

Speech Rhythm Facilitates Syntactic Ambiguity Resolution: ERP Evidence

Maria Paula Roncaglia-Denissen^{1*}, Maren Schmidt-Kassow², Sonja A. Kotz¹

1 Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, **2** Institute of Medical Psychology, Goethe University Frankfurt, Frankfurt am Main, Germany

Abstract

In the current event-related potential (ERP) study, we investigated how speech rhythm impacts speech segmentation and facilitates the resolution of syntactic ambiguities in auditory sentence processing. Participants listened to syntactically ambiguous German subject- and object-first sentences that were spoken with either regular or irregular speech rhythm. Rhythmicity was established by a constant metric pattern of three unstressed syllables between two stressed ones that created rhythmic groups of constant size. Accuracy rates in a comprehension task revealed that participants understood rhythmically regular sentences better than rhythmically irregular ones. Furthermore, the mean amplitude of the P600 component was reduced in response to object-first sentences only when embedded in rhythmically regular but not rhythmically irregular context. This P600 reduction indicates facilitated processing of sentence structure possibly due to a decrease in processing costs for the less-preferred structure (object-first). Our data suggest an early and continuous use of rhythm by the syntactic parser and support language processing models assuming an interactive and incremental use of linguistic information during language processing.

Citation: Roncaglia-Denissen MP, Schmidt-Kassow M, Kotz SA (2013) Speech Rhythm Facilitates Syntactic Ambiguity Resolution: ERP Evidence. *PLoS ONE* 8(2): e56000. doi:10.1371/journal.pone.0056000

Editor: Emmanuel Andreas Stamatakis, University Of Cambridge, United Kingdom

Received: August 23, 2012; **Accepted:** January 4, 2013; **Published:** February 8, 2013

Copyright: © 2013 Roncaglia-Denissen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: M. Paula Roncaglia-Denissen was supported by the Max Planck International Research Network on Aging (MaxNetAging), Maren Schmidt-Kassow by the German Research Foundation (DFG SCHM 2693/1), and Sonja A. Kotz by the German Research Foundation (DFG KO 2268/1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: roncaglia@cbs.mpg.de

Introduction

Over the past decades, several psycholinguistic studies have addressed the importance of prosody in sentence comprehension (e.g., [1–5]). It has been shown that prosody is used in early stages of sentence parsing (e.g., [4–6]) and that it can help to resolve structural ambiguity (e.g., [1–3]). In addition, appropriate prosody can be used as a local cue to facilitate syntactic processing or make it more difficult when inconsistent with syntactic structures (e.g., [1,2,4,7–15]). Furthermore, prosody has been shown to influence several linguistic functions, such as phonology (e.g., [15]), semantics and pragmatics (e.g., [10,16–20]), and syntax (e.g., [21,22]).

Prosody can be understood as the acoustic features of spoken languages, such as duration, amplitude and fundamental frequency [23], manifested in at least two facets: intonation and rhythm. While intonation concerns the speaker-controlled pitch variation in course of an utterance, rhythm regards the temporal organization of the speech, allowing for segmentation of events in the utterance, i.e., sounds and pauses, and structuring them in a pattern of recurrence in time [24–27].

So far, studies investigating the importance of prosody to disambiguate syntactic structure have mainly addressed its intonational facet (e.g., [1–3,7,8,13]). To our knowledge no study has specifically investigated the role of *rhythm* as a sentence segmentation cue to disambiguate syntactic structure and to facilitate sentence comprehension. Regarding the role of speech rhythm in auditory speech and language comprehension, previous

studies suggest that listeners are sensitive to rhythmic regularity in speech (e.g., [28,29]) that a word's metric property influences lexical access (e.g., [30]), interacts with semantics [20,31] and with syntax (e.g., [15,21,32]).

However, speech rhythm should also be investigated as a broader phenomenon rather than just a local one during sentence processing. When speech rhythm operates, it not only organizes sounds into words, but also words into larger prosodic units [6,17] as part of a prosodic hierarchy [33,34], which may constitute units of perception [35–38]. Rhythm allows to segment relevant linguistic information, e.g., sounds, as speech flows, grouping it into meaningful linguistic units, e.g., words. These linguistic units may then be integrated with information from other linguistic domains, such as semantics and syntax, so comprehension is achieved [21,26,31,39]. Given its significant contribution to speech organization, the role of rhythm should be investigated, not only when it operates as a local cue at the lexical level, but also when it serves as a sentence segmentation device, i.e., prior to and during the processing of syntactic complexity.

To our knowledge, there have been no studies investigating the role of rhythm as a sentence segmentation device during syntactic ambiguity resolution using the ERPs. ERPs are of great advantage while investigating unfolding language processes, such as the use of speech rhythm in sentences segmentation, because they capture the exact time course, in which these processes occur [40]. In this sense, the use of ERPs may contribute to a better understanding of ongoing linguistic processing, allowing to expand theories and models of language processing [14,40].

