Does sentence structure boost early word learning?

Bemd, van den, Eva; Mos, M.B.J.; Alishahi, A.; Shayan, Shakila

Does sentence structure boost early word learning? An artificial language learning study

Eva van den Bemd, Maria Mos, Afra Alishahi & Shakila Shayan

Abstract

Many studies show that structural information in a sentence is used as an information source to find the meaning of a novel word in that sentence, a phenomenon known as syntactic bootstrapping (Gleitman, 1990). However, no research thus far has investigated the use of structural cues in the first stages of learning a new language, when both word and structure knowledge are still in development. We conducted an artificial language learning experiment to investigate the interaction between learning word meaning and sentence structure. Participants watched animated scenes while hearing a three-word sentence describing the scene, and were subsequently tested on their linguistic knowledge. Results show that some, but not all participants were able to uncover the word meanings and the underlying sentence structure, and that knowledge of the structure may start to boost vocabulary learning in a later learning phase, which would be in line with the syntactic bootstrapping hypothesis. Follow-up studies are conducted to corroborate these preliminary results.

1 Introduction

How do language learners acquire the meaning of a word? The challenges for word learners are diverse: they have to figure out which part of the sentence refers to which part of the event that the sentence describes (Carey, 1978), and they may perceive aspects that are not mentioned in the utterance (Quine, 1960). Many studies have tried to explain how language users manage to learn words in a new language despite these challenges, showing that we use different mechanisms to narrow down the possible referents of new words. Two of these mechanisms are *cross-situational word learning* and *syntactic bootstrapping*.

Often, a language learner hears a new word in a context with multiple candidate referents and therefore cannot identify the referent of the new word on the basis of this single exposure. *Cross-situational word learning* (CSWL) refers to the learner's ability to combine

information across multiple exposures to determine the referent of an unknown word (Pinker, 1989). CSWL experiments with artificial words show that both children and adults track co-occurrences between unknown spoken words and visual referents across situations to deduce the meanings of these words (e.g., Kachergis, Yu, & Shiffrin, 2010; Smith, Smith, & Blythe, 2011; Smith & Yu, 2008; Yu & Smith, 2011). In such CSWL experiments, participants are typically presented with an unknown pseudoword and a grid of several objects. One object in this grid represents the referent of the word, which is surrounded by random distractor objects. Participants are initially unaware of the correct referent; however, after encountering the word (and its referent) in several grids with different distractor objects, they are able to learn the meaning of the word, which shows that cross-situational learning mechanisms can be employed to learn new words.

A foundational experimental paradigm within word learning research is the Human Simulation Paradigm (HSP) introduced by Gillette, Gleitman, Gleitman and Lederer (1999). In HSP studies, participants watch silent videos of interactions with target words marked by a beep. Their task is to guess these words, being provided with different degrees of information about the word's linguistic context. This and later HSP experiments (e.g., Papafragou, Cassidy, & Gleitman, 2007; Piccin & Waxman, 2007) show that both adults and children use contextual cues to discover the meaning of words. The use of such cues to find the referent of an unknown word is called *syntactic bootstrapping* (Gleitman, 1990).

Other syntactic bootstrapping studies present participants with a new word in a familiar structure of a language they already speak, for example 'she's gorping her over there' or 'she's gorping around'. In these sentences, the linguistic context helps the learner to find the referent of the novel word, as it tells not only that it concerns an action, but also whether it refers to a transitive action (as in the first example) or an intransitive action (the latter example; Fisher, 1996). Results of these studies also show that both children and adults use structural cues to identify the meaning of an unknown word (e.g., Brown, 1957; Fisher, Klingler, & Song, 2006; Gertner, Fisher, & Eisengart, 2006; Lee & Naigles, 2008).

Most of these studies investigate the effect of structural context on *verb* learning, as syntactic bootstrapping in its more narrow interpretation is seen as facilitative for the acquisition of verbs. Learners use the syntactic frames in which verbs are placed as a source of information about their meaning, as in the above example sentences containing the unknown word 'gorping'. Others investigate the effect syntactic bootstrapping on the acquisition of other word categories, assuming that the meanings of words o

utside the verb class can also be predicted when they appear in a familiar structure. For instance, Fisher et al. (2006) show that children use sentence structure to determine whether a new word is an object-category name ('this is a corp') or a spatial-relational term ('this is acorp my box'). These results indicate that the benefits of syntactic bootstrapping go beyond verb learning and are useful for the acquisition of other word categories as well.

Even though these and other studies on CSWL and syntactic bootstrapping revealed interesting insights into mechanisms that are used in word learning, they are limited when it comes to simulating real-world language learning. First, CSWL studies ignore structure, as new words are presented to the participants in an isolated manner, without embedding them in sentences. This does not reflect the linguistic input language learners receive when they come into contact with a new language, and may therefore not show the learning mechanisms that are used in a real-world learning situation, as syntactic bootstrapping studies show that word learning is facilitated by cues in the structural context. Second, to date, syntactic bootstrapping studies have focused on learning new words in a familiar language, with the learner being fully aware of the structure surrounding the unknown word. However, language learners who first come into contact with a new language have to acquire word meanings and underlying structures simultaneously from the linguistic input. A clear picture of how word learning and structure knowledge interact in these first phases of learning a new language is still missing.

