is also an argument omission language, where any argument of the verb can be omitted. So in a situation where an English speaker would use pronouns like “She gave it to him”, a Japanese speaker would typically just produce the verb alone, since the referents are inferable from the context. Another difference is that English noun phrases start with articles, while Japanese phrases do not. Noun phrases in Japanese, however, are marked by particles (“wa”, “ga”, “ni”, “o”) that encode syntactic functions (topic, subject, indirect object, and object, respectively). Unlike English, where syntactic functions are typically marked

using word order, Japanese speakers can also *scramble* the order of particle-marked phrases, which allows them a larger number of word orders. For example, the sentence “the dog chased the cat” (dog-ga cat-o chased) can also occur in the order where the cat is scrambled before the dog as in “cat-o dog-ga chased” and this sentence has the same overall meaning since the particles “ga” and “o” mark the relationship of the dog and cat in the act of chasing. It is important to determine how these differences in the syntax of the language might influence word ordering biases like heavy NP shift.

When examining word order biases such as heavy NP shift, it is important to isolate these biases from other factors that can influence word order and to equate the situations where these biases have been found. Although much of the work on heavy NP shift has used corpus data (Hawkins, 1994; Hawkins, 2004; Lohse, Hawkins, & Wasow, 2004; Rosenbach, 2005; Wasow, 2002), it is easier to isolate the minimal conditions for the shift in an experiment. Several studies in English have found evidence for a short-before-long bias in on-line sentence production. In one study (Stallings, MacDonald, & O’Seaghdha, 1998), speakers had to read a set of phrases from a computer screen and formulate a complete sentence using those phrases after a prompt was given. It was found that speakers were more likely to produce heavy NP shifted structures when given an object phrase that was long (e.g., “the manager presented to Jill the new line of bright summer beach and resort fashions”) than one that was short (e.g., “the manager presented to Jill the styles”). In another production study, Arnold, Wasow, Losongco, and Ginstrom (2000) determined that the proportion of double object dative commands (e.g., “give the rabbit the green spotted pencil”) was higher when the theme (“the green spotted pencil”) was longer than the recipient (“the rabbit”) than when the reverse was true (“give the white rabbit the pencil”).

In contrast to the research using English, work in Japanese has found a long-before-short bias, both in corpora and experiments (Hawkins, 1994). Using a task that was similar to that used by Stallings, MacDonald, and O’Seaghdha (1998), Yamashita and Chang (2001) found that Japanese speakers prefer to place heavy noun phrases earlier by scrambling in both transitive and dative sentences. In addition, Hakuta (1981) found a similar long-before-short bias in a delayed repetition task in Japanese children.

One of the best explanations for the cross-linguistic difference in the direction of the shift comes from distance minimization accounts (Gibson, 1998; Hawkins, 1994; Hawkins, 2004). In these accounts, sentence processing involves the linking of a verb and its arguments. Since linking requires resources (e.g., working memory), the system should prefer configurations that minimize the distance between the verb and its arguments. Therefore, if one of the arguments is long, the system will prefer it to be farther away from the verb, so that the other arguments will be closer to the verb, thus minimizing the overall distance between the verb and its arguments. Since verbs come earlier in English, this theory predicts that English speakers should prefer to postpone long elements, while in a verb-final language like Japanese, speakers should prefer to place long elements earlier.

The motivation for the distance minimization mechanism comes from sentence comprehension. To understand a sentence, it is necessary to link the verb and arguments to determine the role that each argument plays in the event. In contrast in sentence production, one already knows what one wants to say and therefore the message encodes the relationship between the verbs and its arguments. One can simply examine the message to check on verb–argument dependencies if and when this is required. Consequently, it is less clear that an extra processing load is incurred during the production of a sentence with the verb and its arguments widely separated. Just considering production leaves heavy NP shift unexplained.

There are several possible ways that distance minimization in comprehension could influence sentence production. One is that production and comprehension share mechanisms such that the verb–argument linking that occurs in comprehension also takes place in production even though the information is already stored in the message. This seems to be the approach taken by Hawkins (2004, pp. 105–111) who argues that distance minimization has benefits for both the speaker and hearer. For example, he claims that a speaker can choose between candidate structures by using the biases of their sentence parser in combination with the meaning. What is left unexplained in this approach is how the language system generated all of the possible candidate structures in the first place. This claim is also at odds with the assumption in language production theories that speakers plan only one utterance at a time (Bock, 1982; Ferreira, 1996; Ferreira & Dell, 2000).

Another possibility is that language users comprehend heavy NP shifted utterances better than non-shifted utterances, and therefore these comprehended structures are also better learned and therefore more strongly represented in the production system. If this were the case, then verbs that lead to a strong expectation for a noun phrase complement but are followed by an element that is not a noun phrase could be quickly analyzed as a heavy NP shifted structure. The ease in recognizing heavy NP shift with these verbs might strengthen the verb–structure association in language learning, and that might lead to greater use of heavy NP shift with these particular verbs in production. But at present there is evidence against this possibility both in production and comprehension. In production, Stallings et al. (1998) found that heavy NP shift was lower with verbs which had high expectations for a noun phrase complement compared to more promiscuous verbs. In comprehension, Staub, Clifton, and Frazier (2006) found that pairings of verbs that were biased towards noun phrase complements (transitive-biased condition) with post-verbal adjuncts or prepositional phrases that signaled heavy NP shift did not lead to immediate slowing due to reanalysis. Therefore, comprehenders were not biased to expect heavy NP shift in these ambiguous situations and verbs which might provide clear signals for heavy NP shift in language learning did not appear more in shifted structures.

A third possible way that comprehension could influence production would be if speakers were consciously or unconsciously adapting their utterances to make them easier for their listeners to understand (Clark, 1996;

Horton & Gerrig, 2005). These accounts, however, would have trouble explaining why speakers produce heavy NP shifted structures in experimental paradigms without explicit listeners (Stallings et al., 1998; Yamashita & Chang, 2001). Also other work in production has suggested that speakers do not always take their listener into account in selecting words or syntactic structures (Brown & Dell, 1987; Ferreira & Dell, 2000). In summary, the various comprehension-based accounts of heavy NP shift do not clearly explain the behavioral data in production, and therefore it is worthwhile to explore whether there is a way to explain the shift using production-based mechanisms alone.

Several researchers have argued that English heavy NP shift is consistent with the properties of an incremental production architecture (De Smedt, 1994; Kempen & Harbusch, 2003; Wasow, 2002), since short light phrases should be easier to access or assemble than longer phrases. But if the short-before-long preference is linked to the architecture of the sentence production system and the architecture is universal, then it is difficult to explain the opposite long-before-short preference that has been found in Japanese. Therefore, production-based accounts of heavy NP shift do not explain the cross-linguistic data as well as distance minimization accounts. The distance minimization account can explain the cross-linguistic data, because its word order preferences are sensitive to learned language-specific features like verb position. If theories of sentence production were also made to be more sensitive to language-specific features, then it might be possible to provide a production account of heavy NP shift in different languages.

In contrast to heavy NP shift, which differs starkly in English and Japanese, the effect of accessibility in the two languages is more similar. In both languages, factors that make words easier to access in isolation (e.g., in naming experiments) also tend to increase the likelihood that these words will appear early in sentences. For example, words that are discourse or lexically given tend to appear earlier in sentences in both English and Japanese (Bock & Irwin, 1980; Ferreira & Yoshita, 2003). Likewise in both languages, animate nouns are more likely to appear earlier in sentences (McDonald et al., 1993; Tanaka et al., 2008). Results like these support the idea that sentence production is an incremental process, in which speakers plan sentences word-by-word and words that are selected earlier can influence the structure being planned (Bock, 1982). If conceptual factors can influence lexical selection and lexical selection determines the word order in sentences, then we would expect that the word order in sentences would always reflect conceptual factors. But there are findings which run contrary to a pure accessibility account. In particular, the fact that animacy does not influence the order in noun conjunctions suggests that properties of the sentence construction (i.e., transitive or conjunction) can modulate the influence of conceptual factors on word order. The modulation of conceptual factors by syntax seems to be necessary, given that different languages map concepts in different ways (Allum & Wheeldon, 2007).

Theories of sentence production have dealt with the variability in the extent to which conceptual accessibility affects word order by appealing to the distinction between

functional and positional levels that was useful for explaining speech errors (Bock & Levelt, 1994; Garrett, 1988). In these theories, the functional level was a stage in processing where concepts in the message were assigned to syntactic functions (e.g., subject or direct object). In the subsequent positional level, the phrases associated with these syntactic functions were then assigned to surface sentence positions. Since the functional level was directly linked to the message, the competition for syntactic function assignment naturally depended on conceptual factors like animacy or imageability. Since the positional level was not directly linked to the message, it was assumed to be insensitive to conceptual factors. Since the two nouns in a conjunction necessarily share the same syntactic function, they would be similarly represented at the functional level and therefore their relative order would be determined at the positional level.

Although the functional/positional distinction provided a good account of how animacy or imageability influenced word order in the production of English sentences, other studies have found that message-related factors like referential availability and typicality can influence the order of elements in a conjunction (Bock & Irwin, 1980; Kelly, Bock, & Keil, 1986; Onishi et al., 2008). Furthermore, conceptual factors like animacy and referential availability can influence the order of nouns in a conjunction when it is presented in isolation both in adults and children (Byrne & Davidson, 1985; McDonald et al., 1993; Narasimhan & Dimroth, 2008). Another challenge for the theory comes from languages that allow scrambling (Branigan et al., 2008). Scrambling shifts the order of noun phrases without changing their syntactic functions (typically marked with a lexical element like a particle). If conceptual factors do not influence the positional level order of phrases, they should not influence the shifts due to scrambling of noun phrases. But in several languages, it has been found that conceptual factors like animacy and referential availability can influence the order of scrambled noun phrases (Branigan & Fel-eki, 1999; Ferreira & Yoshita, 2003; Kempen & Harbusch, 2004; Prat-Sala & Branigan, 2000; Tanaka et al., 2008). And although animacy can influence the order of scrambled noun phrases in languages like Japanese, it does not influence the order of nouns in conjunctions (Tanaka et al., 2008). So the barrier between aspects of word order that are sensitive to conceptual factors and those that are insensitive to these factors varies with the language, the type of conceptual factor, the sentence construction that it appears in, and the task (e.g., whole sentences vs. isolated phrases). It is hard to explain this variation within a rigid universal language architecture, but it maybe possible to explain this variation if construction-specific and language-specific learning is allowed to change the sentence production system.

Finally, the distance minimization theory and the functional/positional theories are suited to explain the data that they were designed to explain, but it is not clear how these theories are integrated in sentence production. The distance minimization theory seems to require some planning of multiple whole sentences to generate distance measurements, and that seems to be at odds with incremental planning in the functional/positional theory. The

functional/positional theory does not vary with verb position, and hence it would have trouble explaining the different direction of heavy NP shift in English and Japanese. Therefore, there is a need to bridge between these two accounts of word order.

Integrating these phenomena into a single theory seems to require a computational model, which can insure that the theory is internally consistent and has coverage of the data. There are several candidate computational models that might address aspects of the data, but each has deficiencies. To start with, there are incremental models of sentence production for English, but these systems have not been adapted for Japanese production (Dijkstra & De Smedt, 1996; Kempen & Hoenkamp, 1987). Dominey, Hoen, and Inui (2006) presented a model of English and Japanese syntax acquisition that mapped between meaning and word sequences, but their system planned the entire sentence at once, so it cannot account for incremental planning phenomena that we are concerned with here. There is also a great deal of research on natural language generation that is computationally rather than psycholinguistically-motivated. Hence these systems are often not incremental, do not generate from a conceptual representation, or cannot generate Japanese and English with a similar degree of accuracy (e.g., Aikawa, Melero, Schwartz, & Wu, 2001; Langkilde-Geary, 2002).

One model that satisfies our requirements is the Dual-path model (Chang, 2002; Chang et al., 2006), since it can learn to use meaning to produce sentences for different languages in an incremental manner. This current paper introduces a new version of this model and applies it to heavy NP shift and accessibility effects in English and Japanese. The previous versions of the model were only trained to produce single clause structures, but since heavy NP shift often involves multi-clausal stimuli, the model's input was extended so that it included messages with two propositions (see also Fitz, 2009; Fitz & Chang, 2008). These messages were associated with English and Japanese utterances, and two models, one for English and one for Japanese, were trained on the corresponding message–sentence pairs. These models were then tested for heavy NP shift and accessibility behavior in English and Japanese. Finally, the input language of the models was varied in order to determine which features of the input were important for their behavior.

The input language of the model

Children learn language by hearing utterances produced by adults in situations where they can infer something about the meaning that the adult is referring to (Pinker, 1984; Tomasello, 2003). Likewise, the model acquires its language representations from its “parental” input consisting of utterances paired with its meaning or message. To control the nature of this input language, it was generated artificially by an input language generator that is separate from the model proper. The input language was a simple version of English and Japanese that would allow the testing of the sentence structures that are important in heavy NP shift and conceptual accessibility.

The input was a set of message–sentence pairs. There were messages with one argument (intransitives, “the tomato is rolling”), two arguments (transitives, “the dog chases the cat”), and three arguments (datives, “the man gave the woman the apple”). The message was composed of role and concept pairs that encode relationship between the concepts in the event (e.g., who is doing what to whom). There were three roles in the model labeled X, Y, and Z. The Y role encoded the argument that was most affected in an event. This included intransitive subjects as well as transitive and dative patient/themes. The X role encoded transitive and dative agents, and the Z role encoded recipients. The action was encoded in its own role A. The roles were bound to concepts in a particular message as can be seen in the following intransitive message where the theme role Y is linked to the concept for “tomato” (Y = TOMATO).

Message: A=ROLL Y=TOMATO,DEFINITE E=PRESENT, PROGRESSIVE,AA,YY

English: the tomato is roll -ing

Japanese: tomato -wa roll -te iru

In addition to the role–concept bindings, there was a part of the message called the event-semantics (E=PRESENT,PROGRESSIVE,AA,YY), which encoded tense (PRESENT or PAST), aspect (SIMPLE or PROGRESSIVE), as well as information about the number and prominence of the arguments in the message (AA, YY). Noun phrases could be definite (DEFINITE, e.g., the tomato), indefinite (INDEFINITE, e.g., a tomato), or pronominal (PRONOUN, e.g., it). In English, pronominal noun phrases were replaced with case-marked pronouns (he, she, it, him, her), while in Japanese, these noun phrases were omitted to implement the argument omission feature of the language. More detail about this message representation can be found in the “Appendix” as well as in Chang (2002) and Chang et al. (2006).

