THE ROLE OF PROSODY ON THE PERCEPTION OF WORD-ORDER DIFFERENCES BY 14-MONTH-OLD GERMAN INFANTS

Ricardo Augusto Hoffmann Bion, Barbara Höhle, Michaela Schmitz
University of Potsdam
ricardobion@gmail.com, hoehle@rz.uni-potsdam.de, mschmitz@ling.uni-potsdam.de

ABSTRACT

Fourteen-month-old German infants can distinguish between isolated non-finite verb-noun and noun-verb verb-phrases [1]. This study investigated whether prosody is the cause of this differential attention. The prosodic content of 144 verb-phrases was manipulated, and the impact of this manipulation on infants' perception was examined. Our results show that, even though prosody does influence infants' perception, the segmental content of phrases (i.e., morphological markers) is also crucial for German infants' early sensitivity to word-order differences within verb-phrases.

Keywords: prosodic bootstrapping, first language acquisition.

1. INTRODUCTION

At a very early age, children already show knowledge of the way words can be combined in the language they are learning. Word-order mistakes, for instance, are hardly ever made, and children's first two-word utterances are canonical from their onset on [2]. In addition, even before children start producing multi-word utterances, word-order already influences language comprehension [3]. The way this error-free learning takes place has been of fundamental concern for language acquisitionists, and it has been hypothesized that prosody might play an important role in this process [1, 4, 5].

Specifically, infants might start to acquire language-specific word-order rules by paying attention to stress prominence within phrases [4, 5]. This claim is motivated by the existence of a "universal" correlation between word-order regularities and the rhythmic properties of languages. Head-initial languages such as English tend to have its prosodically most prominent element on the left edge of phrases, while the opposite pattern is observed in head-final languages such as Turkish. This hypothesis would be strengthened if empirical studies demonstrated that infants are sensitive to word-order differences within phrases at a very early age, and that this sensitivity is motivated by infants' attention to prosody.

Interestingly, recent studies indicate that 14-month-old German infants are already sensitive to word-order differences within verb-phrases [1]. Infants clearly prefer phrases that follow the order noun-verb (NV) to phrases that follow the verb-noun (VN) order instead (e.g., "Zeitung lesen" vs. "lesen Zeitung"). However, it is still necessary to investigate whether prosody is the cause of German infants' early differential attention to NV and VN phrases.

Therefore, the present study aimed at investigating the role of prosody on the perception of word-order differences by 14-month-old German infants. Towards this end, the original experiment of [1] was replicated, but their phrases were acoustically edited, so that they contained conflicting prosodic and segmental information. That is, NV phrases contained the prosodic cues of VN phrases, and vice-versa. If the results of [1] were caused uniquely by infants' biases towards the prosodic content of NV phrases, when tested on our edited phrases, infants should prefer phrases with VN order and NV prosody, instead of VN phrases, as found by [1]. If different results are found, segmental information must also play a role in German infants' early sensitivity to word-order differences within phrases.

2. THE EXPERIMENT

2.1. Participants

Twenty-eight children of monolingual German-speaking parents were tested. All children were born full term, were healthy, and had no hearing, neurological or developmental problems (following [6]). Eleven of these children were discarded from the data analysis for failure to

complete the test. The remaining group comprised 7 girls and 10 boys, with ages ranging from 13 months and 7 days to 14 months and 30 days.

2.2. Original Stimuli

The same stimuli from [1] were used in the present study. In [1], 72 different NV and VN phrases were created. These phrases contained the same lexical material with different ordering of elements (e.g., "Zeitung lesen" vs. "lesen Zeitung"). For the NV condition, phrases were spliced out of clauses such as "*weil* die Arbeiter Zeitung lessen;" for the VN condition, they were spliced from clauses such as "*denn* die Arbeiter lesen Zeitung." Both clauses are grammatical and have the same English translation (in this case, "because the workers read newspapers"). All nouns and verbs were bisyllabic; all verbs were in the infinitive, and all nouns ended in the derivational affix –ung.