Koehne and Crocker (2010, 2011) conducted artificial language learning experiments to investigate the interaction between cross-situational learning and syntactic bootstrapping. However, in these studies, participants performed an 'isolated verb learning session' before being exposed to linguistic input in structured sentences. This does not allow an investigation of the time line and onset of syntactic bootstrapping effects in a more naturalistic learning scenario. As Piccin and Waxman (2007) indicated, it remains to be clarified when in linguistic development structural information begins to convey its benefits to word learning, and how this information supports the acquisition of various kinds of words.

2 The current experiment

We aim to investigate the impact of gradually emerging knowledge of syntactic structure on word learning. The structural cue for word learning manipulated in our experiment is word order. Word order knowledge has been shown to facilitate the acquisition of the meaning of new words (Gertner et al., 2006) and can therefore be regarded as an aspect of structure knowledge that is employed for syntactic bootstrapping purposes.

An artificial language learning set-up is employed, in which the linguistic input is either structured (VSO word order) or unstructured (random word order). In the structured condition, participants are (at least initially) unaware of the word order. VSO order is very rare, as only 7% of the world's languages have this as the basic word order and it is not the basic order in languages such as Dutch and English (Dryer, 2008). We can therefore assume that most (if not all) of the (Dutch) participants are unfamiliar with this basic word order and will have to learn this order from the linguistic input in the experiment.

The task contains two training sessions (observations of scene-utterance pairs) and two test sessions assessing the participants' knowledge of the meaning of words and full sentences. The use of an artificial language allows the presentation of an input language that is new for all participants. This creates the opportunity to study word learning processes in a language of which both the words and the structures are unknown to the learner. Moreover, artificial languages provide precise control over input issues such as frequency and phonological similarity (Magnuson, Tanenhaus, Aslin, & Dahan, 2003). For our research purposes, the use of an artificial language allows for a manipulation of the underlying structure while keeping all other input aspects constant. A large body of empirical evidence shows that many mechanisms involved in artificial language learning and processing are shared with those of natural language, as artificial language learning studies have replicated findings from natural language learning experiments (Folia, Uddén, De Vries, Forkstam, & Petersson, 2010; Magnuson et al., 2003; Mintz, 2002). This demonstrates the usefulness of artificial language learning experiments as a complementary method to traditional studies for the investigation of specific hypotheses on natural language processing.

We investigate whether the proposed set-up allows participants to learn word meanings and whether structure knowledge affects the word learning process for different word categories. Moreover, we aim to explore the possibility to distinguish learners with structure knowledge on the basis of their eye movements during the processing of sentences in the language. We focus on the following research questions:

(1) Can adult participants successfully learn the meaning of novel words appearing in multi-word sentences, when they are not familiar with the structure of the language?

(2) (How) does structure knowledge affect the word learning process?

(3) Does this effect of structure knowledge differ for verb and noun learning?

(4) Do gaze patterns during listening reveal structure knowledge?

Concerning the first research question we expect word knowledge to improve over time, as the learner has been exposed to more linguistic input. With regard to the second question, we expect that structural cues will not affect the word learning process in the first stages of learning, because the language learner does not have any structure knowledge to employ in this stage. Therefore, the learner has to rely on other learning mechanisms such as cross-situational learning. However, as structure knowledge develops through exposure to the language in the structured-input condition, we expect this knowledge to boost word learning in a later stage. This would be a validation of the computational results of Alishahi and Pyykkönen (2011), who developed a computational model to investigate the syntactic bootstrapping effect in the first stages of word learning. The results of their model indicate that having knowledge of structural information does not affect performance in identifying words in the first stages of learning, but it significantly improves verb learning in later stages, when the processing of linguistic input had lead to the development of a structure knowledge base. This improvement in word identification was present for verbs only: the identification of nouns was not facilitated by structure knowledge.

The third research question investigates whether the syntactic bootstrapping effect differs for verbs and nouns. Alishahi and Pyykkönen's (2011) model showed that the beneficial effects of structure knowledge in the later learning stage apply to the learning of verbs, but not nouns. However, other experimental results indicate that structural information may also facilitate the learning process for other word categories (e.g., Fisher et al., 2006). We therefore expect that the facilitative effect of structure knowledge is present in both verb and noun learning, but that it may be stronger for verb learning.