When conveying a particular meaning, there are often alternative structures that can be used such as the active “the dog chased the cat” and the passive “the cat was chased by the dog”. This ability requires that the message provides information which can be used to delimit the set of structures that can be used to express that message. In the model's message representation, these abilities were provided by a set of prominence units in the event-semantics (XX, YY, ZZ, AA). If a prominence unit was activated above zero, then the corresponding role unit was non-empty (e.g., the prominence unit YY in the above example signals that the Y role is connected to a concept). Therefore, the number of active prominence units signaled the transitivity of the event (e.g., if prominence units XX and YY are activated, then the event is a transitive event). When messages with these prominence units were paired with particular language-specific structures, the model could learn which structures were appropriate for that set of prominence units.

In addition, it was assumed that the activation of the prominence units in messages correlated with the structures used in syntactic alternations. This was done by lowering the activation of the prominence unit for roles that appeared later in sentence structures. For example, a transitive event could be described with an active or a passive

structure. Since actives are generally more frequent than passives, it was assumed that actives were associated with the default prominence unit activations of 1.0 ($XX = 1.0$ and $YY = 1.0$). Then passives were associated with messages where the agent X role was less prominent and this was done by setting the prominence unit XX to 0.6. These variations in prominence are one way to implement discourse/referential availability effects in the literature (Bock & Irwin, 1980; Ferreira & Yoshita, 2003; Prat-Sala & Branigan, 2000). In addition, this approach draws support from work that suggests that anticipatory eye-movements by a speaker are correlated with the order of referents in a speaker's utterance (Bock, Irwin, Davidson, & Levelt, 2003; Griffin & Bock, 2000). If we assume that speakers look at referents to enhance lexical access, then they must have some non-lexical representation for controlling their anticipatory eye-movements and the prominence representation in the event-semantics is one way to model this knowledge.

While we assume that these variations in prominence are correlated with particular structures, these variations do not determine the nature of the structures themselves. For example in English and Japanese, variation in the prominence of arguments is expressed with different syntactic means. While English speakers can alternate between actives and passives to emphasize particular arguments, Japanese speakers often express these same meanings using scrambling. Therefore in the Japanese language, scrambling was used to express differences in prominence (e.g., $XX = 1.0$ $YY = 1.0 \rightarrow$ “dog-ga cat-o chased”; $XX = 0.6$ $YY = 1.0 \rightarrow$ “cat-o dog-ga chased”).

The event-semantics was also important for representing the meaning of utterances with more complex syntactic structures. In the model's input language, there were complex utterances like sentential conjunctions (“the orange rolled and the tomato bounced”), phrasal conjunctions (“the orange and the tomato bounced”), and utterances with a relative clause (“the man gave the orange to the girl that the boy saw”). Utterances with a relative clause can be paraphrased by a sentential conjunction (e.g., the above example can be paraphrased as “the man gave the orange to the girl and the boy saw the girl”). To encode this paraphrasability, utterances with a relative clause were set to have the same two-proposition message as their sentential conjunction counterpart. These two-proposition messages were like the single proposition messages that we presented earlier, but the roles and event-semantics in each proposition were distinguished by placing the number 0 or 1 before the role letter (e.g., main clause agent 0X; relative clause patient 1Y). To represent the difference between the relative-clause utterance and its sentence conjunction counterpart, a special set of event-semantic units called binding nodes were used. These nodes represented the fact that the two clauses in a relative-clause utterance have a role that is shared between the two. For example in the relative-clause utterance above, the girl is the recipient of the giving event (0Z) and at the same time, the patient of the seeing event (1Y). Therefore, the event-semantic binding node 0Z1Y would be activated to signal this binding. When the model heard input message-sentence pairs like this, it could learn that this binding node was associated with a language-specific relative clause structure.

Phrasal conjunctions appear to be single proposition utterances, but they can also be paraphrased by a sentential conjunction (e.g., “the orange and the tomato bounced” can be paraphrased as “The orange bounced and the tomato bounced”) and even young children are able to alternate between the two (Hakuta, de Villiers, & Tager-Flusberg, 1982). Therefore, in the same way as relative clause utterances, phrasal conjunctions were set to have the same two propositions as their sentential conjunction paraphrase, but with extra binding nodes to denote the overlapping meaning. In above example, the action was the same for both propositions (e.g., 0A = BOUNCE, 1A = BOUNCE), so a binding node 0A1A signaled this overlap. To summarize, complex constructions in the model were all two-proposition messages with different degrees of binding between the two propositions (see “Appendix” for more details).

One goal of this project was to distill the key features of natural language that might be important in word order generation. Therefore, the English and Japanese input languages were made as simple as possible, while still incorporating the most frequent features of each language. Even though the input languages were simpler than real languages, the languages were still very expressive. For example, the English sentence “a boy gave a cake to a girl” could be rendered in Japanese as “boy -ga girl -ni cake -o give -ta” or “boy -ga cake -o girl -ni give -ta”. Since Japanese is an argument omission language, marking the various arguments as pronouns would lead to a large number of utterances (i.e., “girl -ni cake -o give -ta”, “boy -ga cake -o give -ta”, “boy -ga girl -ni give -ta”, “cake -o girl -ni give -ta”, “boy -ga give -ta”, “girl -ni give -ta”, “cake -o give -ta”, “give -ta”). Also, if the first noun phrase was definite, then it would be treated as a topic and marked with “wa” (i.e., “boy -wa girl -ni cake -o give -ta”, “boy -wa cake -o girl -ni give -ta”, “girl -wa cake -o give -ta”, “boy -wa cake -o give -ta”, “boy -wa girl -ni give -ta”, “cake -wa girl -ni give -ta”, “boy -wa give -ta”, “girl -wa give -ta”, “cake -wa give -ta”). These variations were present for each of the constructions in the model's input language.

The input language was generated by a program called the input generator. The input generator was used to create language-independent messages and then it produced English and Japanese utterances from those messages. It is important to distinguish the input generator from the sentence production model that will be described in the next section. The input generator was made up of hand-crafted rules for each language, while the model learned its own rules from the input that it was given. The input generator, as a computational device, does something like whole sentence planning, while the model generates sentences incrementally word-by-word. Therefore while the input generator can create utterances from messages, it does not explain how language learning leads to a system that can do incremental sentence production.

The operation of the Dual-path model

Language acquisition is the process where the external input language comes to be instantiated within the brain's neural network and the architecture of this network helps to constrain the representations that are learned. The Dual-

path model (Chang, 2002) implemented a hypothesis about the gross configuration of the neural networks that support syntax in acquisition and production. The model had two important features. One feature was its architecture, which was made up of two pathways or systems: the sequencing system and the meaning system. The sequencing system was a simple recurrent network (Elman, 1990) which learns distributional regularities over word classes, insuring that the sequence of words generated conforms to the syntax of the language being produced. The meaning system encodes the message, constraining the word sequence to reflect the intended meaning. This Dual-path architecture allowed the model to learn separate statistical regularities that could be combined in novel ways and this was the basis of its ability to generalize, that is, to be able to produce novel utterances. A second important feature of the model was its variable-based message. Rather than encoding the message as an activation pattern over neural units (Chang, Dell, Bock, & Griffin, 2000; Dell, Chang, & Griffin, 1999; St. John & McClelland, 1990), the model instantiated the message as a special set of links or weights between roles (e.g., agent) and concepts (e.g., DOG) on analogy with the fast binding that is thought to be needed for binding objects to locations in spatial processing. This variable-based message enhanced the model's ability to generate words in surface positions that were not trained and this allowed the model to consistently acquire an abstract syntax from its limited and idiosyncratic input. The motivations for these assumptions were detailed in Chang (2002) and Chang et al. (2006), and therefore I will concentrate here on describing the aspects of the architecture that are important for the word-order phenomena under consideration.

First, I will start with an example of how the model would produce a sentence given its message. I will use the sentence/message pair that we saw earlier (i.e., sentence “the tomato is roll -ing”; message: A=ROLL Y=TOMATO,DEFINITE E=PRESENT,PROGRESSIVE,AA,YY). In Fig. 1, the meaning and sequencing systems are represented by gray boxes, but the full architecture within these boxes is not shown (particular layers are shown as white boxes). The message was placed in the model by setting the weights between the units in the roles layer (action A, theme Y) and the concepts layer (ROLL, TOMATO, DEFINITE) to a high value. These settings were temporary (dashed lines) and were present only while the model was trying to produce this particular message. The concept units were

linked to word units through other weights that were learned during acquisition (learned weights are shown with black arrows). Setting the activations of event-semantics units PRESENT, PROGRESSIVE, AA, and YY to 1 completes the message placement process. The sequencing system can directly influence word selection, but it can also activate roles in the meaning system. The sequencing system and meaning system received the previously produced word in the word layer (light gray arrows at bottom of Fig. 1), and this information could be used in incremental sentence planning.

After the message is set, the model starts to produce a sentence. Since the event-semantics has activated the action prominence node AA and theme prominence node YY, the sequencing system expects an intransitive structure, and hence it activates the theme Y role in the role layer. The Y role activates the concept TOMATO and DEFINITE, and that causes the words “tomato” and “the” to be activated in the word layer. The sequencing system has recorded the distributional regularity that sentences in English tend to start with articles like “the” and “a”. Since “the” is supported by activation from the concept DEFINITE and distributional regularity in the sequencing system, it is the most activated word and it is produced. Now that the first word has been produced, that word is passed back into the sequencing and meaning systems. Since in general, “the” is followed by nouns, the sequencing system now expects a noun and words such as “tomato” and “apple” should be activated. Meanwhile, the Y role is still activated, because its content has not been fully produced, and therefore the concepts TOMATO and DEFINITE are still activated. The combined influence of the meaning and sequencing systems causes “tomato” to be produced. Now this word is passed back to the input layer and that causes the model to deactivate the Y role, because its content has been produced. Next, the sequencing system activates the word “is”, because it has learned that when the PRESENT and PROGRESSIVE nodes are activated and the subject has been produced, a present tense auxiliary should then be produced. After “is” is passed as input for the sequencing system, the model then, because of its experience with English, knows to retrieve the verb by activating the action A role. That activates the ROLL concept causing the verb “roll” to be produced. Finally, the model uses its knowledge that the verb has been produced and the PROGRESSIVE event-semantic feature to activate the morpheme “-ing”. Since all of the message content has been produced,

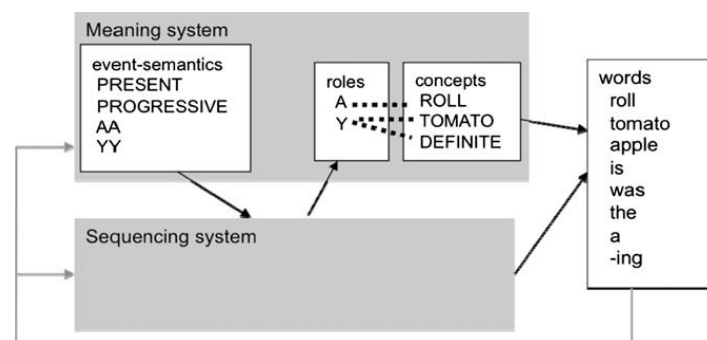


Fig. 1. Dual-path architecture with message for “the tomato is roll -ing”.

the model knows that it is time to end the utterance. This example illustrates how different kinds of information influence the model's behavior at each point in a sentence. Learning is crucial to setting up this system of constraints such that the right information is involved at each position in each construction.

The Dual-path architecture has several features that are important in its word order behavior. The first feature is that it plans sentences incrementally, and hence lexical selection plays an important role of the model's structural decisions. Therefore, meaning and syntax are both just constraints on word selection and the lexical competition determines how these factors influence sentence generation (Dell, 1986; Gordon & Dell, 2003). Syntactic knowledge in the model is represented in a simple recurrent network that maps from the previous word to the next word in a sequence. Incremental planning is made possible by another feature of the model that is not shown in Fig. 1, a reverse message that maps from concepts to roles. This reverse message helps the model to map from a word that is selected (and then fed back to the model as input) to its role in a particular message and is used to keep that role from being produced again (deactivating the past; Dell, Burger, & Svec, 1997). For example, when people are given a picture of lightning striking a church, they can describe it with an active sentence "the lightning is striking a church" or a passive "the church is struck by lightning". This structural choice can be influenced by making the word "lightning" or "church" more likely to be produced first (Bock, 1986). The Dual-path model implements this behavior by using the reverse message. If the model happens to start the sentence with the word "church", this will be fed back into the reverse message to activate the concept CHURCH and the patient Y role that it is linked with in this particular message. An English model will have learned that when a message has two arguments and starts with the Y role, it should then employ the passive structure. In a Japanese model, the same situation will require that the model use a scrambled structure where the object argument precedes the subject argument.

The assumptions of the Dual-path model were found to be useful in creating a model of incremental sentence production that could learn abstract syntactic representations using neural network learning and processing mechanisms. These assumptions are supported by the fact that the model can explain behavior in several different domains of psycholinguistics. For example, it can explain double dissociations in aphasia (Chang, 2002; Gordon & Dell, 2003) and the relationship between preferential looking and elicited production in syntactic development (Chang et al., 2006). The statistics in the Dual-path model were found to be better than several other common statistical models in predicting the order of words in utterances in 12 typologically-different languages (Chang et al., 2008). The model was also used to test the idea that structural priming was a type of implicit learning (Bock & Griffin, 2000; Chang et al., 2000) and it was able to account for 14 structural priming results in adults and children (Chang et al., 2006). The assumptions in the model led to several predictions about the nature of the mechanisms that support structural priming and evidence in support of those predictions has been accumulating (Bencini & Valian, 2008; Ferreira, Bock, Wilson, & Cohen, 2008;

Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008; Huttenlocher, Vasilyeva, & Shimpi, 2004; Jaeger & Snider, 2008; Kaschak & Borreggine, 2008; Konopka & Bock, 2009; Savage, Lieven, Theakston, & Tomasello, 2003; Thothathiri & Snedeker, 2008). However, since the model was developed with English psycholinguistic data in mind, it is still an open question whether these assumptions are appropriate for other languages. Therefore by testing the model's behavior in Japanese, we are checking whether the model's assumptions are appropriate for learning and processing syntax in a typologically-different language.

