Abstract

An experiment examined the role of word order and case inflection in the interpretation of Czech transitive sentences by children and adults. Participants listened to simple transitive sentences while watching picture pairs with the same characters performing the same action, but with the opposite assignment of the subject and object roles. Sentences varied in word order (SVO, OVS), and in whether the initial noun was case-ambiguous or not. Word order and case inflection appear to have only weak, if any, immediate influence on 3-year-olds sentence interpretation. The performance in 5-year-olds revealed sensitivity to both word order and inflection, similar to the performance of adults. However, 5-year-olds appear to be driven by word order more than adults. Surprisingly, processing of OVS sentences in adults was not different from SVO sentences as long as the initial noun was unambiguously inflected for case. The results are discussed from the viewpoint of the competition model, and a processing basis for some of the children’s deficits is proposed.

Keywords: Language acquisition; case; word order; competition model; preferential looking; visual world.

One of the most important steps in the process of language acquisition is learning the specific ways in which the child’s native language marks basic syntactic relations, such as those of subject and object. The acquisition of subject and object relations has been the topic of much research both in children’s production and comprehension (e.g. Akhtar & Tomasello, 1997; Hirsh-Pasek & Golinkoff, 1996). Among the linguistic devices used to mark subject and object relations, word order and case morphology play a prominent role. The present study examines the use of word order and case forms to interpret sentences during an on-line preferential looking task in Czech children and adults.

Slobin and Bever (1982) examined relative importance of word order and nominal inflection in children’s sentence interpretation in Serbo-Croatian and Turkish, and the effect of word order in English and Italian. In English and Italian, children showed a tendency to interpret the sentence-initial noun as the subject. This tendency became stronger with increasing age, reflecting development towards the adult-like performance. On the other hand, even the youngest Turkish-speaking children showed strong reliance on nominal inflection. In Serbo-Croatian, the pattern of performance changed with age: younger children appeared to rely on word order more than on case forms. With age, their interpretations were increasingly driven by case form which corresponds to the adult performance.

The present study was designed to test whether Czech children are similar to Croatian children in Slobin’s and Bever’s study (1982). It can be viewed as a study testing the use of different linguistic cues in comprehension, as defined by the competition model (Bates, McNew, MacWhinney, Devescovi, & Smith, 1982; Bates et al., 1984), in Czech. At the same time, it broadens the scope of competition model studies by focusing on the on-line aspects of sentence interpretation.

The objective of the study was to test children’s ability to use grammatical cues, and evaluate the relative importance of word order and case forms in children’s sentence comprehension. Unlike the act-out or picture matching tasks that were used in most research on competition model, the present study uses the on-line preferential looking method (“looking-while-listening”; Fernald, Swingley, Weinberg, & McRoberts, 1998). This provides an opportunity to study whether children perform sentence interpretation incrementally, i.e. whether they interpret the grammatical cues immediately.

Slobin’s and Bever’s (1982) data on the development of Serbo-Croatian can be used to derive expectations about Czech, as the languages are related and structurally similar. Just like in Croatian, the case form is a strong cue in Czech but it cannot be used exclusively because many nouns have the same form in the nominative and accusative cases. The word order is available any time the nouns are expressed explicitly in a sentence, but it is not an absolutely reliable cue. Even though the SVO word order is considered standard (Šgall, Hajičová, & Panenková, 1986) in the absence of strong information structure constraints, practically any permutation is possible under appropriate pragmatic conditions. Children in Slobin’s and Bever’s study initially relied mainly on word order, and only later shifted to case forms as the main cue. Czech children should be similar if the development in Czech and Croatian follows the same path. It should be noted that the preference of 2- and 3-year-olds for word order in Slobin’s and Bever’s study was above chance but fairly weak. The present study can suggest whether this was due mainly to the off-line act-out task, or whether children have more serious limitations in the use of grammar.

Experiment

The general question in the present study was whether children understand verbal picture descriptions differing only in their assignment of subject/object roles, and how different they are from adults in this task. Participants saw pictures with the same characters performing the same action but with the opposite assignment of agent/patient roles to the charac-
Method

Participants The participants in this study were 53 preschool children in the age range 30 to 71 months, and a group of 20 adults. All children were monolingual Czech and were from Prague or vicinity. They were recruited using fliers placed in doctors’ offices or preschools and via on-line ads. The parents received ca. $10 as a reimbursement for the time spent with the lab visit. All children were typically developing without any major health or developmental problems, and with normal or corrected vision.