The fourth, exploratory research question investigates whether structure knowledge is reflected in gaze patterns of language learners as they observe scenes while listening to sentences describing the scenes. As eye movements have been shown to be systematically related to auditory linguistic processing (Huettig, Rommers, & Meyer, 2011), we investigate whether eye movements during the training sessions reflect word order knowledge by showing the same pattern of fixation on the animation as in the order of the words they were processing. If this would be the case, we expect participants in the VSO condition who have acquired the word order to fixate on the action after hearing the first word, on the subject after hearing the second word, and on the object after hearing the third word.

3 Method

In order to investigate the influence of word order on word learning, an artificial language learning experiment was run with word order (VSO or random) as a between-participant variable. First, participants were exposed to a training session of 24 visual scenes accompanied by three-word sentences. All sentences are instances of a transitive construction. This training was followed by 28 multiple-choice test trials (including 16 word and 12 sentence meaning tests). This training-test sequence was performed twice, to allow a measure of linguistic knowledge at different stages in the learning process. Dependent measures consist of the number of correct answers and reaction times in the test phase, and eye movements during the training phase as an extra exploratory measure.

3.1 Participants

All 50 participants (34 females, 16 males) were undergraduate students of Communication and Information Sciences at Tilburg University, who participated for course credits. The average age was 22.4, ranging from 18 to 55. All participants were native speakers of Dutch who are in regular contact with English; 7 of them were raised multilingually (Dutch and Czech, Croatian, French, Vietnamese, Turkish, English, and one participant was raised with Dutch, English, Spanish and Papiamento). All had normal hearing and normal or corrected-to-normal vision.

3.2 Materials

Each training item consisted of an animation with a duration of 7 seconds containing 3 objects (subject, object, distractor) and an action that the subject performed towards the object (Figure 1). In Adobe Flash CS6, unfamiliar objects and actions were created for which participants would not yet have any names in their own language(s). This reduces the likelihood of interference because familiar objects already trigger certain words in the participants' previous languages. The allocation of the subject, object and distractor roles in each animation was randomly assigned in such a way that each of the 12 objects had 6 occurrences (2 as subject, object and distractor), and each of the 4 actions was displayed 6 times. Both the color and the position of the different objects and roles varied throughout the animations.
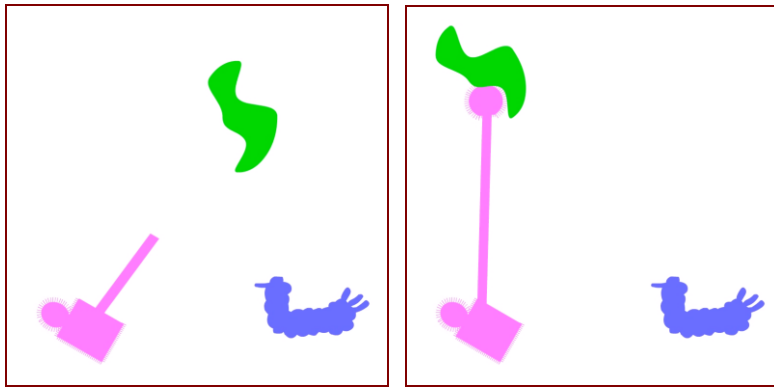
Figure 1: Two snapshots of an animation, in which the subject's 'arm' (bottom left) moves towards the object (top) and moves it around. The third object (distractor; bottom right) does not move and is not mentioned in the sentence accompanying the animation, reflecting some referential uncertainty.

The artificial words, as displayed in Figure 2, consisted of 2 syllables in a CVCV (consonant-vowel-consonant-vowel) format and were created on the basis of suggestions for pseudo-words in Dutch provided by the Wuggy application (Keuleers & Brysbaert, 2010). Existing words in possibly familiar languages (English, German, French, Spanish) and common names and abbreviations were filtered out of the word list generated by Wuggy. Then 4 pretesters listed their associations with the words; words that elicited associations for more than 2 out of 4 pretesters were also deleted. Of the remaining list, the 16 most auditorily distinguishable words were recorded in a soundproof booth by a female speaker who was trained on the Indonesian pronunciation with a synthesizer (SitePal, n.d) to prime a pronunciation that would be unfamiliar to the participants. Each word was recorded in the prosody of the first, second and third position in a Dutch declarative sentence. The separate words were cut out (in all 3 prosody positions) and the pace was reduced 30%, without affecting pitch. The words were combined to form 24 three-word sentences, separated by a short break (150 ms) between the words. There were 4 unique sets of mappings of the 16 words to the objects and actions, as certain referents, words or their mappings may be easier to recognize or learn than others (Smith & Yu, 2008). For the structured condition, words were combined in a VSO-order which is initially unfamiliar for the Dutch participants. For the unstructured condition, the positions of the words in the sentences were varied randomly.

fape	hojo	pafo	waba
foga	jase	rada	wani
haka	lowi	taga	zafi
hazi	mazi	tawi	zasu

Figure 2: List of the 16 artificial words (12 nouns, 4 verbs).