2.3. Acoustic analysis of the phrases

In order to evaluate whether there were prosodic differences between the NV and VN phrases, all the original phrases from [1] were submitted to an acoustic analysis. The parameters analyzed were duration, maximal pitch, and maximal smoothed intensity of the stressed syllable of the first and second word of each phrase. Paired-samples T tests were used to evaluate whether the difference between the VN and NV condition reached statistical significance. Our results indicate that the NV and VN phrases contained clear differences in prosody. These differences were present in all the acoustic cues evaluated, except for the duration of the stressed syllable of the second words of phrases, probably due to final lengthening in sentence final position. *Figure 1* shows the means and standard deviations of the prosodic cues of the stressed syllable of the first and second words of the NV and VN phrases. Asterisks signal statistically significant differences.

Table 1. Prosodic characteristics of phrases.

acoustic cue and word	VN	NV
* pitch 1st word (Hz)	321 (93)	245Hz (96)
* pitch 2nd word (Hz)	215 (81)	303Hz (77)
* amplitude 1st w. (dB)	77 (12)	80 (16)
* amplitude 2nd w. (dB)	78 (15)	73 (11)
* duration 1st w. (ms)	310 (73)	347 (84)
duration 2nd w. (ms)	345 (72)	325 (84)

2.3. Acoustic manipulation of the phrases

The main objective of the acoustic manipulation of the stimuli was to make phrases contain conflicting prosodic and syntactic information. Thus, NV phrases should contain the prosodic information of VN phrases, and vice versa. Towards this end, the verb-phrases from [1] had their prosodic information measured, and then modified with the PSOLA algorithm.

The median values for the duration and maximum amplitude of the syllables of the NV and VN conditions were calculated. As for pitch, 10 values per syllables were measured, and a normalized pitch contour was obtained by getting the median of these values for the NV and VN phrases. This normalized contour and the median duration values for the verb-phrases appear in *Figures 1* and *2*.

Figure 1. Median pitch contour and syllable duration for the NV condition

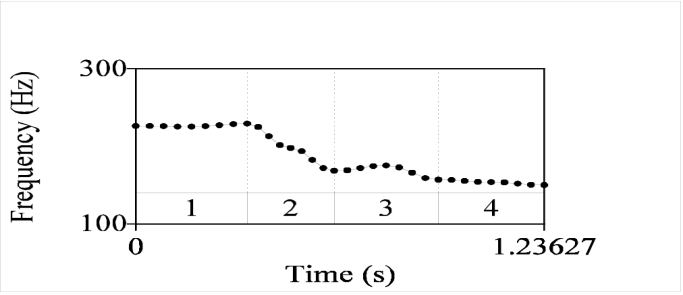
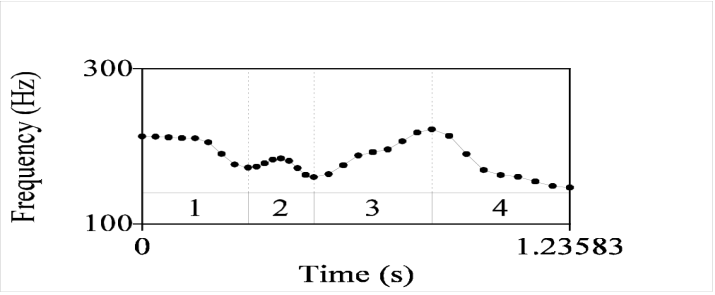


Figure 2. Median pitch contour and syllable duration for the VN condition



As a next step, the original NV phrases were edited so that their individual syllables contained the prototypical pitch contour of the VN phrases. The duration and maximum amplitude of the syllables of the NV phrases were also replaced by the median duration and amplitude of the syllables of the VN phrases. The opposite was done with the VN phrases, with syllables containing the median

pitch contour, duration, and intensity of the NV sentences. After this manipulation, the phrases were organized in blocks. As in [1], for each experimental condition, 8 trials were constructed, each with 9 phrases with pauses of 600 ms between phrases. Each trial had the duration of around 16 s.