The children were divided in two groups. The group of 5-year-olds consisted of 26 children (9 girls), with mean age 59 months (range 48–71, SD=6). The group of 3-year-olds consisted of 27 children (15 girls), with mean age of 39 months (range 30–47, SD=6). The group of adults consisted of volunteers and university students who received credit for their participation.

Stimuli The whole experiment consisted of 18 trials. Of those trials, 16 are directly relevant for the research questions discussed in this paper (see Table 1); the remaining two focused on sentences without any inflectional cues. All trials shared the same format. In each trial, participants were simultaneously presented with a pair of pictures. Both pictures showed the same action involving the same cartoon characters, but the agent/patient roles of characters were inverted. While watching the picture pair, participants heard a recorded sentence with meaning matched to one of the pictures.

For each picture pair, there were four possible sentences differing in the assignment of agent/patient roles and word order. Two sentences matched each picture of the pair. These two sentences differed in word order, with one being subject-initial and the other being object-initial. All sentences involved animate participants, animals or humans, so that the pictured events were approximately equally (un)likely (e.g. “The fox is pushing a kitty” and vice versa). The four versions of each sentence were used to develop four protocol variants.

Table 1: Stimulus sentences used in the experiment.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stimulus structure</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Sub\textit{junambig} – Verb – Ob\textit{junambig}</td>
<td>4</td>
</tr>
<tr>
<td>B</td>
<td>Sub\textit{junambig} – Verb – Ob\textit{junambig}</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>Ob\textit{junambig} – Verb – Sub\textit{junambig}</td>
<td>4</td>
</tr>
<tr>
<td>D</td>
<td>Ob\textit{junambig} – Verb – Sub\textit{junambig}</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>Sub\textit{ambig} – Verb – Ob\textit{junambig}</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>Ob\textit{ambig} – Verb – Sub\textit{junambig}</td>
<td>2</td>
</tr>
</tbody>
</table>

Twelve items addressed the first research question, i.e. compared the subject-initial (A,B) and object-initial (C,D) unambiguous sentences. Four items examined the second research question, comparing sentences with initial case-ambiguous noun that were disambiguated towards the subject-initial (E) or object-initial (F) reading.
Apparatus and Procedure A PC computer with dual screen and speakers was used to present the pictorial and auditory stimuli. Two 19" LCD screens were placed on a table, and a pair of speakers connected to the computer was placed between the screens. The distance between midpoints of the monitors was 68 cm and the participants viewed them from the distance of ca. 1.25 m. The camera was situated behind the table. We used SONY digital camera with PAL resolution 720 × 576 with 25 fps.

In each trial, the stimulus pictures appeared first; 2 s later, the stimulus sentence presentation started. The picture disappeared 5 s after the sentence onset. The participants were asked to point to the picture matching the presented sentence. The pointing responses were not strictly required and if a child failed to follow the instruction, she was not prompted to. For this reason, the pointing data were not evaluated.

The stimulus presentation was controlled via PXLab script (Irtel, 2007). Dual-head video card was used to synchronize the two-monitor presentation. The camera recordings were synchronized with the experiment script by an alternative audio track. The script presented different signals into each channel of the stereo signal. One channel presenting the stimulus sentences was connected to the speakers. The other channel contained short beeps marking the sentence beginning and the end of picture presentation; this signal was sent to the camera’s audio input.

Coding Videotapes of the participants were digitalized and coded frame-by-frame for gaze direction. For each trial, 6 seconds of the recording were coded, beginning 1 s before the onset of the stimulus sentence. This provided 150 frame codes per trial, i.e. 2400 per participant for the trials analyzed here. For each frame, the observed gaze direction was coded. Four codes were possible: left, middle (between the screens), right, or other, which included gazes away from the screens and frames where the gaze direction was not apparent, such as during blinks. The data were coded by a research assistant who also administered the experiments. The coder was thus not totally blind to the content of the stimuli. However, the audio track in the recordings did not contain the stimulus sentences, only the beeps marking the sentence onset and the end of picture presentation. Therefore, the coder did not routinely know which particular stimulus she was coding.