3.3 Procedure

After a calibration of the eye tracking apparatus, the experimenter left the booth and the participants read on the screen that they were going to listen to a new language that was describing what was happening on the screen. No explicit mention was made of any learning task, nor of the structure or lexical categories in the language. Then 2 training blocks, each followed by a word test and a sentence test, started. Both the training and test trials appeared in a random order and were preceded by a fixation cross (2000 ms).

The training sessions consisted of 24 trials in which participants were presented with an animation on the screen while listening to a three-word sentence that was played two times, separated by a pause (1800 ms). In the 16 trials of the word test, the participants heard one word (twice, interceded by a 1800 ms pause), while watching 4 boxes on the screen, showing 2 moving objects and 2 actions. All objects and actions were familiar from the training phase, though they were now all colored in grey. Before the start of the test, a written instruction told the participants to mouse-click on the box matching the word. In the following sentence test, the participants heard 12 of the sentences from the training phase and for each sentence were shown two training animations at the same time. They were instructed to mouse-click on the box matching the sentence.

After a recalibration, this block of training and tests was performed a second time. After finishing the second block, participants filled in a short questionnaire concerning their age, gender, language use, hearing or vision impairments, and their feeling of knowledge (how many words they thought they had learned).

Participants conducted the experiment in a soundproof booth, using a mouse and headphones. Eye tracking was conducted with an SMI Hi-Speed xView eye-tracking camera centered at the base of the screen (22 inch, 1680 x 1050 pixels). Stimulus randomization, answers and reaction times were administered using E-Prime (Schneider, Eschman, & Zuccolotto, 2002). The entire procedure lasted approximately 25 minutes.

4 Results

For each participant, the number of correct answers on the word tests and sentence tests was counted, resulting in two scores for each of the two test moments, ranging from 0 to 16 (word tests) and 0 to 12 (sentence tests). A first exploration of the test results of the second test block shows that on average, the participants had learned 6.74 of the 16 words (ranging from 1 to 13) and 8.26 out of 12 sentences (ranging from 4 to 12). Figure 3 shows a scatter plot of the number of correct answers of each participant in both the second word test and the second sentence test. Scores in the top right corner indicate a good understanding of both word and sentence meanings. Not all learners mastered the language, as indicated by scores on and below chance level (4 on the word test and 6 on the sentence test). However, scores in the top right corner show that it is possible to acquire the language in the setting of the experiment.
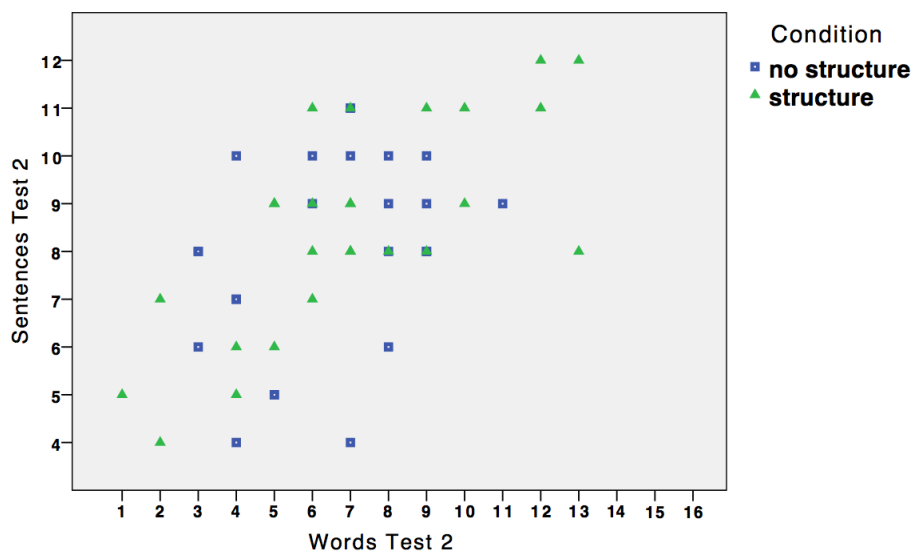

Figure 3: Scatter plot of the number of correct answers of each participant in Word Test 2 (maximum score = 16) and Sentence Test 2 (maximum score = 12).

4.1 Effects of time and word order condition

To statistically test the effect of time and word order on word learning, separate repeated measure ANOVAs were run for word tests and sentence tests, with Test Moment (2 levels) as a within-participant variable and Word Order (2 levels: VSO, random) as a between-participant variable. For both the word and sentence tests, Tables 1 and 2 display the average number of correct answers and the average reaction times respectively, divided by condition (Structure bend No Structure) and time (Test Block 1 and 2).

	Words Correct		Sentences Correct	
Condition	Block 1	Block 2	Block 1	Block 2
Structure	5.81 (1.86)	7.04 (3.28)	7.31 (1.98)	8.50 (2.20)
No Structure	5.75 (1.78)	6.42 (2.30)	6.96 (2.07)	8.00 (2.11)

Table 1: Average number of correct answers (and standard deviations) in the word tests (maximum score = 16) and sentence tests (maximum score = 12), divided by condition and test block.