Learning English and Japanese in the model

People can differ because of their input experience and the same is true for models that learn their representations from variable input. To insure that results were replicable, multiple model subjects were created and differences between test conditions were tested with inferential statistics. To create the model subjects, 10 different random seeds were used to create 10 different sets of 40,000 message-sentence training pairs for each language. Each of the 10 English and 10 Japanese models was trained for two passes through its own training set for a total of 80,000 patterns, with weights being updated after each message-sentence pair. The model was tested after every 5000 training patterns. For each message in the test set, the model attempted to predict the utterance that was paired with that message, and two measures of accuracy were used (see "Appendix" for details). The first measure was grammaticality, which involved labeling the utterance that the model produced for syntactic categories, and then checking whether that sequence was a valid sequence for the language. The second measure was message equivalence, which involved transforming both the target utterance and the utterance that the model produced into order-invariant messages, and then comparing whether the same message was achieved.

Fig. 2 shows the average message and grammatical accuracy for the models when it was trained on English and Japanese as a function of the number of training patterns up to 80,000. For this and the other analyses in the paper, the data from the model at the end of training (80,000 patterns) was used to represent the model's adult behavior. The adult grammaticality accuracy on the test set was 93% in English and 95% in Japanese, and message-equivalence accuracy was 91% in English and 93% in Japanese. Errors in the model's output arose from a heterogeneous set of utterance changes. Some of these "errors" would be unremarkable in normal conversation, such as repeating a word (e.g., "the uncle throw -ed the the kite") or omitting an argument in English (e.g., the model expects all English dative verbs to have two arguments, so "he was throw -ing it" would be considered ungrammatical). Higher accuracy could be achieved by training with larger input sets, but for the present purposes, these results were considered sufficient.

Heavy NP shift in English and Japanese

To examine heavy NP shift in the model, test items whose phrases varied in weight were created. The test

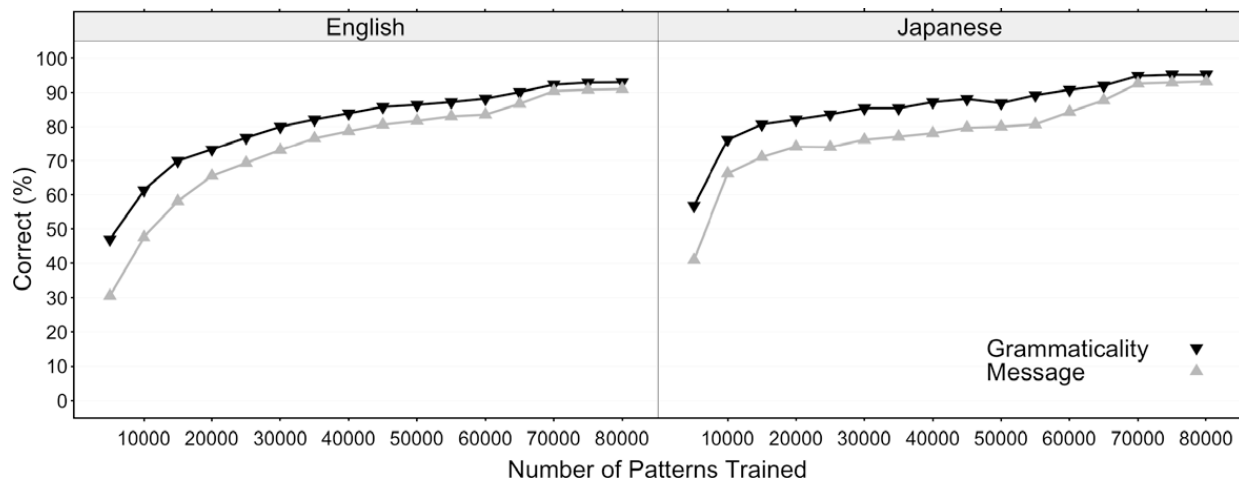


Fig. 2. Accuracy in grammaticality and message equivalence on the test set after every 5000 training patterns.

items were datives with three simple noun phrase heads (indefinite articles and no adjectives). Three conditions were created: Long Recipient, Long Patient, and All Light. The Long Recipient condition had a relative clause attached to the recipient (OZ) and it had a binding node to signal the link between the main clause and relative clause role (e.g., in the example below, the relative clause patient 1Y is linked, and the binding node OZ1Y is on). The Long Patient condition had a relative clause attached to the patient (OY) and the corresponding binding node (e.g., OY1Y). The All Light condition had no relative clauses.

Long Recipient Message:

OA=GIVE OX=MAN,INDEFINITE OY=TELEPHONE,INDEFINITE
OZ=GIRL,INDEFINITE OE=PRESENT,
PROGRESSIVE,AA,XX,ZZ,YY,OZ1Y
1A=TOUCH 1X=DOG,INDEFINITE 1Y=GIRL,INDEFINITE
1E=PAST,SIMPLE,AA,XX,YY

English: a man is give -ing a girl that a dog touch -ed a telephone.

Japanese: man -ga dog -ga touch -ta girl -ni telephone -o give -te iru.

Long Patient Message:

OA=GIVE OX=MAN,INDEFINITE OY=TELEPHONE,INDEFINITE
OZ=GIRL,INDEFINITE
OE=PRESENT,PROGRESSIVE,AA,XX,ZZ,YY,OY1Y
1A=TOUCH 1X=DOG,INDEFINITE
1Y=TELEPHONE,INDEFINITE 1E=PAST,SIMPLE,AA,XX,YY

English: a man is give -ing a girl a telephone that a dog touch -ed.

Japanese: man -ga girl -ni dog -ga touch -ta telephone -o give -te iru.

All Light Message:

OA=GIVE OX=MAN,INDEFINITE OY=TELEPHONE,INDEFINITE
OZ=GIRL,INDEFINITE
OE=PRESENT,PROGRESSIVE,AA,XX,ZZ,YY

English: a man is give -ing a girl a telephone.

Japanese: man -ga girl -ni telephone -o give -te iru.

For each condition, 100 examples were randomly generated. Half of the test items had prominence information in the message that was biased towards the recipient before patient form, and half had prominence information that was biased towards the reverse order. Therefore, if

weight does not influence word order, the model should produce both double object and prepositional datives 50% of the time. The model was tested with this test set after every 5000 patterns during training and the utterances that the model produced were coded for whether they used the recipient before patient order and whether the utterance produced matched the target message. The average percentage of the Recipient-before-Patient order out of all message-appropriate utterances is graphed in Fig. 3 for the three conditions. Mixed-effect models were fitted to the empirical logit of the Recipient-before-Patient order as the dependent measure with language (English, Japanese) and weight condition (Long Recipient, Long Patient) as predictor variables and model subject as a random effect (Baayen, Davidson, & Bates, 2008). Mixed effect models allowed us to use multiple different random effects such as model subject and number of training patterns. The empirical logit mapped the dependent measure into the range of the real numbers, but kept the values away from extreme values (Agresti, 2002). Markov chain Monte Carlo sampling was used to compute *p*-values (Baayen, 2008).

For the adult models, a significant interaction between language and weight condition was found ($t = 5.5$, $p < 0.001$). This interaction was due to the Japanese model's stronger preference for the recipient before patient order when the recipient was long (long-before-short bias, difference = 15.4%, $t = 3.8$, $p = 0.001$) and the English model's preference for this order when the patient was long (short-before-long bias, difference = 15.9%, $t = 4.1$, $p < 0.001$). This result suggests that the cross-linguistic difference in heavy NP shift in English and Japanese can be simulated in a system that does not implement an explicit distance-minimization structure-selection algorithm.

In both languages, the appropriate language-specific heavy NP bias is evident throughout development (Fig. 3). To test if the pattern was stable over development, a mixed effect regression including number of training patterns and model subject as random factors was applied to the data collected at each point in development (every 5000 patterns). The weight * language interaction was significant over the whole of development ($t = 17.2$, $p < 0.0001$). Therefore, the model's heavy NP shift bias is

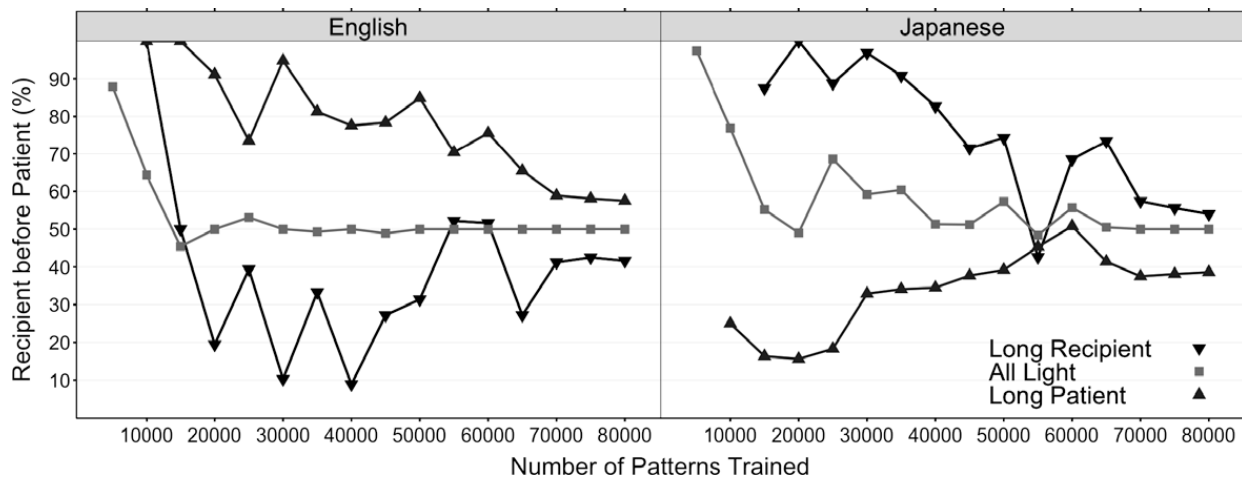


Fig. 3. Weight-based differences in dative production in English and Japanese.

not something that is learned late in syntax acquisition, but rather it is a feature of production behavior throughout development. Experimental evidence supports this result, as language-appropriate heavy NP shift biases have also been found in 2–5 year old English and 3–6 year old Japanese children (de Marneffe et al., 2007; Hakuta, 1981).

To better understand how heavy NP shift is created in the model, the model's internal representations were analyzed at the point where the alternative structures diverge. Previous work with the Dual-path model has suggested that the role activations are strongly related to the structure chosen at the choice point (Chang, 2002). In the English dative sentence “a man is give -ing a girl that a dog touch -ed a telephone”, the choice point is at the point where “girl” is produced, because if the word “telephone” is produced instead, then the model must produce a short-before-long prepositional dative “a man is give -ing a telephone to a girl that a dog touch -ed”. Since these words are activated by their corresponding concepts which are linked up to their thematic roles, the choice between these two orders depends on the activation of the main clause patient OY and recipient OZ roles at the post-verbal position. In Japanese, embedded clauses come before their heads and therefore a different set of roles enter into the choice point competition. For example in the Japanese test sentence “man -ga girl -ni dog -ga touch -ta telephone -o give -te iru”, the choice point is at the word “girl”, because if embedded agent “dog” is produced there instead, a scrambled long-before-short utterance is produced “man -ga dog -ga touch -ta telephone -o girl -ni give -te iru”. Therefore, the choice between the two word orders depends on the winner of the competition between the main clause role (patient OY or recipient OZ) and the embedded clause role (action 1A or agent 1X). To refer to the role that determines the choice between the two word orders, we will talk about the patient choice role or the recipient choice role.

The first step in tracing the model behavior would be to see if the relative activation of the words at the choice point is in fact related to the relative activation of the choice point roles. To index the word chosen, the difference in the activation of the words associated with the pa-

tient and recipient choice point roles (*worddiff*) was computed (if the *worddiff* is positive, then the word for the patient choice point role was selected). Likewise, the difference in the activation of the choice point roles was also computed to create the *rolediff*. A linear regression was used to predict the *worddiff* based on the *rolediff* and a strong relationship was found ($t(3986) = 120$, $p < 0.001$, $R^2 = 0.78$). Therefore in both languages, structure selection during heavy NP shift is due to differences in the activation of roles at the choice point.

These results show that the model has internal representations that generate the language-dependent heavy NP shift preferences. Since the heavy NP shift conditions differ mainly in the event-semantics information that marks the role of the long phrase, the roles difference should be due to the influence from the event-semantics. This influence is difficult to trace, because its effect is passed through a hidden layer where a variety of other influences are combined. Therefore to trace this influence, the activation of the event-semantics units that were important for the long phrase (e.g., binding nodes, prominence units) was traced as it spread along the weights through the hidden units until they reached the patient and recipient choice point role units. This tracing is akin to the normal process of spreading activation, except that the logistic function was not applied at the hidden layer, because the inputs were only a subset of the normal inputs to the hidden units and hence the application of the logistic function would distort the output. The weighted input to the patient and recipient choice point roles was computed separately and then the difference between the two was computed to create the *tracediff*. The *tracediff* was positively correlated with the *rolediff* at the choice point for each utterance in each model ($t(3986) = 11.6$, $p < 0.001$), but only a small portion of the variance is explained ($R^2 = 0.03$). This is to be expected, as heavy NP shift involves a small shift in structural preferences and therefore the parts of the event-semantics that the *tracediff* indexes should only have a small influence on the role activations. However, if we look at the correlations separated by language, the correlations between *rolediff* and *tracediff* are stronger in Japanese than in English (Japanese $R^2 = 0.175$;

English $R^2 = 0.007$). This suggests that the event-semantics plays a greater role in Japanese heavy NP shift behavior than in English. Also, since so little variance is explained by the event-semantics, English heavy NP shift must be due to other factors, such as the changing representations in the hidden layer due to the syntactic context.

The key issue then is why does the model learn these biases. Since long noun phrases use a heterogeneous set of structures and these phrases occurred equally often in different sentence positions, these biases could not be learned directly from the frequency of whole structures in the input. However, word-prediction models like a simple recurrent network often learn statistical regularities that are smaller than a whole sentence. These kinds of sub-sequence regularities could be useful in explaining English heavy NP shift, because the short-before-long order (e.g., “give the book to the man that he saw last week”) has more structural overlap with simple main clause sentences (e.g., “give the book to the man”) than the long-before-short order (e.g., “give the book that he saw last week to the man”). To examine whether the relative frequency of simple main clauses to complex clauses can influence the model's heavy NP shift behavior, a new set of models were trained on a language where the relative proportion of complex dative clauses was increased. In Fig. 4, the data from this new Frequent Complex model is shown next to the data from the Standard model. There is a three way interaction between language, weight, and model, $t = 4.0$, $p < 0.001$. The English models no longer show the short-before-long preference ($t = 1.3$, $p = 0.19$), while the Japanese model still has a strong long-before-short preference ($t = 3.6$, $p = 0.002$). Since it is possible to remove the English heavy NP shift bias by making complex structures unnaturally frequent, the production of these complex structures in the Standard model may depend on “help” from other similar structures, such as main clause structures. But since that same manipulation in Japanese does not remove the similar magnitude of heavy NP shift there, it would seem that the Japanese shift depends less on main clause similarity. One possible reason for this difference is that English heavy NP shift in datives occurs after the verb and this position provides a good cue for recording structural similarity. In contrast, Japanese heavy NP shift occurs in various contexts before the verb (sometimes after the subject, topic, or at the

beginning of the sentence when an argument is omitted), and this can make it harder to find structural similarities between relative clause structures and main clause structures.