2.4. Apparatus

During the experiment, the infant was seated on the lap of her caregiver on a chair in the center of a booth measuring 1,90 by 1,78 by 2,0 meters. In front of the infant, 85 cm above the floor, a green lamp was fixed. On the left and right sides of the booth, red lamps were fixed at the same height as the green lamp. Loudspeakers were placed behind the red lamps.

The experimenter was located in a room outside the testing booth. The equipment in this room consisted of a computer, a video monitor, and a response box with two buttons. The movements of the infant were projected to the video monitor by a camera inside the testing booth. By looking at the video screen and pressing the buttons on the response box, the experimenter could control the presentation of the acoustic stimuli, the flashing of the red and green lamps, and start and stop the timer for measuring the child's head turn direction.

Both the experimenter and the caregiver were "deaf" in regards to the sound being played in the testing booth. The sound played in the booth was not transmitted to the control room, and the caregiver was wearing noise-canceling headphones with masking music.

2.5. Design and Procedure

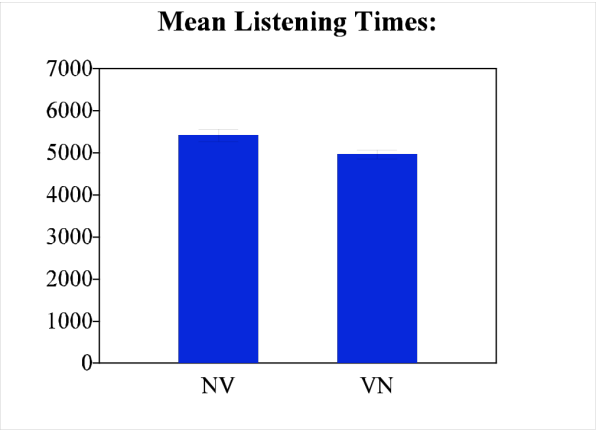
The design of the experiment followed the headturn preference paradigm. As a warm up, the child listened to one trial on each possible presentation site (i.e., the red lamp on the left, and the red lamp on the right).

Then, the infant was confronted with 14 trials (7 trials with NV phrases and 7 with VN phrases), in two different orders. The trials started when the experimenter pushed the button on the response box, causing the green lamp in front of the room to start flashing. The response box was connected to the computer, and the pressing of the buttons controlled a timer that measured the child's head direction. Once the infant looked at this lamp, the experimenter pushed a button, causing the green lamp to stop flashing, and one of the red lamps on the sides to start flashing. Once the infant looked at this red lamp, the experimenter pressed another button, and the acoustic stimuli started to be played by the loudspeaker behind the flashing lamp. The experimenter then should push a button in case the infant looked away from the presentation side, stopping the timer. Then, in case the infant looked back at the presentation side within 2 seconds, the button should be pressed to restart the timer. If the infant looked away for more than two seconds, the trial was stopped, and another trial started. In case the infant held his/her head towards the presentation side constantly, the entire trial was presented. The maximum looking time per trial could be the entire file, the sum of several head turns, or a single head turn. This procedure lasted around 6 minutes.

2.6. Results

The listening times for each trial of both conditions were added and a mean listening time for each condition was calculated for each child. As can be seen in *Figure 3*, on the average, the infants listened to each trial with NV phrases for 5422 ms, and to each trial with VN phrases for 4966 ms (the trials had around 16 sec.). A t-test for dependent samples indicated this difference not to be significant ($t(16) = 1.37$; $p > 0.19$). In contrast, a statistically significant difference was found by [1], when the stimuli with the original prosodic content was played to the infants.

Figure 3. Mean listening times of the whole sample



Therefore, the prosodic content of the VN and NV phrases did influence infants' perception, as

different results from [1] were found after the manipulation of the prosodic content of the original phrases. However, if infants were motivated exclusively by the prosodic content of the phrases, a switch in attention should have been observed, with infants preferring phrases with the VN order and NV prosody, rather than VN phrases, as found by [1]. Instead, our manipulation of prosody made infants' unable to differentiate the VN and NV phrases. Thus, prosody is not the only cause of German infants' ability to discriminate between NV and VN verb-phrases. An alternative source is discussed in the next session.