	RT Word Tests		RT Sentence Tests	
Condition	Block 1	Block 2	Block 1	Block 2
Structure	6843 (1738)	6290 (1680	7768 (1747)	7288 (1928)
No Structure	7022 (1655)	6307 (1163)	7685 (1968)	7239 (1516)

Table 2: Average reaction times (RT) in milliseconds (and standard deviations) in the word tests (maximum score = 16) and sentence tests (maximum score = 12), divided by condition and test block.

There is a significant effect of test block on the number of correct answers for both the words ($F (1, 48) = 6.13$, $p = .017$, $\eta^2 = .11$) and sentences ($F (1, 48) = 6.94$, $p = .011$, $\eta^2 = .13$): knowledge of the word and sentence meanings increases through time. This effect, as visualized for both tests in Figure 4, is consistent with the expectation that linguistic knowledge improves as a function of time (i.e., amount of input).

Moreover, there is a significant main effect of test block on the reaction time for both the words ($F (1, 48) = 12.77$, $p = .001$, $\eta^2 = .21$) and sentences ($F (1, 48) = 8.17$, $p = .006$, $\eta^2 = .15$). Reactions were faster in the second block than in the first block (see Figure 5).

The results of the word tests (Figure 4, left) might indicate a trend that after the second block of training, the participants in the structured condition performed better in the word meaning test compared to those in the random condition. However, this difference does not reach significance ($F (1, 48) = 0.37$, $p = .546$). Similarly, there is no effect of structure condition on the number of sentences correct, ($F (1, 48) = 1.06$, $p = .308$). Contrary to our expectation, learners who received a structured language as input (VSO) did not learn more of the language than those who received an unstructured linguistic input.
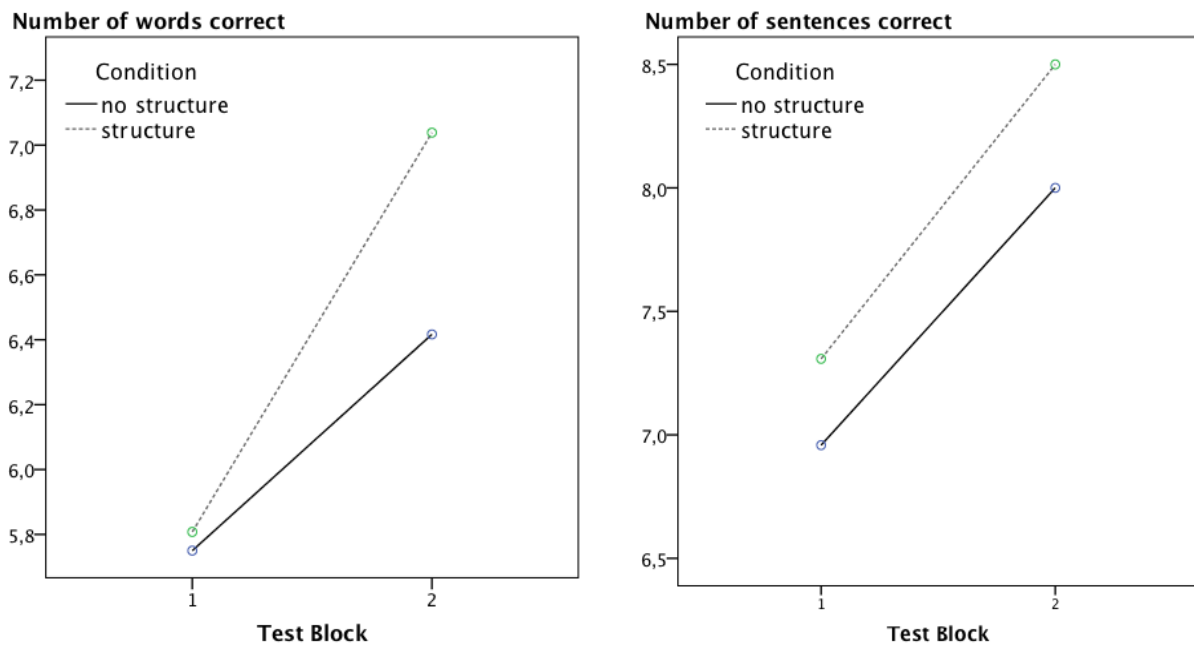
Figure 4: Graphical representations of the average number of correct answers in the word tests (left) and sentence tests (right), divided by condition and test block.
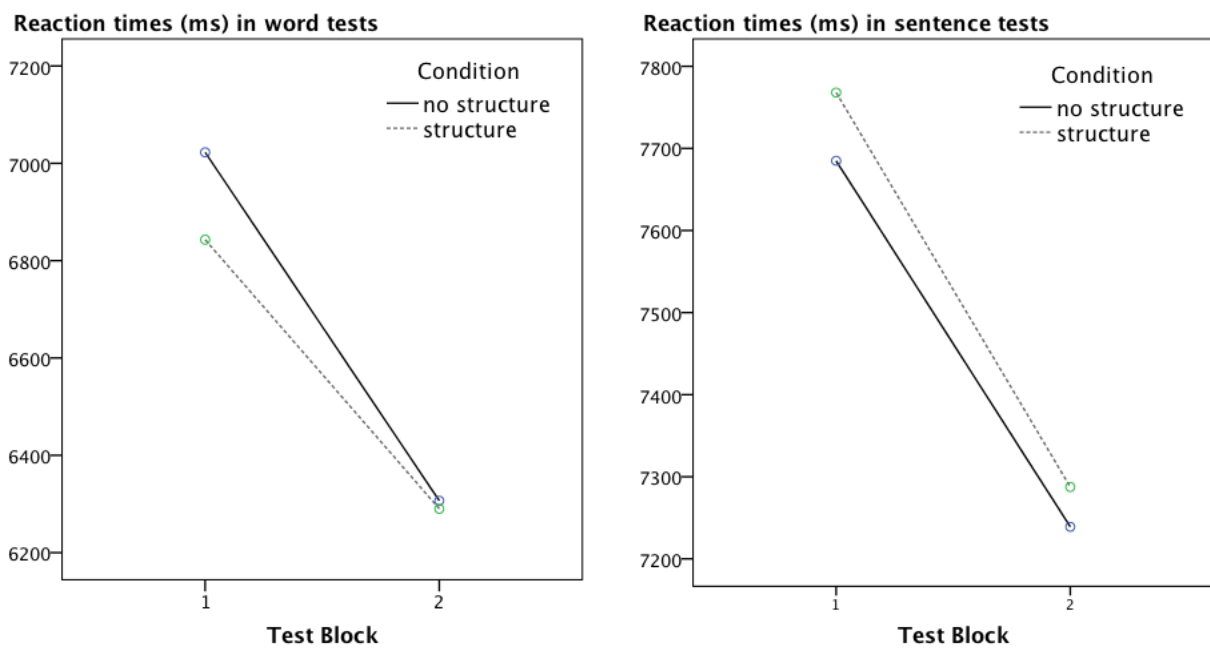