Analysis The structure of Czech grammar makes it difficult to fit the complete data into a factorial design. A series of analyses was thus performed separately for each research question and each age group. In general, in the analysis we used binomial (logistic) mixed models with subjects and stimulus sentences as crossed random factors (Baayen, Davidson, & Bates, 2008). The binary dependent variable in the models was the gaze direction toward the target picture. The models estimated the probability of fixing the target picture as a function of experimental condition and time interval.

For the analysis we used time units defined relatively to the start and end of the words in the audio recordings. We used four time intervals: three corresponding to the words in the stimulus sentence, and the fourth representing the 400 ms interval after the end of the last word. To measure the gaze behavior initiated during these time intervals, the gaze data were shifted by 200 ms compared to the audio recordings. The 200 ms delay was used as an estimate of the saccade latency (e.g. Melcher & Colby, 2008).

The analyses of preferential looking data must take into account the looking preferences of children before the beginning of the data presentation. Therefore, rather than comparing the looking data to the chance level of 50 %, the analyses compared the probability of target picture fixation during the sentence presentation against a baseline that was minimally influenced by the sentence content. This baseline comprised gazes initiated during the presentation of the first noun of the target sentence, i.e. the video frames starting 200 ms after the sentence onset and ending 200 ms after the first word offset. Analyses used planned orthogonal contrasts to compare the results in the subject-initial condition on the second, third and fourth intervals to the baseline, and in the object-initial condition to the results in the subject-initial condition during the same interval.

Results

Unambiguous Sentences

The proportion of 3-year-olds’ looks to the target picture during the baseline and the subsequent presentation intervals is shown in the Figure 1 (top left). The graph suggests a small increase in the proportion of looks towards the target picture in the subject-first condition in the third interval.

The analysis revealed no significant difference between the subject-first and object-first condition during the baseline period. In the subject-first condition, the probability of looks towards the target during the second interval was not significantly different from the baseline. During the third interval, the probability was significantly higher (z = 3.667, p < 0.001). In the object-first condition, the probability of looks towards the target picture during the second and third interval was significantly lower than in the subject-first condition (second interval z = −2.425, p < 0.05, in third interval z = −3.012, p < 0.01). The results from 3-year-olds suggest that the children have some ability to interpret sentences with canonical subject-first word order but show no such ability in the noncanonical object-initial sentences. Actually, the significant difference between conditions on the second interval (gazes initiated during the second word) indicates that children tend to shift away from the target picture after they hear the first word in the sentence. Such a pattern indicates that 3-year-olds rely primarily on word order and ignore the inflectional cues. Overall, even though the results from the youngest group of children show significant changes in looking behavior related to the sentences presented to children, the changes are small and may reflect a successful performance in a small portion of the group only.
In five-year-olds, the pattern of results is somewhat more clear, as Figure 1 (top middle) suggests. There is no significant difference between the conditions during the first, baseline interval. In the subject-first condition, the probability of looks towards the target increases in the second interval ($z = 4.396, p < 0.001$) and remains significantly higher than baseline during the third ($z = 6.304, p < 0.001$) and fourth ($z = 5.648, p < 0.001$) intervals. In the object-first condition, the proportion of looks towards the target picture is significantly lower than in the subject-first condition during the second ($z = -3.966, p < 0.001$) interval, and higher in the fourth ($z = 2.388, p < 0.05$) interval with no difference in the third interval. The pattern indicates that children rely more on word order during the initial interpretation of sentences: the gazes initiated during the second word tend to shift to the picture in which the character mentioned sentence-initially plays the agent role. However, 5-year-olds appear to differ from 3-year-olds in that they switch their preference and look to the correct target picture in the third interval with about the same probability in the subject-initial and object-initial condition. Compared to 3-year-olds, this reveals better ability of 5-year-olds to use inflectional cues and deal with non-canonical word orders. The decrease in the second interval might be caused by higher cognitive demands of processing a noncanonical object-initial sentence.