Another way to examine how the model learns its heavy NP shift biases is to examine the model's behavior with languages that are intermediate between English and Japanese. By looking at which features are critical for the language-specific direction of the shift, we can better understand how learning produces these biases. To examine this question, the English language was gradually changed into a Japanese-like language by varying verb position, articles, and relative clause position. The simplest set of changes modified a single property. Verb-final English was like normal English, except verbs appeared at the end of clauses (e.g., “The man the book to the girl is give -ing”). No-articles English had all articles removed (e.g., “Man is give -ing book to girl”). RC-before English placed relative clauses before their heads (the normal English “The man is give -ing the book to the girl that he saw” would be expressed as “The man is give -ing the book to the girl that he saw the girl”; notice that “that” was also omitted). Argument omission was not co-varied with these other factors, because preliminary testing found that adding argument omission led to models that were highly accurate at producing the structure specified in the message and therefore these models were not sensitive to other factors like weight. In addition to these three single-change languages, combinations of these single feature changes resulted in three double-change languages and one triple-change language, the most Japanese-like language, a combination Verb-Final, No Articles, and Relative Clause Before, which had sentences like “man book to he saw girl is give -ing”). This Verb-Final, No Articles, and Relative Clause Before language still differed from the model's Japanese language in that it did not have case markers, scrambling, argument omission, and appropriate verb formation rules.

The use of Recipient-before-Patient order for Long Recipient and Long Patient conditions was averaged over the whole of development for each type of language (Fig. 5). All of the data was use here, because the heavy NP shift differences were larger and the influence of each of the language factors was clearer. The bottom left corner

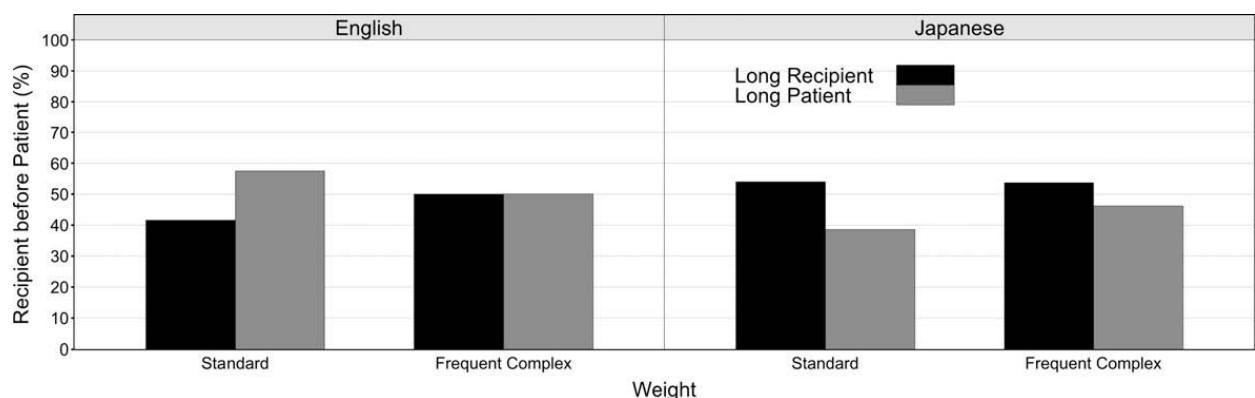


Fig. 4. Comparison of heavy NP shift in Standard and Frequent Complex models.

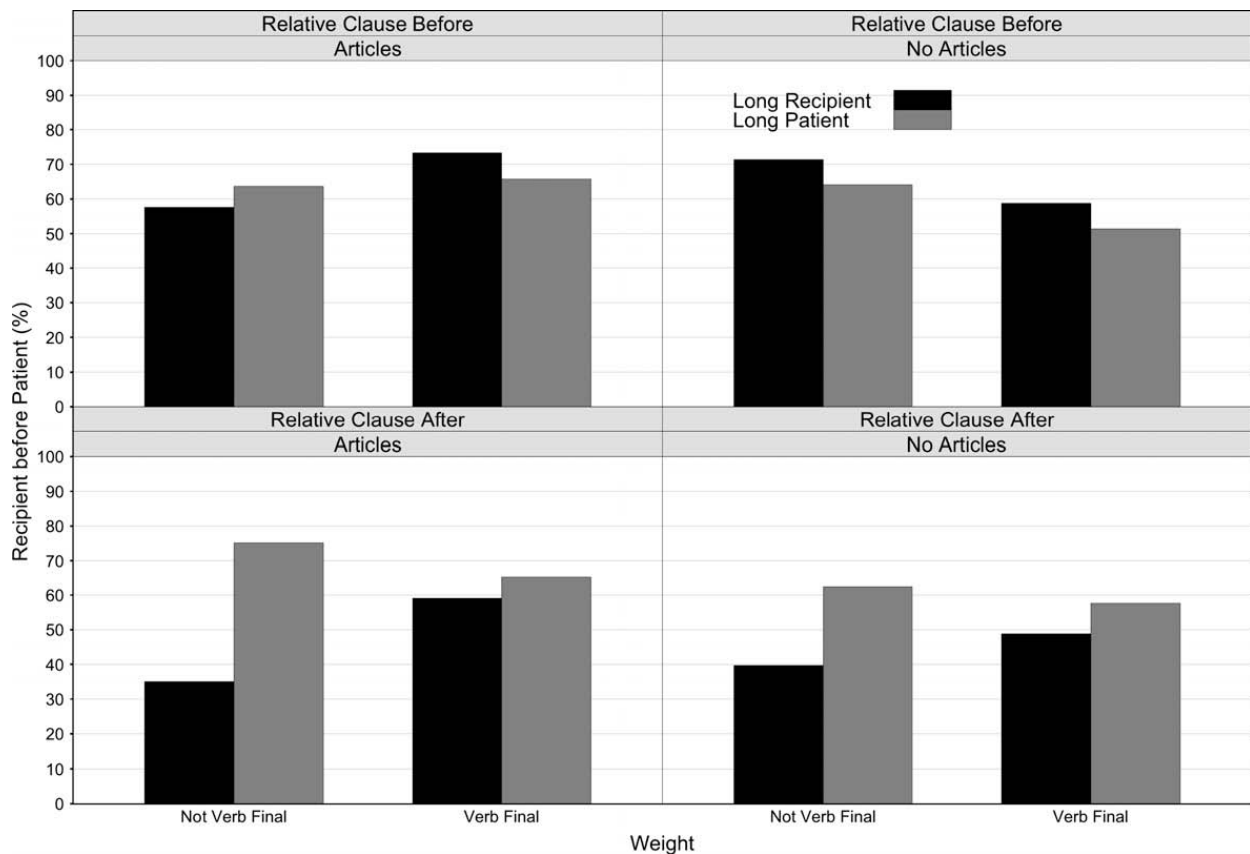


Fig. 5. Weight-based differences in languages that vary in verb position, articles, and relative clause position.

(Relative Clause After, Articles, Not Verb-final) represents the standard English results with a greater preference for Recipient-before-Patient order when the patient was long. The top right corner represents the language that is closest to Japanese (Relative Clause Before, No Articles, Verb-final) and it has the opposite preference (Long Recipient is preferred in the Recipient-before-Patient order). A mixed effects regression yielded strong interactions between weight and each of these features (verb position * weight, $t = 8.4$, $p < 0.001$; articles * weight, $t = 3.9$, $p < 0.001$; relative clause position * weight, $t = 8.7$, $p < 0.001$) as well as several higher order interactions (verb position * articles * weight, $t = 4.2$, $p < 0.001$; verb position * relative clause position * weight, $t = 4.7$, $p < 0.001$; verb position * articles * relative clause position * weight, $t = 2.1$, $p = 0.04$). These results suggest that each of these three syntactic features is important in creating the language-specific directions of heavy NP shift.

This analysis allows us to formulate an account of the model's heavy NP shift behavior in each language. The short-before-long bias in English was due to the similarity of the short-before-long structures to main clause structures and the fact that in English, the choice point occurred later in sentences in a fixed position that facilitated the recording of structural information. This position has good structural cues such as the verb and articles, and the choice between structures involves the main clause roles (since relative clauses go after their heads). The long-before-short bias in Japanese was due to strong links between the

event-semantics for long phrases and the roles. These links were strong, because the same role units were used for embedded clauses on different main clause heads, and therefore all embedded clause inputs helped to train the same set of weights. In addition, since the choice point in Japanese used the roles of the long phrase and was in an early position where it was difficult to record structural similarity, the model learned to use event-semantics more for these structural choices in Japanese. Therefore, unlike obligatory syntactic features like verb position, heavy NP shift is an optional weak bias because it is an emergent effect of the differences in the type of information that are helpful in sentence planning at different points in utterances in different languages.

Accessibility in English and Japanese

The other cross-linguistic phenomenon that is implemented in different ways in English and Japanese is the effect of conceptual factors like animacy on the order of words. Animacy influences the selection of syntactic functions in English, but it influences the *order* of syntactic functions (i.e., scrambling) in Japanese. But in both languages, this influence is attenuated in noun phrase conjunctions. The human data can be seen in Fig. 6, which summarizes McDonald et al.'s (1993) experiment 1 and Tanaka et al.'s (2008) experiment 1 in terms of the percentage of word order inversions. This figure, like the other

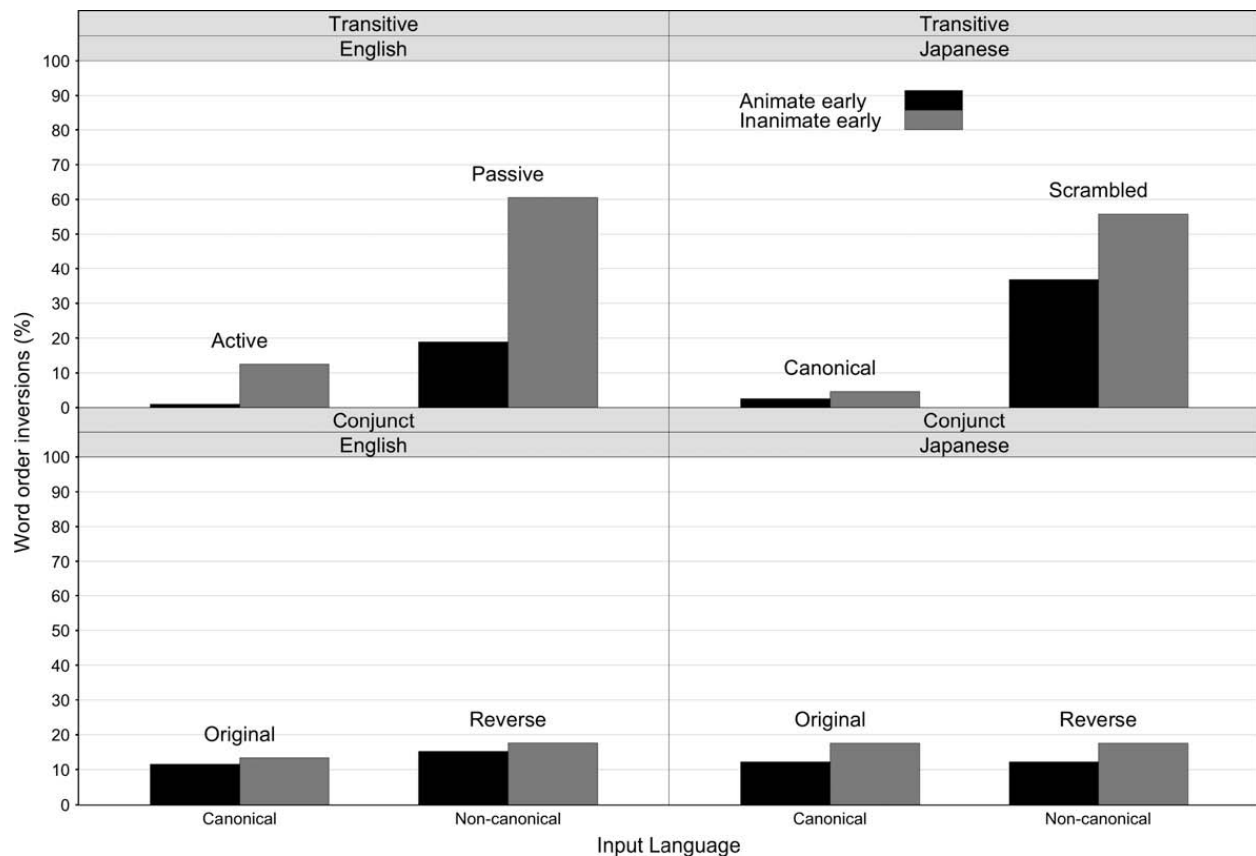


Fig. 6. Human animacy order data from McDonald et al. (1993, from their Table 2) and Tanaka et al. (2008, from their Table 1).

figures in this section, is divided into four sections by language and construction (transitive/conjunction). Within each box, there are four bars which represent the crossing of canonical and non-canonical structure with animacy condition. In addition, since canonicity differs by language and construction, a language-construction-specific label is placed above each set of bars.

In these sentence recall experiments, participants hear one word order, but sometimes invert that word order when asked to recall the original sentence from memory. The English transitive results (top left box in Fig. 6) shows that speakers are more likely to change a non-canonical passive with inanimate subject (e.g., “A refrigerator was purchased by a farmer”) into a canonical active than they were to change a passive with an animate subject (e.g., “The students were frightened by the sound”). The Japanese transitive results (top right box of Fig. 6) show a similar pattern of results, but these results index the order of syntactic functions. So speakers are more likely to change a non-canonical scrambled order with an inanimate object (e.g., “boat-o five people-ga carried”, which means “the boat was carried by five people”) to the canonical order (e.g., “five people-ga boat-o carried”) than they are to convert a non-canonical order with animate object (e.g., “five people-o boat-ga carried”, which means “five people were carried by the boat”) to the canonical order (e.g., “boat-ga five people-o carried”). While the overall pattern of results in both languages is similar, the syntactic function for the first element in the sentence in each language is different

and therefore an account based on either the order or assignment of syntactic functions would not be sufficient to explain the results in both languages.