Figure 5: Graphical representations of the average reaction times in the word tests (left) and sentence tests (right), divided by condition and test block.

4.2 Differences between verb and noun learning

Table 3 displays the percentage of correct answers in the word tests for both verbs and nouns, divided by time (block 1 and 2) and condition (structure or no structure). Similar to the complete set of words, nouns show a significant effect of test block on correctness

(F (1, 48) = 5.52, p = .023, η^2 = .10), but not of structure (F (1, 48) = 0.11, p = .740). However, for verbs, neither time nor structure has a significant effect on test results (time: F (1, 48) = 0.45, p = .506; structure: F (1, 48) = 0.81, p = .372).

	Verbs		Nouns	
Condition	Block 1	Block 2	Block 1	Block 2
Structure	30.75	35.50	38.17	46.83
No Structure	28.25	29.25	38.58	43.75

Table 3: Percentage correct in the word tests for both verbs (N = 4) and nouns (N = 12), divided by condition and test block.

4.3 Fixation patterns

Eye movement analysis was conducted as an extra measure to explore whether fixation patterns during observation of scene-utterance pairs are an informative measure of implicit structure knowledge. As it takes 200 milliseconds to plan a saccade (Matin, Shao, & Boff, 1993), we expect participants in the VSO condition that have acquired the word order to fixate on the action 200 ms after hearing the first word, on the subject 200 ms after hearing the second word, and on the object 200 ms after hearing the third word (and not paying any attention to the distractor object). This fixation pattern would indicate awareness of the word order in the language.

An analysis of the number of fixations showed that, as expected, hardly any fixations are directed to the distractor object. Gaze patterns also showed that participants consistently fixated on the aspect that was showing movement in the animation, irrespective of the word order. The small amount of fixations on the distractor object is therefore likely to reflect the fact that the distractor object was not showing any movement, rather than knowledge about the sentential structure. Fixations appear to be driven by movement in the visual modality, rather than input in the auditory modality, and are therefore not suited as a way to gauge structure knowledge in the current experimental setting using moving images.

5 Discussion

With regard to the first research question, the significant effect of test block on both number of correct answers and reaction time shows that in the second learning phase, participants have more word knowledge and are able to decide faster on the meaning of words and

sentences. As expected, more time (and therefore more input) is beneficial to word and sentence knowledge in the new language.

The second research question concerns the effects of structure knowledge on word learning. As indicated, at the second test moment, the difference in word test scores of participants in the structured condition compared to those in the random condition did not reach significance. This is contradictory to our expectations based on previous research showing that structural information facilitates word learning (e.g., Fisher et al., 2006). Our results show that in contrast to the common assumption that syntactic bootstrapping is effective in word learning, structured input does not facilitate word learning immediately.