In adults (see Figure 1, top right) there is no significant difference between both conditions in any inspected interval. In the subject-first condition the probability of looks towards the target picture increases in the second interval ($z = 2.435, p < 0.05$) and further increases during the third ($z = 11.163, p < 0.001$) and fourth interval ($z = 13.850, p < 0.001$). The performance of adults suggests that the inflectional cues present in the first word of an unambiguous sentence are sufficient to find the target picture. Although the object-first sentences are less common and pragmatically marked in Czech, the results show that the adults use the inflectional cues and are not mislead by the non-canonical word order.

**Temporarily Ambiguous Sentences**

The proportion of 3-year-olds’ looks to the target picture while listening to the sentences with first ambiguous noun is shown in Figure 1 (bottom left). The graph suggests the overall performance decrease with no apparent trend towards the target picture in both conditions.

The analysis revealed no significant differences in the baseline interval. In the subject-first sentences there was a significant decrease compared to the baseline during the second interval ($z = -2.228, p < 0.05$), but the difference is no longer significant in the third and fourth interval. In the object-first sentences, children looked towards the target picture more frequently during the baseline, even though this difference was not significant.

This nonsignificant advantage of object-first sentences was preserved over the second and third interval but it changed in the fourth interval. The proportion of looks towards the target picture was lower than predicted by the remaining predictors.
The results indicate generally low performance in 3-year-olds when they are presented with an ambiguity in the sentence-initial noun. This condition is probably too confusing and the case inflection on the sentence-final noun does not help children to determine the target picture. There is no evidence that 3-year-olds are capable of using word order or case inflection to drive their sentence interpretation while listening.

The results of 5-year-olds (see Figure 1, bottom middle) show almost a symmetrical pattern. In the subject-first condition there is a significant preference for the target picture starting in the second interval \((z = 2.961, p < 0.01)\) and further increasing in the third \((z = 9.279, p < 0.001)\) and fourth interval \((z = 8.834, p < 0.001)\). The baseline preference for target picture in the object-first condition is significantly higher \((z = 2.667, p < 0.01)\), but in the subsequent intervals it markedly decreases and is lower than the preference in the subject-first condition (second interval \(z = -8.594, p < 0.001\), third \(z = -18.109, p < 0.001\), fourth \(z = -14.191, p < 0.001\)). The pattern of results in 5-year-olds suggests that the children use word order in the absence of inflection. The nominative inflection on the sentence-final noun does not appear to change the interpretation within the observed time, even though Figure 1 suggests a tendency to shift towards the target in the fourth interval.

The results of adults suggest that both cue types are used (see Figure 1, bottom right). In the analysis we observed small but significant baseline preference for the target picture \((z = 2.215, p < 0.05)\). In the subject-initial sentences the probability of looks towards the target picture increases in the third \((z = 6.733, p < 0.001)\) and in the fourth interval \((z = 7.483, p < 0.001)\). In the object-initial condition the preference for the target picture decreases in the following intervals (in second interval \(z = -6.088, p < 0.001\), in third \(z = -10.123, p < 0.001\)). In the fourth interval the difference between conditions is smaller but still significant \((z = -4.022, p < 0.001)\). The performance of adults shows that after the initial ambiguity they use the word order cue, but this assumption is revisited when they hear the third word (the unambiguous noun) and finally they shift towards the target picture.

**Discussion**

The data presented in our study suggest that some reorganization takes place in the language processing system between the child and adult age. Both adults and 5-year-olds are influenced by case inflection and word order but to a somewhat different degree and at different points during the comprehension process. The results also suggest that 3-year-olds have limited ability to process the grammatical structure of sentences on-line, at least if they have to rely exclusively on formal grammatical devices.