The English and Japanese conjunction results are shown in the bottom half of Fig. 6. To make the conjunctions similar to transitives, McDonald et al. separated their conjunction items into two orders and one order was arbitrarily treated as the original order and the other order was treated as the reverse order (e.g., “the key and the manager” and “the manager and the key”). In Fig. 6, the original order was labeled as “canonical” and the reverse order was labeled as “non-canonical”. In the Japanese study, results were not divided in the same way, thus the same data was used for the canonical and non-canonical results in the bottom right box in Fig. 6. Regardless of this treatment of what is canonical or not with the conjunctions, the results in both languages were similar, in that there was no robust influence of animacy on the order of elements in conjunctions, in contrast with the findings with transitives. Thus, these studies show that sentence construction (transitive, conjunction) modulates how accessibility influences word order.

The conjunction data shows another important result. The base rate of word order inversions (the average height of the bars) differs between transitives and conjunctions. In sentence recall tasks, speakers are more likely to invert the word order of non-canonical transitives (passives or scrambled sentences), but they are less likely to do so with conjunctions or canonical transitives. Since the influence of

animacy on structure selection depends on the ability to freely alternate between structures, the difference in the base rate of inversions in transitives and conjunctions may help to explain the difference in the influence of animacy on these constructions. From the point of view of existing theories of sentence production, it is not clear how to explain the low rate of word order inversions in conjunctions. It is often thought that the order of items in a conjunction can be predicted from lexical properties of the component words (Benor & Levy, 2006; McDonald et al., 1993; Pinker & Birdsong, 1979), but since the division into “original” and “reverse” order was arbitrary, this account would predict that both orders should be equally distributed in each category and the inversion rate should be around 50%. If the division into these categories was inadvertently correlated with these lexical properties, then we would expect high accuracy in one category and low accuracy in the other, but that also does not seem to be the case. This ability to remember the arbitrary order of nouns in conjunctions seems to require that the message store some short-term memory of the given order, which is independent of the long-term knowledge that is stored about words (e.g., conceptual features, phonological form).

To explain the low rate of word order inversion for conjunctions in the model, the model was given some message information that simulated the memory for the “recalled” word order. This information made use of the prominence units, because the order of nouns in conjunctions have been shown to be sensitive to factors related to referential/discourse prominence (Bock & Irwin, 1980; Gleitman, January, Nappa, & Trueswell, 2007; Narasimhan & Dimroth, 2008). Just as the prominence units signaled the order of arguments in transitive and dative structures, the prominence units were also set to signal the order of nouns in conjunctions. For example, take the utterance “the man chase -ed the cat and the dog”. Since phrasal conjunctions were represented with two propositions, the conjunction “the cat and the dog” would be related to the same patient role in two different propositions (e.g., $OY = CAT$, $1Y = DOG$). To distinguish the two orders, the patient prominence unit OYY was used. If it was set to 1.0, then the default order was used (arbitrarily, OY before $1Y$ was the default order). When OYY was set to 0.6, then the opposite order was used ($1Y$ before OY order → “the dog and the cat”). By testing the model with this prominence information for conjunctions, we can see if this information can create the low base rate of inversions in human conjunction production.

To examine conceptual accessibility effects in the model, transitive and noun conjunction sentences with animate and inanimate nouns were generated. To mimic the sentence imitation task that McDonald et al. (1993) used, where speakers were given a sentence in a particular structure and asked to produce that sentence at a delay, the model was tested with messages in both canonical and non-canonical orders. The prominence unit activation was used to implement the short-term memory of the given order. For transitives, the non-canonical order (e.g., passive) was signaled by setting the agent prominence unit XX to 0.6 (for canonical orders, it was 1.0). For phrasal conjunctions, the non-canonical “reverse” order was signaled

by setting the prominence unit activation of the appropriate role (XX or YY) to 0.6 (the canonical “original” order conjunctions would have a value of 1.0). Below are examples of a transitive and a conjunction message–sentence pair in both canonical and non-canonical order.

Transitive

Canonical Message: $OA = CARRY$ $OX = CAT, INDEFINITE$

$OY = CUP, INDEFINITE$ $OE = PAST, SIMPLE, AA, XX, YY$

English: a cat carry -ed a cup.

Japanese: cat -ga cup -o carry -ta.

Non-canonical Message: $OA = CARRY$ $OX = CAT, INDEFINITE$

$OY = CUP, INDEFINITE$ $OE = PAST, SIMPLE, AA, XX = 0.6, YY$

English: a cup was carry -par by a cat.

Japanese: cup -o cat -ga carry -ta.

Conjunction

Canonical Message: $OA = FALL$ $OY = BALL, INDEFINITE$

$OE = PAST, SIMPLE, AA, YY, OA1A$

$1A = FALL$ $1Y = BROTHER, INDEFINITE$ $1E = PAST, SIMPLE, AA, YY$

English: a ball and a brother fall -ed.

Japanese: ball and brother -ga fall -ta.

Non-canonical Message: $OA = FALL$ $OY = BALL, INDEFINITE$

$OE = PAST, SIMPLE, AA, YY = 0.6, OA1A$

$1A = FALL$ $1Y = BROTHER, INDEFINITE$ $1E = PAST, SIMPLE, AA, YY$

English: a brother and a ball fall -ed.

Japanese: brother and ball -ga fall -ta.

The test set included 200 transitives with half made up of messages with animate agents and inanimate patients and the other half with the opposite pattern. The conjunction sentences were transitive sentences with the conjunction on either the agents or the patients of both clauses and there were equal numbers for each pairing of animacy and clause role (in the above example, OY is inanimate and $1Y$ is animate). Half of each set was presented with canonical or non-canonical event-semantics information. The dependent measure was percentage of word order inversions from the given utterance out of all message-appropriate utterances in the adult models. Mixed effects regressions were performed on the empirical logit of the proportion of word order inversions with animacy (animacy early, animacy late), canonicity (canonical, non-canonical), and language (English, Japanese) as treatment factors.

Fig. 7 suggests that animacy has its effect on non-canonical transitives and this is supported by a three-way interaction of animacy, construction, and canonicity ($t = 2.4$, $p < 0.021$). So the model mimics the distinction between transitives and conjunctions in the human data, although it does not have an architectural distinction that would naturally segregate these two constructions. Also, the non-canonical transitives tend to flip structures more than the other conditions ($t = 7.5$, $p < 0.001$), and this is also consistent with the behavior in the human studies. To insure that there was no effect of animacy on the ordering of nouns in conjunctions, a mixed model was applied to just the conjunctions. There was no main effect of animacy ($t = 0.65$, $p = 0.68$) and no interaction of animacy and canonicity ($t = 0.97$, $p = 0.53$). Generally speaking, the model is able to match the human word-order invertibility and animacy order behavior for both English and Japanese. To better understand how these results arise within the model, the input languages to the model were manipulated to

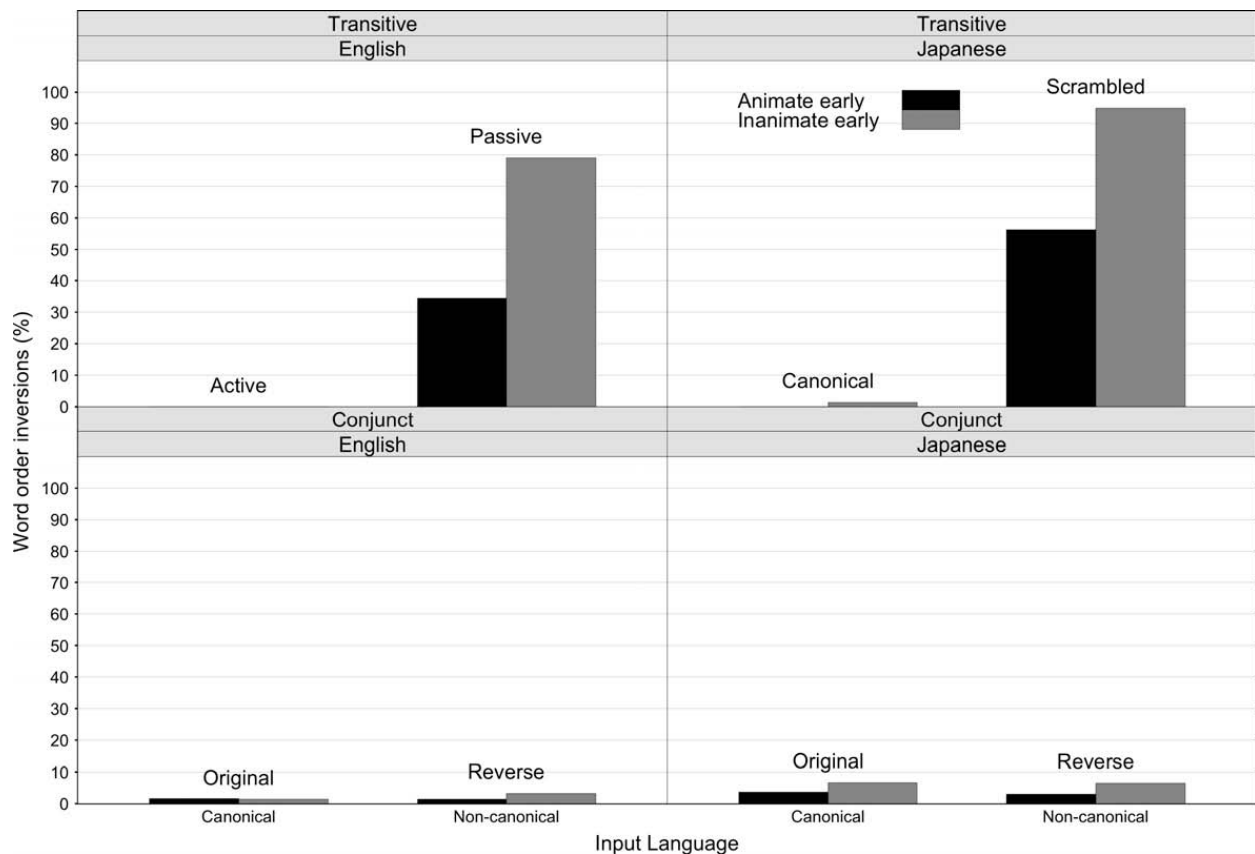


Fig. 7. Animacy order results with English and Japanese Standard model.

see what features of the input were related to these behaviors.

Given that almost every sentence in the input has an animate or inanimate element, a system that records the distribution of words in the input will have a lot of information for ordering these categories. The Dual-path model's knowledge of grammar comes in part from trying to predict the distribution of sentences in the input, and therefore, animacy order preferences in the model should also mirror those in its input. The above model that matched the human data well was trained with input that included animacy order biases. The input was manipulated so that utterances which had inanimate agents and animate patients were likely to occur in passive structures (Ferreira, 1994) and utterances which had animate agents and inanimate patients were more likely to occur in active structures (see "Appendix" for details). To see how this animate-first bias in the input influenced this Standard model, a new No-Animate-Bias model was created which did not have this animate-first bias in its input.

Fig. 8 shows that, in the No-Animate-Bias model, animacy no longer influences word order in non-canonical transitives (three-way interaction, $t = 0.10$, $p = 0.92$). There was also no main effect of animacy or interactions between it and other factors. Since the only difference between this model and the Standard model was the animacy bias in the input, this comparison demonstrates that the model's animacy bias is learned from the input.

Earlier, it was hypothesized that the difference between transitives and conjunctions in their sensitivity to animacy could be due to differences in the baseline rate of word order inversions. It was argued that the order of elements in conjunctions and transitives were guided by discourse prominence and both structures in the model were associated with event-semantics units that signaled their prominence. The main difference between the two structures was that the transitive alteration had a strong frequency asymmetry, where actives were more common than passives, while the conjunctions did not have this asymmetry (since the two orders have the same meaning, it is hard to know what internal ordering people are using when they produce conjunctions). This difference between transitives and conjunctions was implemented in the model by setting two parameters in the generation of the input language. The conjunction parameter was 50%, which means both of the conjunction orders were equally likely. The passive parameter was 20% to make actives more common than passives. This percentage of passives is higher than the 5–15% that have been found in corpora (Roland, Dick, & Elman, 2007). It was assumed that the percentage of passives in the model also simulated the influence of other structures which are known to share structure with passives and whose frequency helps to enhance passive production (e.g., intransitive locatives, Bock & Loebell, 1990; Chang et al., 2006).