It is important here to indicate the difference between structured input (as in the VSO condition) and structure knowledge: participants in the structured condition might not have acquired the structure underlying the sentences, as part of the participants in this group still scored on or below chance in the second test block. Post-trial interviews confirm the presumption that most of the participants in the structured condition did not acquire the word order even after the second block. The structured input did not (yet) lead to structure knowledge, which explains the absence of a syntactic bootstrapping effect: learners were not able to employ structural cues to facilitate word learning. Syntactic bootstrapping as a word learning mechanism is not present in word learning immediately, because learners need time to acquire knowledge of the structure underlying the sentences in the new language before being able to employ it to facilitate word learning.

The third research question focuses on the differences in learning effects between verbs and nouns. Results of the current experiment show an effect of time (i.e., amount of input) on noun knowledge, but no such effect on verb knowledge. This may be explained by the fact that the input language contained only four verbs (versus twelve nouns); this might have been a too small amount to find any effects. Moreover, previous studies shaped our expectation that structure knowledge may be a more important cue for verb learning than for noun learning (e.g., Alishahi & Pyykkönen, 2011; Gleitman, 1990; Lee & Naigles, 2008). This at least applies to concrete nouns that refer to a physical object, as the objects in the experiment, which can also be acquired using other strategies, such as cross-situational learning. However, the meaning of verbs is relational rather than referring to a concrete physical object, and therefore verbs cannot be learned by solely observing the situations in which they are used (Gleitman, 1990). As the structural pattern in which a verb occurs provides an additional informative cue with respect to its meaning, the acquisition of verb meanings may rely more strongly on structure knowledge than the acquisition of (concrete)

nouns. The absence of structure knowledge in most of the participants in our experiment then explains the fact that verb knowledge did not improve through time, whereas noun knowledge did improve.

Finally, the fourth research question investigates gaze patterns during observation of the scene-utterance pairs. If these patterns reflect structure knowledge, this measure would allow a differentiation of participants on the basis of their structure knowledge and subsequently a comparison of the word learning performance of learners that did and did not acquire structure knowledge. However, fixations show to be linked to movement in the animation, irrespective of (knowledge of) the word order. Therefore, the eye tracking data is not an adequate measure to differentiate participants on the basis of structure knowledge. The systematic linking of eye movement patterns to auditory linguistic processing shows to be effective for still images only (Huettig et al., 2011), as these images do not contain any movements that guide the fixations, rather than the processing of the linguistic input.

6 Future work

The current results indicate several directions for future research. First of all, maximum scores of 13 out of 16 words and 12 out of 12 sentences after an experimental session of only 20 minutes show that it is possible to learn the artificial language on the basis of this bi-modal input. However, as the task of learning the structure appeared to be too difficult for the majority of participants, we suggest several changes for future work. Straightforwardly, the duration of the experiment might be lengthened, to increase the amount of input the language learners receive. It is possible that this alone allows a much higher number of the participants to learn the underlying word order, which permits an investigation of syntactic bootstrapping effects. However, the task may remain too difficult and a longer duration of the experiment may create other problems such as fatigue and loss of concentration. It may therefore be advisable to adjust the experiment in other ways.

In our future work within this experimental paradigm, we implement several changes. First, we will use a smaller vocabulary and therefore increase the word frequencies in the training trials. Second, the follow-up experiment will contain more than two (shorter) training and test blocks, which creates a more detailed image of the course of word knowledge through time. Third, the training input will contain minimally different pairs of training items to promote structure learning. Finally, we will add new fixed word order conditions containing SVO and SOV order, which are closer to the word order of the Dutch participants' first language and may therefore be acquired more easily. We expect these changes to

facilitate the acquisition of the underlying word order and therefore to allow a more detailed investigation of the onset and presence of syntactic bootstrapping processes in the first stages of learning a new language.

As indicated, fixation patterns did not offer an objective measure to distinguish learners with and without structure knowledge. Our follow-up work will therefore include a sentence production task to measure whether the participants have managed to acquire the word order. Participants are given an animation and the words corresponding to the objects and action, and are asked to combine these words into a sentence. The order of the sentences they generate will show whether they acquired the word order (if it was present). This permits separating the group of learners with structure knowledge from the group without structure knowledge, and comparing their word learning development in light of the syntactic bootstrapping hypothesis.

Acknowledgments

References

Alishahi, Afra & Pyykkönen, Pirita (2011): 'The onset of syntactic bootstrapping in word learning: Evidence from a computational study'. In: L. Carlson, C. Hoelscher, & T. Shipley (Eds.): *Proceedings of the 33rd Annual Conference of the Cognitive Science Society (CogSci)* (pp. 587–592). Austin, Tx: Cognitive Science Society.

Brown, Roger W. (1957): 'Linguistic determinism and the part of speech'. In *Journal of Abnormal and Social Psychology* 55, 1–5.