4. DISCUSSIONS AND CONCLUSION

In the conclusion of their paper, [1] suggest that German infants' ability to differentiate NV from VN phrases could be an indication of infants' early sensitivity to rhythmical information at the level of the phonological phrase. Our results show that infants' are indeed sensitive to this source of information, but that prosody is not the only source of the differential attention found in their study.

Therefore, the segmental content of their sentences must have influenced infants' attention as well. Previous studies have shown that have shown that infants process inflectional endings [7]. Importantly, there were clear morphological differences between the verbs and nouns chosen by [1]. Interestingly, all verbs were in the infinitive, and all nouns had the affix "ung." Thus, in addition to biases towards the prosodic content of the basic syntactic structure of the German verb phrase, infants might also exhibit morphological biases. Specifically, infants might be paying attention to the morphological marker of verb-infinitive at the right edge of phonological phrases.

These findings have implications for hypotheses on bootstrapping mechanisms in first language acquisition. Bootstrapping proposals claim that information in the input might contain clues to fundamental syntactic distinctions, and that infants might rely on these cues during language acquisition. In the specific case of the acquisition of word order regularities, provocative proposals argue that infants might rely on prosodic information to trigger parameters that regulate language-specific word-order regularities [4, 5]. Thus, infants might know the basic syntactic structure of their language before any knowledge of the lexicon.

Our results indicate that infants are sensitive to this source of information from very early on. However, it is important to point out that this early sensitivity to "universal" correlations, as argued by [4, 5] might be acquired in the same way sensitivity to language-specific morphological correlations are: by paying attention to frequently co-occurring elements in the input. In addition, infants' sensitivity needs not to be seen as a way to set innate parameters; instead, it can be seen as a way to find syntactic units and relations in the input

In order to find these units and relations, infants might rely on several phonological cues. This supports the view that the integration of information from different sources facilitates the acquisition of syntax [8]. Additionally, the unreliability of prosodic cues, often used to discredit hypotheses of prosodic bootstrapping, is a smaller problem if the analysis of the speech signal is supplemented with a distributional analysis of morphological markers.

5. REFERENCES

[1] Hofmann, M., Höhle, B., & Weissenborn, J. (2003). The prosody-syntax interface in the acquisition of word order in German: University of Potsdam.

[2] Penner, Z., Weissenborn, J., & Wymann, K. (2001). On the prosody/lexicon interface in learning word order. A study of normally developing and language impaired children. In J. Weisseborn & B. Höehle (Eds.), *Approaches to bootstrapping* (Vol. 1). Amsterdam: John Benjamins.

[3] Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: evidence from early language comprehension*. Cambridge, Mass: MIT Press.

[4] Christophe, A., Guasti, M. T., Nespor, M., & van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-complement parameter. *Developmental Science, 6*, 213-222.

[5] Guasti, M. T., Nespor, M., Christophe, A., & van Ooyen, B. (2001). Pre-lexical Setting of the Head Complement Parameter through Prosody. In J. Weissenborn & B. Höhle (Eds.), *Approaches to Bootstrapping* (Vol. 1). Amsterdam: Benjamins.

[6] Brandt, I. (1983) Griffiths Entwicklungsskalen (GES) zur Beurteilung der Entwicklung in den ersten beiden Lebensjahren. Basel: Beltz Verlag.

[7] Pelzer, L., & Höhle, B. (2005). The Impact of Morphological Markers on Infants' and Adults' Speech Processing. Proceedings of the BUCLD 2005.

[8] Christiansen, M., & Dale, R. (2001). Integrating Distributional, Prosodic and Phonological Information in a Connectionist Model of Language Acquisition. Proceedings of the BUCLD 2001.