To see how the rate of word order inversion influenced the transitive/conjunction distinction, the conjunction

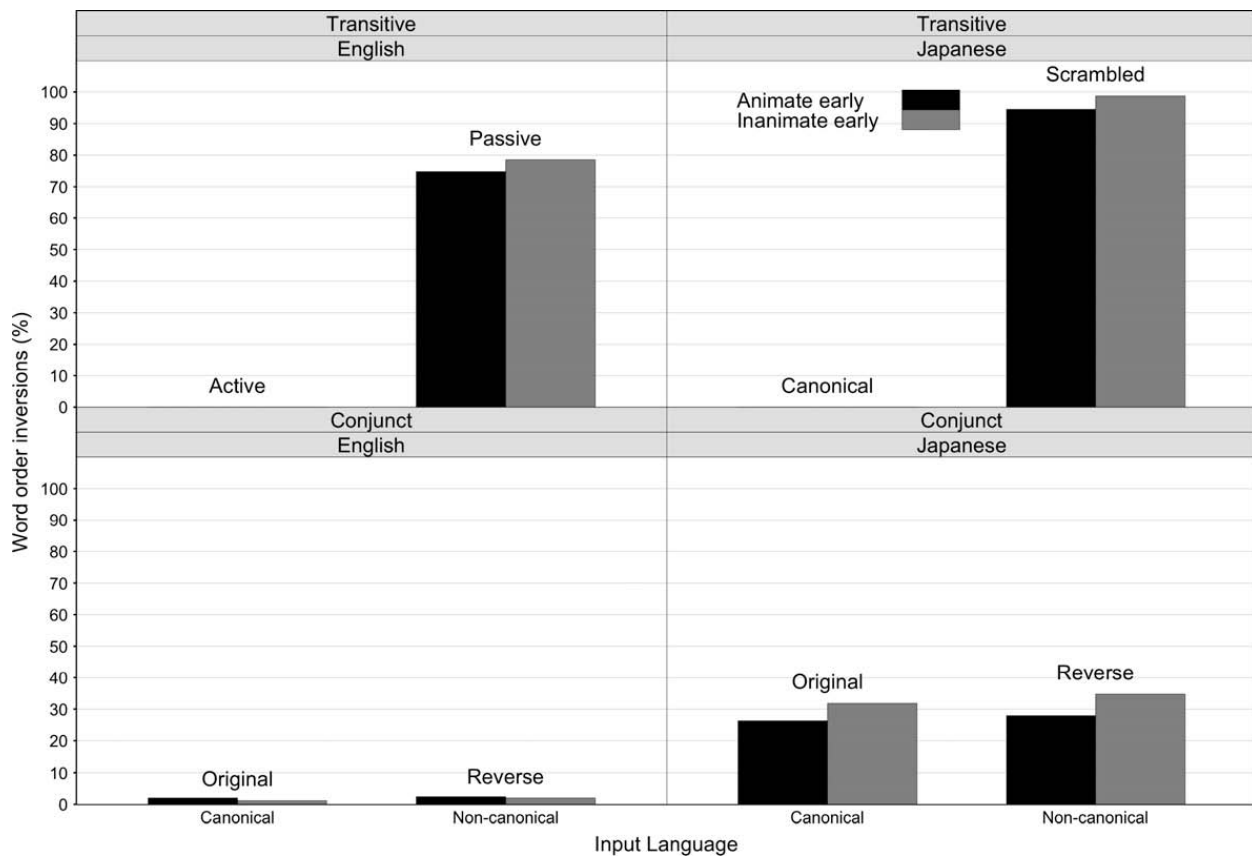


Fig. 8. Animacy order results with English and Japanese No-Animate-Bias model.

parameter was set to same value as the passive parameter (20%). Therefore, this new Conjunction=20% model was the same as the Standard model, except the conjunction alternation was biased towards one order just like the transitive alternation. The results showed that the Conjunction=20% model was more likely to invert the word order in conjunctions, and animacy had a robust effect on their ordering (Fig. 9). As in the Standard model, there was a three-way interaction of animacy, construction, and canonicity ($t = 7.31$, $p < 0.001$). But unlike that model, the Conjunction=20% models had an animacy effect in the conjunctions ($t = 3.56$, $p = 0.003$) and a significant interaction with canonicity ($t = 3.7$, $p = 0.002$). The higher base-rate of word order inversion made conjunctions more sensitive to the influence of animacy. Although both transitives and conjunctions use the same 20% parameter, the conjunctions seem to be less sensitive to animacy than the transitives. This is mainly due to the fact that the separation into original and reverse order was done arbitrarily and hence the lower frequency conjunction order, which was the order that was most sensitive to animacy, occurred equally in both orders and therefore animacy effects were also spread over both orders. In sum, these results suggest that the difference between conjunctions and transitives in human behavior could be due to differences in the balance of structures in each of these alternations. The lower frequency structure in asymmetric constructions comes to depend on other factors like animacy. But symmetric structures are learned

equally well and therefore do not depend as much on other factors for production.

To summarize, the model provides a syntactic learning-based account of accessibility effects in sentence production. Since the input was biased to place animate elements earlier in both English and Japanese sentences, the model learned representations that replicated this bias in its own production planning. This animate-early bias was modulated by the discourse prominence features in the message. In conjunctions, the prominence features had a strong influence on the order of words and therefore animacy had little influence. But the low frequency of passives and scrambled structures means that the prominence constraints on these structures are weaker, and hence animacy constraints had a greater influence on transitive production.

Conclusion

Word order biases in different languages provide an important test of theories of sentence processing and language learning. The weak optional nature of these biases suggested that they were not normal syntactic rules, but rather weak processing constraints. But the fact that these biases differ across languages means that learning must play a role in creating these biases. The current work provides an explicit unified model that demonstrates that learning within an incremental production architecture

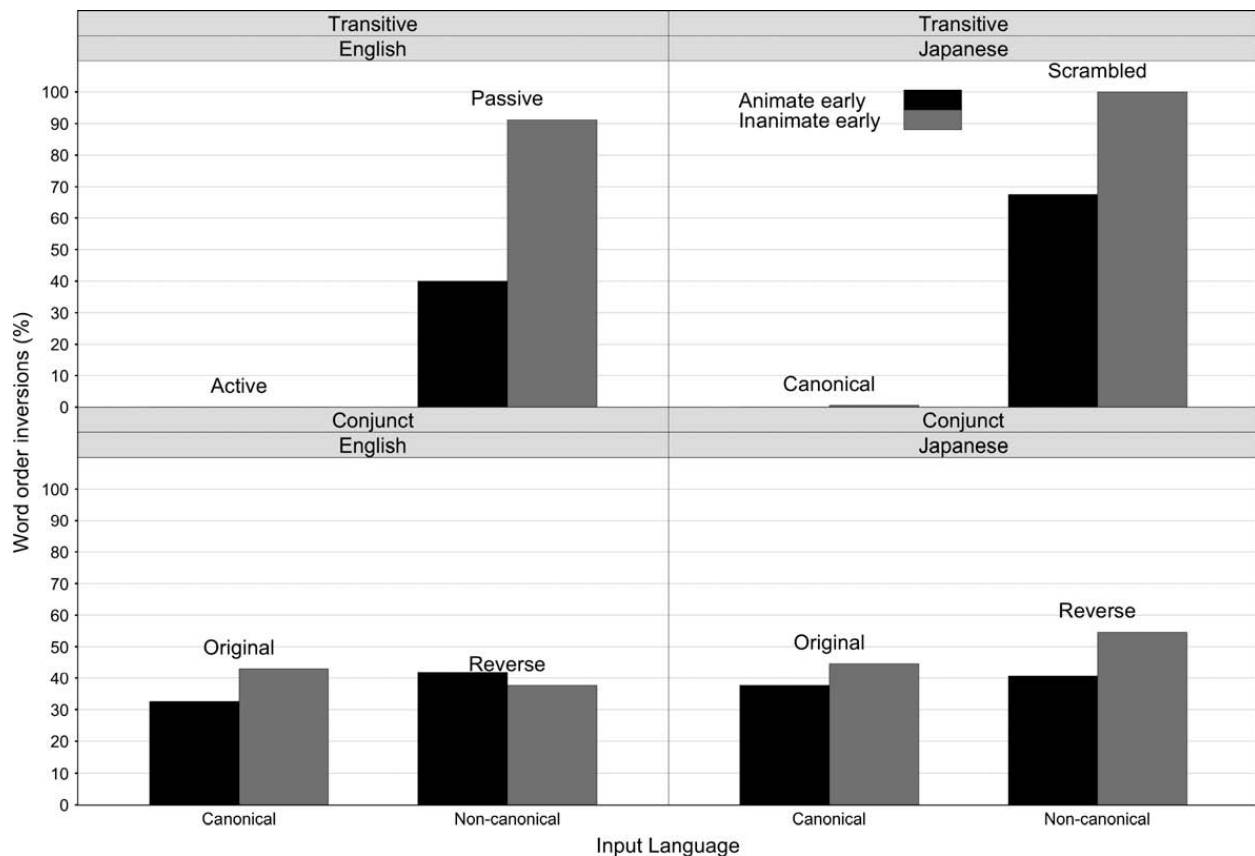


Fig. 9. Animacy order results with English and Japanese Conjunction = 20% model.

can explain both the language dependence and optionality of these word order biases.

How is the Dual-path model related to the distance minimization accounts of heavy NP shift (Gibson, 1998; Hawkins, 1994; Hawkins, 2004)? Both theories can account for the difference in the direction in heavy NP shift in English and Japanese and both posit that language features like verb position are important in this difference. The model does not directly generate multiple utterances and measure the verb–argument distances, so in this sense, it is not a direct implementation of a distance minimization account. But there is a sense in which the model does do a form of minimization through learning. The model tries to find a compact representation that can generate a sentence. Since English short-before-long structures share structure with simple main clause structures, the model would naturally try to use a shared representation for both. By sharing representations, the frequency of simple main clauses and short-before-long utterances is combined and the shared representation becomes stronger. Then when a choice must be made between it and a long-before-short utterance, the model will tend to prefer the structure with the shared representation. The model's account of heavy NP shift is therefore consistent with the overall claim of Hawkins (2004) that grammar is shaped by processing, but it suggests that this shaping is implemented through learning processes, rather than parsing or production, per se (Chang et al., 2006) and this learning-based account might help to explain other minimization phenomena in language use.

The Dual-path model makes several predictions about heavy NP shift. Since the model learns incremental syntactic representations that create the conditions for the shift, it predicts that heavy NP shift will be sensitive to context-specific information. For example, when choice points follow the verb, the model predicts that the distribution of structures with particular verbs will influence heavy NP shift. There is some evidence that supports this prediction (Stallings et al., 1998). In addition, since this model can explain adult structural priming data (Chang et al., 2006) with the same mechanism that is used to learn the representations for heavy NP shift, it predicts that heavy NP shift alternations should be modifiable through structural priming. In support of this prediction, Konopka and Bock (2009) found that structural priming could influence an alternation that has been thought to be governed mostly by weight considerations, verb–particle placement (“call X up” versus “call up X”, Lohse et al., 2004). Finally, another prediction arises from the simulation study that varied the properties of the language being trained. In that study, a language that was not verb-final, had no articles, and where relative clauses appeared before their heads showed the long-before-short preference. Chinese languages like Cantonese and Mandarin have this set of syntactic features and Cheung (2006) found that Cantonese speakers are better able to remember utterances with long theme arguments in earlier positions than in later positions. These results were argued to support distance minimization accounts (e.g., Hawkins, 2004), since Cantonese has a double object structure where themes precede recipients.

However, Mandarin Chinese has a different double object structure (recipients precede themes) and the distance minimization account does not bias towards the long-before-short preference as much in this type of language. Therefore, while a distance minimization account predicts some difference between Mandarin and Cantonese heavy NP shift in datives, the Dual-path model predicts a long-before-short preference in both languages. An explicit comparative study is needed to answer this question.

How does the Dual-path model's account of accessibility effects relate to the functional/positional theory of sentence production? Both the model and the functional/positional theory share the assumption of incremental planning and the distinction between lexical and structural information. They differ in how they implement accessibility effects. In the functional/positional theory, animacy is represented in the conceptual system. In the Dual-path model, the distribution of animate and inanimate words is learned and instantiated in the model's sequencing system. Therefore, rather than being completely conceptual, the model treats animacy as a syntactic distinction. This syntactic distinction acts like a filter that is biased for utterances where animate elements come earlier. This filter is however too weak to control the order in conjunctions, because the order in conjunctions is strongly controlled by the prominence features in the message.

There are two types of evidence that support this input-based distributional-learning account. First, there is corpus evidence that animate entities do tend to be mentioned earlier in constructions that alternate (Bresnan, Cueni, Nikitina, & Baayen, 2004) but no animate-first bias in conjunctions (Kelly, 1986, cited from McDonald et al., 1993). Thus the corpus data suggest that the input could help to create the distinction between conjunctions and transitives.

Second, work in structural priming suggests that animacy has its effect on word order at a surface level rather than at a higher level. Bock, Loebell, and Morey (1992) found that the proportion of actives with inanimate subjects was increased by a passive prime with an inanimate subject. This animacy order passive priming effect was the same magnitude as the priming effect for actives, which suggests that this priming effect is not sensitive to variations in structure or in thematic roles. If animacy effects are implemented as a filter on surface word orders and can be strengthened due to primes, then it is straightforward to explain these priming results.

It is also possible to extend this distributional learning account for animacy to explain how other conceptual factors influence word order. For example, the typicality of normal categories can influence the word order in conjunctions, but not the typicality of ad hoc categories (Onishi et al., 2008). Since the typical before atypical order for normal categories is better represented in corpora (Kelly et al., 1986), it might be possible for people to learn to order the members of typical categories from the input, but not the members of ad hoc categories, since the combination of these categories is by definition unusual and unlikely to be ordered relative to each other in the input. Typicality effects also seem to have a stronger effect on the order in conjunctions rather than major constituents. One possible

reason for this difference is that in an incremental learning system like the one presented here, it is easier to learn shorter distance dependencies like the relative order of elements in a conjunction compared to a longer distance dependency like the relative order of major constituents. Finally, typicality is culture-dependent and it would be surprising if speakers could acquire this culture-specific typicality structure without using distributional learning on conjunctions. For example, since “shrimp and mayonnaise” is a common topping on pizza in Japan, a person who does not think of these as being in the same category can still quickly learn that these are members of the same pizza topping category and the relevant ordering relationship between the two. The distributional learning account of conceptual influences on word order shows that it is not necessary to assume that conceptual distinctions only have their influence only at the conceptual level. Rather, distributional learning will cause these culture-specific distinctions to be represented at a surface structural level and the current account can help to explain some of the variability across structures in accessibility effects.

The model was able to explain animacy-order biases, even though the model does not have a conceptual unit for animate elements. In fact, even if the model had a unit that represented animacy, the weight between this animate conceptual unit and individual words must be learned and if the input for learning was unbiased with respect to animacy, then this animate conceptual unit would not be able to create an animacy order bias. As long as conceptual features can influence words only through learned links that are set by the input, meaning-form regularities such as animacy order will be highly constrained by the nature of the input, regardless of the type of conceptual representations that is used.

Although the model does not use conceptual features directly to implement its animacy bias, it does suggest that conceptual factors could influence language through their effect on the types of messages that are created and the effect of those messages on the input distribution. One way that this might occur is suggested in Bock et al.'s (1992) account of why animate nouns tend to precede inanimate nouns. They suggest that the preference for animate leaders is due to the fact that more things can be predicated about animate elements than inanimate elements. If more things can be predicated about animate elements, then those elements will occur more often in utterances in the most common sentence positions (e.g., subject). If those utterances form the input for the next generation of learners, then those learners can incorporate the animate-subject regularity into their language system and they can use the distributional category of animacy to generate word order, without having to embed the complete conceptual network that encodes predicability into their language system.

The unique feature of the model's account of adult production behavior in different languages is that it is defined in terms of a language acquisition device. One reason that this language acquisition account is successful is because it requires that learning takes place within the sentence production architecture itself, so that the representations that are learned are fitted to needs of an incremental message-based production process. Another important feature of

this account is that it uses a powerful general-purpose learning algorithm (back-propagation), that has been used for modeling of non-language phenomena like making coffee or driving a car (Botvinick & Plaut, 2004; Pomerleau, 1991). Non-language phenomena also differ across cultures – when parking cars, Americans typically park forward, while Japanese tend to park backwards – and therefore humans must have a learning mechanism that is powerful enough to deal with various types of culture-specific learning. The model leveraged these powerful mechanisms to learn two very different languages and explain syntactic phenomena that have not been previously explained in terms of general learning mechanisms.