Carey, Susan (1978): 'The child as word learner'. In: M. Halle, J. Bresnan, & G. A. Miller (Eds.): *Linguistic Theory and Psychological Reality* (pp. 264–293). Cambridge, MA: MIT Press.

Dryer, Matthew S. (2008): 'Order of subject, object and verb'. In: M. S. Dryer & M. Haspelmath (Eds.): *The World Atlas of Language Structure Online* (Chap. 81). Retrieved from http://wals.info/chapter/81

Fisher, Cynthia (1996): 'Structural limits on verb mapping: The role of analogy in children's interpretations of sentences'. In *Cognitive Psychology* 31, 41–81.

Fisher, Cynthia, Klingler, Stacy L., & Song, Hyun-joo (2006): 'What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions'. In *Cognition* 101, B19–B29.

Folia, Vasiliki, Uddén, Julia, De Vries, Meinou, Forkstam, Christian, & Petersson, Karl M. (2010): 'Artificial language learning in adults and children'. In *Language learning* 60(s2), 188–220.

Gertner, Yael, Fisher, Cynthia, & Eisengart, Julie (2006): 'Learning words and rules: Abstract knowledge of word order in early sentence comprehension'. In *Psychological Science* 17(8), 684–691.

Gillette, Jane, Gleitman, Henry, Gleitman, Lila, & Lederer, Anne (1999): 'Human simulations of vocabulary learning'. In *Cognition* 73, 135–176.

Gleitman, Lila (1990): 'The Structural Sources of Verb Meanings'. In *Language Acquisition* 1(1), 3–55.

Huettig, Falk, Rommers, Joost, & Meyer, Antje S. (2011): 'Using the visual world paradigm to study language processing: A review and critical evaluation'. In *Acta Psychologica* 137, 151–171.

Kachergis, George, Yu, Chen, & Shiffrin, Richard M. (2010): 'Cross-Situational statistical learning: Implicit or intentional?' In: S. Ohlsson & R. Catrambone (Eds.): *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2362–2367). Austin, TX: Cognitive Science Society.

Keuleers, Emanuel & Brysbaert, Marc (2010): 'Wuggy: A multilingual pseudoword generator'. In *Behavior Research Methods* 42(3), 627–633.

Koehne, Judith & Crocker, Matthew W. (2010): 'Sentence processing mechanisms influence cross-situational word learning'. In: S. Ohlsson & R. Catrambone (Eds.): *Proceedings of the 32nd Annual meeting of the Cognitive Science Society* (pp. 2458–2564). Austin, TX: Cognitive Science Society.

Koehne, Judith & Crocker, Matthew W. (2011): 'The interplay of multiple mechanisms in word learning'. In: L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.): *Proceedings of the 33rd Annual meeting of the Cognitive Science Society* (pp. 1930–1936). Austin, TX: Cognitive Science Society.

Lee, Joanne N. & Naigles, Letitia R. (2008): 'Mandarin learners use syntactic bootstrapping in verb acquisition'. In *Cognition* 106, 1028–2037.

Magnuson, James S., Tanenhaus, Michael K., Aslin, Richard N., & Dahan, Delphine (2003): 'The time course of spoken word learning and recognition: Studies with artificial lexicons'. In *Journal of Experimental Psychology: General* 132(2), 202–227.

Matin, Ethel, Shao, K. C., & Boff, Kenneth R. (1993): 'Saccadic overhead: Information-processing time with and without saccades'. In *Perception & Psychophysics* 53(4), 372–380.

Mintz, Toben H. (2002): 'Category induction from distributional cues in an artificial language'. In *Memory & Cognition* 30(5), 678–686.

Papafragou, Anna, Cassidy, Kimberley, & Gleitman, Lila (2007): 'When we think about thinking: The acquisition of belief verbs'. In *Cognition* 105, 125–165.

Piccin, Thomas B., & Waxman, Sandra R. (2007): 'Why nouns trump verbs in word learning: New evidence from children and adults in the human simulation paradigm'. In *Language Learning and Development* 3(4), 295–323.

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

Quine, Willard (1960): *Word and object*. Cambridge, MA: MIT Press.

Schneider, Walter, Eschman, Amy, & Zuccolotto, Anthony (2002): *E-prime 2.0 user's guide.* Psychology Software Tools, Inc., Pittsburg.

SitePal (n.d.): SitePal text-to-speech demo. Retrieved from http://www.oddcast.com/home /demos/tts/tts_example.php?sitepal

Smith, Kenny, Smith, Andrew D. M., & Blythe, Richard A. (2011): 'Cross-situational learning: An experimental study of word-learning mechanisms'. In *Cognitive Science* 35, 480–498.

Smith, Linda & Yu, Chen (2008): 'Infants rapidly learn word-referent mappings via cross-situational statistics'. In *Cognition* 106, 1558–1568.

Yu, Chen & Smith, Linda B. (2011): 'What you learn is what you see: Using eye movements to study infant cross-situational word learning'. In *Developmental Science* 14(2), 165–180.