The effects of stimulus sentences on three-year-olds were rather weak. The younger children shifted towards the incorrect pictures if the first word was in accusative: this would suggest reliance on word order as the main cue, at least in the initial stages of processing. Also, they showed some evidence of shifting towards the target picture while listening to subject-first sentences with unambiguous initial noun, which would suggest some sensitivity to word order. However, the results from sentences with ambiguous initial noun are not consistent with any useful strategy. This contrasts with offline studies that show evidence of sensitivity to grammatical cues in children as young as 25 months (Slobin & Bever, 1982). The similarity of the pictures and the limited preview time in our study may have contributed to the confusion in children. However, comparing the present results with typical preferential looking studies suggests some deficit in the quick on-line processing (cf. Hirsh-Pasek & Golinkoff, 1996; Fernandes, Marcus, Nubila, & Vouloumanos, 2006; Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008). Apparently, young children around 3 years of age need time to process the grammatical structure of the incoming language.

The data from 5-year-olds suggest good sensitivity to case cues. In subject-initial sentences with unambiguous initial noun, children start shifting to the correct picture while processing the second word in the sentence. This suggests immediate processing of the nominative cue. The object-initial sentences with unambiguous initial noun appear to be processed more slowly, which indicates that processing non-canonical word order is more difficult. On the contrary, adults appear to rely on case forms exclusively and do not show any difference between sentences with subject-initial and object-initial word orders. This is an interesting contrast with other studies that suggested problems with non-canonical word orders (Ferreira, 2003; Scheepers & Crocker, 2004). The inflection appears to be so strong for Czech adults that it overrides the word-order heuristics.

In sentences with ambiguous initial noun, both adults and 5-year-olds initially interpret the first noun as the subject. After listening to the disambiguating sentence-final noun, greater percentage of adults than children shift to the target picture, and the shift occurs faster in adults. The results suggest almost immediate effects of the linguistic input in adults. In 5-year-olds, however, there is actually no evidence of recovering from the incorrect interpretation.

To summarize, children in the younger group in this study showed no on-line sensitivity to formal grammatical cues under investigation. In five-year-olds and adults, both case morphology and word order play a role. When case cues are available, they override word order, but this happens faster in adults than in children. When case cues are not available, both adults and 5-year-olds use word order. However, adults can change their interpretation very fast when confronted with the disambiguating inflection. Inflection is thus a stronger cue in adults than in children, although it plays an important role by 5 years of age. Word order is weaker in adults but still strong in the absence of inflectional cues. This
pattern is consistent with other findings from studies based on the competition model, especially with Sokolov’s (1988) proposal that children first rely on cues that are highly available, and only later switch to highly reliable, such as case inflection.

While the competition model offers good tools to characterize the children’s performance, the preference for word order in children might also be interpreted purely on a processing basis as a result of children’s limited ability to revise initial parsing commitments (cf. Trueswell et al., 1999). The word order is the first available cue: while children are listening to the initial noun and retrieving it from the lexicon, they can start constructing a syntactic structure for the sentence. It is likely that by the end of the presentation of the first word, a commitment has already been made to treat the word as the sentential subject. This may be similar in children and adults, but children take longer to revise the initial assignment. This is especially apparent in temporarily ambiguous sentences, but the delay in children’s shifts towards the target picture in unambiguous sentences also points in this direction.

The interpretation based on incrementality of processing is not, as such, in conflict with the competition model. The authors of the model counted processing cost to the relevant properties of grammatical cues (Bates et al., 1984), and other features of the processing mechanism can be reflected in the model as well. However, the interpretation emphasizes that the strength and validity of grammatical cues is not a function of linguistic input only but can be due to processing limitations at a particular developmental level. This view opens an interesting area for further research: if the early preference of word order in sentence interpretation is due to slow and unreliable revisions, the difficulty with non-canonical word orders should disappear in sentences where case is marked pre-nominally.

Overall, the present experiment demonstrated that 5-year-olds, but not 3-year-olds, immediately use both word order and case morphology to identify pictures that correspond to auditorily presented sentences. Children around 3 years of age show only weak evidence of grammatical knowledge when confronted with reversible action pictures: it remains open whether this is due to the lack of corresponding grammatical knowledge or to a difficulty in processing the pictures and sentences in an on-line task. The performance of 5-year-olds reveals that revisions of initial parsing decisions in children take longer than in adults, but the basic grammatical cues for interpreting transitive sentences are in place.

Acknowledgment

The study was supported by the Grant Agency of the Academy of Sciences of the Czech Republic, grant no. KJB700250801 "Processing of agreement and case forms in children".

References