Finally, this account takes a radical approach to the nature of syntactic universals. In most linguistic and psycholinguistic theories, there is an implicit assumption that it is possible to characterize human behavior in different languages with the same set of syntactic categories or distinctions (for some typological arguments against this idea, see Croft (2001), and for developmental issues with innate categories, see Tomasello (2003)). For example, it was assumed that heavy NP shift in different languages could be explained by a universal verb–argument distance metric or that Japanese must have a functional/positional distinction. Several syntactic theories assume that English and Japanese differ in the setting of a head-direction parameter, and that heads in English and heads in Japanese are the same category for the setting of this parameter (for some issues with this approach, see Mazuka, 1998). What the model demonstrates is that it is possible to explain the learning and use of different languages without requiring that the internal representations that support those behaviors be made up of the same syntactic categories or distinctions across languages. This is a radical approach, because universal syntactic categories are the basic building blocks for most syntactic theories. Since the model is learning its abstract syntactic representations by doing incremental prediction on utterances from languages that differ in many ways, it cannot help but learn different categories and dependencies in each language. It is interesting that this approach is able to explain the behavioral data, while at the same time, other approaches to syntax acquisition, such as linguistic theories or computational linguistic models, have not yet yielded explicit models of incremental meaning-based sentence production for these two languages. It is possible that the requirement that syntactic knowledge in different languages be instantiated in a common syntactic theory actually hinders the development of explicit models. Ironically, by making fewer assumptions about the universal nature of syntactic knowledge and allowing learning to determine the constraints for each language, we might actually get closer to a universal account of human language behavior.

Acknowledgments

I thank Kathryn Bock, Gary Dell, Hartmut Fitz, John Hawkins, Tomoko Kawaguchi, Tessei Kobayashi, Tadahisa Kondo, Ryoko Mugitani, Kristine Onishi, David Plaut, Hiroko Yamashita, Anthonio Cheung, and two anonymous reviewers for helpful comments on the manuscript. Previ-

ous versions of this work were presented at the CUNY Sentence Processing Conference.

Appendix

The simulations were implemented using the LENS neural network software version 2.6 (Rohde, 1999). The same Dual-path architecture that was used in Chang (2002) and Chang, Dell, and Bock (2006) was used, and those papers provide a more detailed motivation for the architecture. Fig. A1 displays the architecture vertically (production proceeds upward) and to be consistent with the labeling used in the earlier papers, the concept layer has been labeled the *what* layer and the roles layer has been labeled the *where* layer. The layers that operate in the comprehension direction are prefixed with the letter “c” (e.g., *cword*) to distinguish them from the layers in the production direction (e.g., *word*). There were 15 *compress* units, 15 *ccompress* units, 90 *hidden* units, 90 *context* units, 56 *what* units, 56 *cwhat* units, 9 *where* units, 9 *cwhere*, 9 *cwherecopy*, and 27 *event-semantic* units. The English model had 74 *word* and *cword* units, and the Japanese model had 68 *word* and *cword* units.

Although not shown in Fig. A1, the output of several layers are copied to other layers. The context layer activations are a copy of the previous hidden layer activations combined with the previous context layer activations. The *cword* layer receives the previous word layer activations. The *cwherecopy* units hold a running average of the *cwhere* activations. These copy connections allow the sequencing system to record various types of information about the preceding sequence and it can use that to help it to make the right prediction about the next word.

The learning rate was initially 0.1 for the first 60,000 patterns, and then it was gradually reduced to 0.001 after 10,000 more patterns and the model was trained for a further 10,000 patterns. Weights were initialized to random values from –0.2 to 0.2 (the weights in all models were randomized with the same seed). During training, weights were updated after every pattern (batch size = 1). A version of back-propagation was used to train the model where derivatives were clipped at 1.0 (doug momentum; Rohde, 1999) and a zero error radius of 0.1 was used. The rest of the settings for the units and the parameters were identical to those in the Chang, Dell, and Bock (2006) paper (see their “Appendix”).

Input language for the models

The input language was generated by an input generator that was separate from the model. The input generator first generated a language-independent message representation and then used a set of language-specific rules to create English or Japanese utterances from those messages. The messages were generated from abstract templates for three simple constructions: intransitive, transitive, and dative constructions. Below is an example of a transitive construction.

MESSAGE: A=CARRY X=BROTHER,DEFINITE Y=STICK,
INDEFINITE E=PAST,SIMPLE,AA,XX,YY

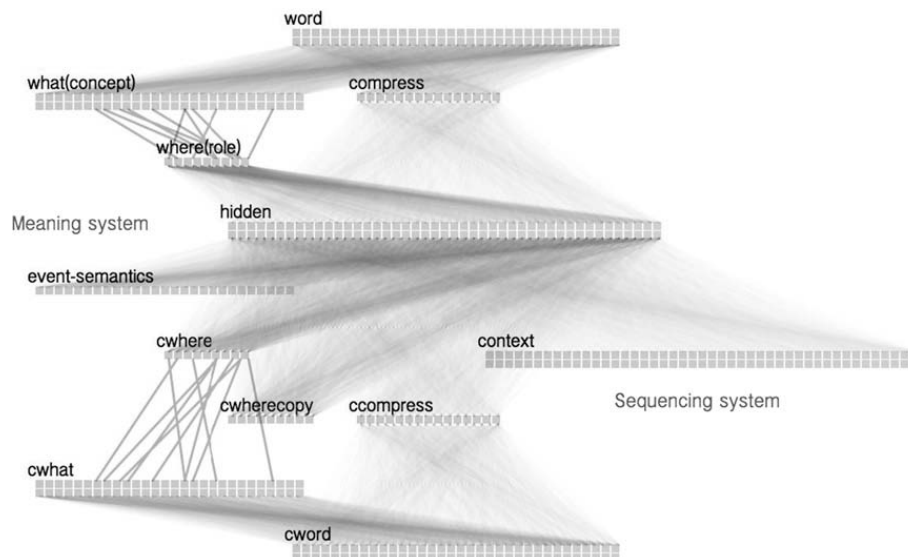


Fig. A1. Full architecture of Dual-path model (copy connections not shown).

English: the brother carry -ed a stick
Japanese: brother -ga stick -o carry -ta

The message was made up of role/concept pairs that encode the relational information in the utterance. As mentioned in the main text, the message was made up of concepts (e.g., BROTHER), roles (e.g., A, X, Y), and event-semantics ($E = \text{PAST, SIMPLE, AA, XX, YY}$). In the word sequences, verbs were made up of their base form (e.g., “carry”) and inflectional suffixes were treated as separate words (e.g., “-ed”, “-ta”). Irregular verbs were treated the same way, so “threw” would be represented in the model as “throw -ed”. Each argument in the message could be definite, indefinite, or pronominal (e.g., DEFINITE → “the”, INDEFINITE → “a”), but the actual implementation of this distinction was more complicated and is described later in this section.

The constructional templates had restrictions on the types of arguments that could occur in them. For example, the dative template required that the agent (X) and recipient (Z) be animate, while the patient (Y) was required to be inanimate. An example dative is shown below:

MESSAGE: A=GIVE X=GIRL,DEFINITE Y=KITE,INDEFINITE
Z=BOY,DEFINITE $E=\text{PAST,SIMPLE,AA,XX,ZZ,YY}$
English: The father give -ed the boy a kite
Japanese: father -wa boy -ni kite -o give -ta

This example also illustrates the canonical order for syntactic functions in Japanese datives, where the topic (wa) or subject (ga) precedes the indirect object (ni) and that precedes the object (wo). In English and Japanese, datives typically alternate between two orders and these two orders are signaled by information in the prominence units. The canonical ordered dative above has an event-semantics where XX, ZZ, and YY are all set to the default activation level of 1. To signal that the Z role is less prominent than the Y role, we set the ZZ unit to 0.6, and this activation level is associated with different structures in each language, as below:

MESSAGE: A=GIVE X=GIRL,DEFINITE Y=KITE,INDEFINITE
Z=BOY,DEFINITE $E=\text{PAST,SIMPLE,AA,XX,ZZ=0.6,YY}$
English: The girl give -ed a kite to the boy
Japanese: girl -wa kite -o boy -ni give -ta

In English, the dative alternation involved different surface structures (double object and prepositional dative). In Japanese, speakers typically use scrambling to yield the same effect. In transitives, the situation was even more complicated, since Japanese has several passive forms, and also allows scrambling of arguments in both active and passive structures. However, Japanese speakers tend to mainly use scrambling in situations where English speakers would use passives. Therefore to keep our simulations simple and to make the transitive and dative similar both within and across languages, Japanese passives were not implemented, and instead scrambling was used as the equivalent of English passives. Therefore, if the XX unit was set to 0.6, a passive was generated in English, and the direct object argument was scrambled to the beginning of the sentence in Japanese. A special verb ending (-par) was used to mark the past participle in English. In both languages, active canonical structures tend to be much more frequent than passives or scrambled structures, so the percentage of passive/scrambled utterances was set to be 20% of all transitives. An example passive/scrambled utterance is shown below.

MESSAGE: A=CARRY X=BROTHER,DEFINITE Y=STICK,
INDEFINITE $E=\text{PAST,SIMPLE,AA,XX=0.6,YY}$
English: a stick was carry -par by the brother
Japanese: stick -o brother -ga carry -ta

Arguments in a message could be pronominal (marked by the PRONOUN feature) and would be replaced in English with their pronoun form. Japanese is a pro-drop language, which means that in situations when an English speaker would use a pronoun, a Japanese speaker will often omit an argument. For example, if the stick in the previous

example was inferable from the discourse, then the message–sentence pair might appear like this.

MESSAGE: A=CARRY X=BROTHER,DEFINITE Y=STICK,
PRONOUN E=PAST,SIMPLE,AA,XX=0.6,YY
English: it was carry -par by the brother
Japanese: brother -ga carry -ta

In addition to these simple constructions, complex constructions were assembled out of simple ones. One type of complex construction was a sentential conjunction, which simply combined two utterances with the conjunction “and”. Sentence conjunctions in Japanese can be made by placing the verb ending “-te” after the verb of the first utterance. Since these messages have two propositions, the roles in each proposition must be distinguished, and so the number 0 was placed before the first clause role and the number 1 before the second clause role (all of the previous examples were 0 clause roles).

MESSAGE: 0A=PUSH 0X=GIRL,DEFINITE 0Y=ORANGE,
DEFINITE 0E=PRESENT,SIMPLE,AA,XX,YY
1A=TOUCH 1X=MOTHER,DEFINITE 1Y=APPLE,DEFINITE
1E=PRESENT,SIMPLE,AA,XX,YY
English: the girl push -es the orange and mother touch -
es the apple.
Japanese: girl -wa orange -o push -te mother -ga apple
-o touch

Another type of complex construction was the phrasal conjunction (e.g., “the girl pushes the orange and the apple”). A phrasal conjunction was similar to its sentential conjunction paraphrase (e.g., “the girl pushes the orange and the girl pushes the apple”) in that it was represented by two propositions, but in addition it had binding nodes that marked the elements that were overlapping between the two propositions. In the above phrasal conjunction example, the pusher in both sentences is the same girl and the pushing action is the same pushing action, so the message had binding nodes that marked the overlap in the agent X role (0X1X) and the action A role (0A1A). A matched example is shown below (they differ only in terms of the binding nodes).

Sentential Conjunction Message:
0A=PUSH 0X=GIRL,DEFINITE 0Y=ORANGE,DEFINITE
0E=PRESENT,SIMPLE,AA,XX,YY
1A=PUSH 1X=GIRL,DEFINITE 1Y=APPLE,DEFINITE
1E=PRESENT, SIMPLE,AA,XX,YY
English: the girl push -es the orange and girl push -es the
apple.
Japanese: girl -wa orange -o push -te girl -ga apple -o push

Phrasal Conjunction Message:
0A=PUSH 0X=GIRL,DEFINITE 0Y=ORANGE,DEFINITE
0E=PRESENT,SIMPLE,AA,XX,YY,0A1A,0X1X
1A=PUSH 1X=GIRL,DEFINITE 1Y=APPLE,DEFINITE
1E=PRESENT, SIMPLE,AA,XX,YY
English: the girl push -es the orange and the apple.
Japanese: girl -wa orange and apple -o push

To implement the high rate of recall of the given order (McDonald et al., 1993), the activation of the appropriate prominence unit was used to control the order in conjunctions. In the above example, the conjunction involves the patient Y role, so the YY prominence role in the 0 clause event-semantics controlled the order. When the YY was the normal activation of 1.0, the order placed the 0 clause element first (e.g., “the orange and the apple”). When YY was set to 0.6, then the 1 clause element came first (e.g., “the apple and the orange”).

The last complex construction was a sentence with a relative clause. In these sentences, there is a linked referent between the main and relative clauses. To mark the linked referent in the message, the appropriate binding node was set (e.g., 0Y1Y). The examples below differ in terms of which main clause role the relative clause is attached to

Relative clause construction with binding between 0Y and 1Y:
0A=GIVE 0X=BOY,DEFINITE 0Y=ORANGE,DEFINITE
0Z=DOG,DEFINITE 0E=PRESENT,SIMPLE,AA,XX,ZZ,YY,0Y1Y
1A=PUSH 1X=GIRL,DEFINITE 1Y=ORANGE,DEFINITE
1E=PRESENT,SIMPLE,AA,YY

English: the boy give -es the dog the orange that the girl push
-es.
Japanese: boy -wa dog -ni girl -ga push orange -o give.

Relative clause construction with binding between 0Z and 1Y:
0A=GIVE 0X=BOY,DEFINITE 0Y=ORANGE,DEFINITE
0Z=DOG,DEFINITE 0E=PRESENT,SIMPLE,AA,XX,ZZ,YY,0Z1Y
1A=PUSH 1X=GIRL,DEFINITE 1Y=DOG,DEFINITE
1E=PRESENT,SIMPLE,AA,YY

English: the boy give -es the dog that the girl push -es the
orange.
Japanese: boy -wa girl -ga push dog -ni orange -o give.

Relative clause construction with binding between 0X and 1Y:
0A=GIVE 0X=BOY,DEFINITE 0Y=ORANGE,DEFINITE
0Z=DOG,DEFINITE 0E=PRESENT,SIMPLE,AA,XX,ZZ,YY,0X1Y
1A=PUSH 1X=GIRL,DEFINITE 1Y=BOY,DEFINITE
1E=PRESENT,SIMPLE,AA,YY

English: the boy that the girl push -es give -es the dog the
orange.
Japanese: girl -ga push boy -wa dog -ni orange -o give.

To summarize, the model's language had three simple constructions (intransitive, transitive, dative) and all combinations of these structures in phrasal conjunctions, sentential conjunctions, and relative clauses (with all possible bindings between roles). There were a total of 50 constructions in the language (3 simple constructions, 9 sentential conjunctions, 6 phrasal conjunctions, 32 structures with relative clauses).

The language was made up of a set of concepts that included 14 animate concepts, 14 inanimate concepts, 6 adjectives, 6 intransitive actions, 6 transitive actions, and 6 dative actions. Noun phrases occurred with adjectives 25% of the time, and they had a feature that selected whether it was definite, indefinite, or a pronoun (each equally likely). Tense could be PRESENT or PAST, and aspect could be SIMPLE or PROGRESSIVE. One difference from previous work is that the present language has sets of roles for each of two propositions. The first proposition roles (0A, 0X, 0Y, 0Z) were used for simple single clause

propositions and the main clause roles of two proposition utterances. The second proposition roles (1A, 1X, 1Y, 1Z) were used for relative clause roles or the second clause in conjunctions. Also the event-semantic features were distinguished for first or second proposition (e.g., the feature PAST is different for 0E = PAST and 1E = PAST). In addition there was one set of binding nodes that would signal the links between the first and second proposition roles. Every combination of roles had its own binding node (0A1A, 0X1X, 0X1Y, 0X1Z, 0Y1X, 0Y1Y, 0Y1Z, 0Z1X, 0Z1Y, 0Z1Z). Whenever the concept in the first proposition referred to the same concept in the second proposition (e.g., utterances with relative clauses), the appropriate binding node was activated. Phrasal conjunctions were assumed to be two proposition messages with binding nodes activated for all pairs of roles except for the roles that were part of the conjunction.

After a message was generated, its event-semantics was changed to implement the structural alternations. Initially, the event-semantics in the message were set to the default values for that construction. Then, the animacy bias was applied. This was done by setting the prominence activation to produce actives for 40% of the utterances with animate agents and inanimate patients, and the equivalent activations for passives for 40% of the utterances with animate patients and inanimate agents. Since about half of the utterances had animate and inanimate arguments, this means that about 20% of the utterances were forced to appear in animate before inanimate order. This percentage of animacy flips was found to generate an animacy pattern that was similar to the magnitude of animacy order biases in the human studies. The remaining 80% of the utterances were allowed to alternate. In transitive structures, the prominence activation and structure was changed to passive/scrambled order 20% of the time and the remaining 80% stayed in active/canonical order. This percentage of passives was needed to insure that in testing of the model, the model would produce a passive some part of the time. For datives, both structures (e.g., English double object, prepositional dative) occurred 50% of the time. For conjunctions, both orders were equally likely.

Next, the sentences were generated that corresponded to the messages that were created. This was done using a set of language-specific rewrite rules that converted the message into an appropriate English or Japanese sentence. In English, prepositions and articles were added. In Japanese, particles were added. In both languages, tense and aspect was translated into the appropriate verb morphology. In English, noun phrases marked as pronouns were converted into case-marked pronouns (he, she, it, him, her) and in Japanese, these noun phrases were omitted. After the sentence was generated, the final step was to create some mismatch between the message and the utterances, so that it would be possible for the model to alternate between structures. To do this, 25% of the messages were randomly changed so that the message predicted the alternative structure (e.g., a passive sentence would be paired with a message that would signal an active). This random flipping parameter was important in creating the base-rate of word order inversion in conjunctions.

After a message–sentence pair was generated, it was then converted into the model's internal representation. The model had a localist word representation and words were presented one at a time as a target for the *word* layer or as an input for the *cword* layer. The model uses weights to instantiate the message, so the link between the appropriate role node in the *where* layer and the corresponding concept in the *what* layer was set to 6 (the corresponding *cwhat*–*cwhere* links were also made). Concepts used a localist representation, but one exception to this was the representation of the articles and pronouns. Rather than having three units for indefinite, definite, and pronoun, the strength of the role to DEFPRO unit was used to encode these distinctions. Definite noun phrases had a role-DEFPRO link that was 60% as strong as a normal message link (weight = 4) and indefinite noun phrases were 30% as strong (weight = 2). Pronouns had a role-DEFPRO link that was 100% strength (weight = 6), but the activation of the lexical content of the noun phrase was half as strong (weight = 3). This activation pattern helps to explain why pronouns are used in English (the activation of DEFPRO is correlated with pronoun use) but argument omission is used in Japanese (if the lexical content is less activated, then the model can omit it).

To make our language both learnable and similar to human language, each structure had a particular frequency for selection during training and test set generation. To insure that each structure in the language was trained, the training set included 10 versions of each construction (500 message–sentence pairs). Since the set of constructions includes many more relative clause constructions than main clause constructions, the remaining 39500 utterances in the language were sampled such that the main clause constructions were 50 times more likely than the relative clause structures. To equate the factors influencing transitives and conjunctions for the conceptual accessibility comparison, the frequency of conjunctions in a particular construction was made to be equal to main clause frequency for that construction. This training set seemed to be sufficient to yield a high level of accuracy in both languages, and yet had some similarity to the input that humans use to learn language. The 400 sentences in a separate test set were generated with a similar methodology to the training set (with at least 1 utterance from each construction).

Grammaticality was tested by labeling the produced utterance with syntactic categories and then checking if the sequence of categories was present in a database of grammatical sequences. The database was created by labeling all of the target utterances in the training set with syntactic categories and then recording the syntactic category sequences in the database. To insure full coverage of the English grammar, additional sequences were added to the database so that each noun phrase could occur as a pronoun, article noun, and article adjective noun. For the Japanese grammar, a similar set of additions were made, except that pronouns were omitted. Message equivalence accuracy involved transforming the produced utterance into a form that was similar to the messages in the model, where the different surface structures in alternations had a common message structure. Message accuracy

ignored minor deviations such as incorrect articles and verb morphology.

The model and the input-generator are available on the web at <http://sites.google.com/site/sentenceproduction-model/>.

References

- Agresti, A. (2002). *Categorical data analysis*. Hoboken, NJ: John Wiley & Sons, Inc.
- Aikawa, T., Melero, M., Schwartz, L., & Wu, A. (2001). Sentence generation for multilingual machine translation. In *8th European workshop on natural language generation*, Toulouse, France.
- Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: General*, 137, 791–810.
- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264.
- Arnold, J. E., Wasow, T., Losongco, A., & Ginstrom, R. (2000). Heaviness vs. Newness: The effect of structural complexity and discourse status on constituent ordering. *Language*, 76(1), 28–55.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. New York: Cambridge University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Baker, M. C. (2005). Mapping the terrain of language learning. *Language Learning and Development*, 1(1), 93–129.
- Bencini, G. M. L., & Valian, V. (2008). Abstract sentence representation in 3-year-olds: Evidence from comprehension and production. *Journal of Memory and Language*, 59, 97–113.
- Benor, S. B., & Levy, R. (2006). The chicken or the egg? A probabilistic analysis of English binomials. *Language*, 82(2), 233–278.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1–47.
- Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 575–586.
- Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, 26(2), 119–137.
- Bock, J. K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129, 177–192.
- Bock, J. K., & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior*, 19(4), 467–484.
- Bock, J. K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, 48(4), 653–685.
- Bock, J. K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego: Academic Press.
- Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition*, 35, 1–39.
- Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99, 150–171.
- Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21, 47–67.
- Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111, 395–429.
- Branigan, H. P., & Feleki, E. (1999). Conceptual accessibility and serial order in Greek speech production. In *Proceedings of the 21st cognitive science society conference*. Vancouver: Canada.
- Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118(2), 172–189.
- Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2004). Predicting the dative alternation. In *Proceedings of the royal Netherlands academy of science colloquium on foundations of interpretation*, Amsterdam.
- Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19, 441–472.
- Byrne, B., & Davidson, E. (1985). On putting the horse before the cart: Exploring conceptual bases of word order via acquisition of a miniature artificial language. *Journal of Memory and Language*, 24, 377–389.
- Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science*, 26, 609–651.
- Chang, F., Dell, G. S., & Bock, J. K. (2006). Becoming syntactic. *Psychological Review*, 113, 234–272.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29, 217–229.
- Chang, F., Lieven, E., & Tomasello, M. (2008). Automatic evaluation of syntactic learners in typologically-different languages. *Cognitive Systems Research*, 9, 198–213.
- Cheung, A. (2006). Processing factors in language comprehension and production: The case of Cantonese dative constructions. Unpublished master's thesis. The University of Hong Kong, Hong Kong.
- Clark, H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford, UK: Oxford University Press.
- de Marneffe, M.-C., Grimm, S., Priva, U. C., Lestrade, S., Ouzbek, G., Schnoebelen, T., et al. (2007). *A statistical model of grammatical choices in children's production of dative sentences*. University of York, England: Formal Approaches to Variation in Syntax.
- De Smedt, K. (1994). Parallelism in incremental sentence generation. In G. Adriaens & U. Hahn (Eds.), *Parallel natural language processing* (pp. 421–447). Norwood, NJ: Ablex.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, 104, 123–147.
- Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23, 517–542.
- Dijkstra, A., & De Smedt, K. (Eds.). (1996). *Computational psycholinguistics: AI and connectionist models of human language processing*. London: Taylor & Francis.
- Dominey, P. F., Hoen, M., & Inui, T. (2006). A neurolinguistic model of grammatical construction processing. *Journal of Cognitive Neuroscience*, 18, 2088–2107.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Ferreira, F. (1994). Choice of the passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33, 715–736.
- Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, 35(5), 724–755.
- Ferreira, V. S., Bock, K., Wilson, M., & Cohen, N. J. (2008). Memory for syntax despite amnesia. *Psychological Science*, 19, 940–946.
- Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40, 296–340.
- Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, 32, 669–692.
- Fitz, H. (2009). *Neural Syntax*. Amsterdam: ILLC publication series.
- Fitz, H., & Chang, F. (2008). The role of the input in a connectionist account of the accessibility hierarchy in development. In *Proceedings of the 32nd Boston University Conference on Language Development*. (pp. 120–131).
- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey. Language: Psychological and biological aspects* (Vol. 3, pp. 69–96). Cambridge, England: Cambridge University Press.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1–76.
- Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57(4), 544–569.
- Gordon, J. K., & Dell, G. S. (2003). Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science*, 27, 1–40.
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274–279.
- Hakuta, K. (1981). Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition*, 9, 197–236.
- Hakuta, K., de Villiers, J. G., & Tager-Flusberg, H. (1982). Sentence coordination in Japanese and English. *Journal of Child Language*, 9, 193–207.

- Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, 58(2), 214–238.
- Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge, UK: Cambridge University Press.
- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. New York City: Oxford University Press.
- Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes*, 40(1), 1–35.
- Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, 50(2), 182–195.
- Jaeger, T. F., & Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In *Proceedings of the cognitive science society conference*, Washington, DC.
- Kaschak, M. P., & Borreggine, K. L. (2008). Is long-term structural priming affected by patterns of experience with individual verbs? *Journal of Memory and Language*, 58, 862–878.
- Kelly, M. (1986). On the selection of linguistic options. Unpublished doctoral dissertation, Cornell University, Ithaca, NY.
- Kelly, M. H., Bock, J. K., & Keil, F. C. (1986). Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language*, 25(1), 59–74.
- Kempen, G., & Harbusch, K. (2003). Word order scrambling as a consequence of incremental sentence production. In H. Härtl & T. Tappe (Eds.), *Mediating between concepts and grammar* (pp. 141–164). Berlin: Mouton De Gruyter.
- Kempen, G., & Harbusch, K. (2004). A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann (Ed.), *Language production*. Berlin: Mouton.
- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–258.
- Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, 58, 68–101.
- Langkilde-Geary, I. (2002). An empirical verification of coverage and correctness for a general-purpose sentence generator. In *Proceedings of the international natural language generation conference*, New York City, NY.
- Lohse, B., Hawkins, J. A., & Wasow, T. (2004). Domain minimization in English verb–particle constructions. *Language*, 80, 238–261.
- Mazuka, R. (1998). *The development of language processing strategies: A cross-linguistic study between Japanese and English*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, 25, 188–230.
- Narasimhan, B., & Dimroth, C. (2008). Word order and information status in child language. *Cognition*, 107, 317–329.
- Onishi, K. H., Murphy, G. L., & Bock, K. (2008). Prototypicality in sentence production. *Cognitive Psychology*, 56, 103–141.
- Pinker, S. (1984). In *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1989). In *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Pinker, S., & Birdsong, D. (1979). Speakers' sensitivity to rules of frozen word order. *Journal of Verbal Learning and Verbal Behavior*, 18(4), 497–508.
- Pomerleau, D. A. (1991). Efficient training of artificial neural networks for autonomous navigation. *Neural Computation*, 3(1), 88–97.
- Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, 42(2), 168–182.
- Rohde, D. L. T. (1999). *LENS: The light, efficient network simulator* (No. CMU-CS-99-164). Pittsburgh, PA: Carnegie Mellon University, Department of Computer Science.
- Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, 57, 348–379.
- Rosenbach, A. (2005). Animacy versus weight as determinants of grammatical variation in English. *Language*, 81(613–644).
- Ross, J. R. (1967). Constraints on variables in syntax. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
- Savage, C., Lieven, E. V. M., Theakston, A. L., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, 6(5), 557–567.
- Sridhar, S. N. (1988). *Cognition and sentence production: A cross-linguistic study*. New York: Springer-Verlag.
- St. John, M. F., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, 46(1–2), 217–257.
- Stallings, L. M., MacDonald, M. C., & O'Seaghdha, P. G. (1998). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language*, 39(3), 392–417.
- Staub, A., Clifton, C. J., & Frazier, L. (2006). Heavy NP shift is the parser's last resort: Evidence from eye movements. *Journal of Memory and Language*, 54(3), 389–406.
- Tanaka, M., Branigan, H. P., & Pickering, M. J. (2008). Conceptual influences on word order and voice in Japanese sentence production. Unpublished manuscript.
- Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, 108, 51–68.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Wasow, T. (2002). *Postverbal behavior*. Stanford: CSLI Publications.
- Yamashita, H., & Chang, F. (2001). "Long before short" preference in the production of a head-final language. *Cognition*, 81, B45–B55.