So far, a few studies have used ERPs to investigate the role of prosodic breaks, as a local cue and influencing the syntactic parser during ambiguity processing (e.g., [14,41]). In these studies, the ERP component Closure Positive Shift (CPS) was associated with the occurrence of prosodic breaks, while an enlarged N400 was elicited by the less-preferred syntactic structure, object-first sentences. This enlarged N400 was previously associated with difficulty in lexical integration (e.g., [42–44]), such as the encounter of an intransitive verb when a transitive one would be preferred (e.g., [9,14]). In addition, an enlarged P600 elicited by object-first structures was found (e.g. [14,41]), which was linked to the re-analysis of this less-preferred syntactic structure (e.g., [45–47]).

In the current study, we investigated the role of *rhythm* as a sentence segmentation cue, grouping words together in regular rhythmic chunks so as to facilitate the processing of syntactically ambiguous sentences. In previous experimental work, it has been suggested that the parser makes use of prosodic information, in our case rhythm, to create low-level syntactic structures, grouping words in “chunks” [5,48,49]. These chunks would remain unattached until enough morphosyntactic information is provided, reducing memory load, without forcing the listener to commit to a possibly wrong syntactic analysis. Our view is consistent with the existence of a prosodic representation available already during early stages of sentence processing (e.g., [4,7,8,16,22]) that interacts with the syntactic parser prior to, during, and after syntactic ambiguity is encountered [7,8,16].

Therefore, we presented participants with German sentences containing syntactic ambiguity, spoken in either regular or irregular rhythmic patterns. Rhythmic regularity was established by using one stressed syllable followed by three unstressed ones that created clitic groups (groups of grammatical words carrying one primary stress only [34]) of constant size.

In order to focus on syntactic re-analysis and avoid lexical integration difficulty, we used only transitive verbs (i.e., verbs requiring an accusative argument). In this sense, we expected to find a P600 response, which has been interpreted to indicate syntactic re-analysis of a less-preferred structure, i.e., object-first order (e.g., [14,45,47]).

By presenting ambiguous sentences in rhythmically regular context, we provide a reliable segmentation cue, namely stress patterns, creating rhythmic chunks. These rhythmic chunks operate clustering linguistic constituents, such as morphemes and grammatical words sharing one common primary stress (i.e., a clitic group; [33]). As a result of their acoustic salience, i.e., shared primary stress, these clusters constitute perceptual units in the speech stream. Perceptual units may guide the syntactic parser [6,17,35–38] when structures of greater syntactic complexity are encountered (i.e., object-first sentences), facilitating their processing.

It could be the case that rhythm facilitates the processing of both syntactic structures, i.e., subject-first and object-first order, however, its benefits should be more valuable and, therefore, more apparent during the processing of sentences with enhanced processing costs (i.e., object-first sentences), as in such cases, any facilitation cue can be used. Such facilitation should be confirmed by a significant reduction in the P600 mean amplitude response to object-first rhythmically regular sentences compared to the same structure in a rhythmically irregular context. Furthermore, behavioral results, such as higher accuracy rates and faster response times, should also be found for the less-preferred syntactic structure, i.e., object-first sentences, in rhythmically regular context in comparison to their rhythmically irregular counterparts.

Methods

Ethics Statement

This study was approved by the ethics committee of the University of Leipzig. All individuals in this study gave their written informed consent for data collection, use, and publication.

Participants

Thirty-two participants (17 males; $M_{\text{age}} = 25.59$, $SD = 2.53$) participated in an initial rating study of the material, while twenty-four different participants (12 female; $M_{\text{age}} = 26.33$, $SD = 1.97$; all right-handed) took part in the EEG experiment. Participants from both studies were students of the University of Leipzig, native speakers of German, and were paid for their participation. None of the participants reported any neurological impairment or hearing deficit, and all had normal or corrected-to-normal vision.

Material

Originally, we created 480 sentences using 60 transitive verbs (requiring an accusative complement combined with 120 different common and proper nouns. By using transitive verbs instead of intransitive ones (i.e., verbs requiring dative complements) we focused on sentence reanalysis (P600; [14,47,50]), avoiding responses to difficulties in lexical integration (N400; [42,44]). Half of the sentences constituted experimental items, whereas the other half were filler sentences. Experimental sentences consisted of one main clause followed by a relative clause, i.e., the clause of interest, and were presented in a 2×2 design, with the factors *argument position* (subject-first vs. object-first order) and *rhythm* (irregular vs. regular rhythm). This resulted in sentence quadruplets, with each sentence corresponding to one of the four experimental conditions: subject-first rhythmically irregular, SFI; subject-first rhythmically regular, SFR; object-first rhythmically irregular, OFI; object-first rhythmically regular, OFR. Fillers and experimental sentences were between 17 and 19 syllables long ($M = 17.1$, $SD = 0.36$).

Rhythmic regularity was established by a constant metric pattern of one stressed syllable followed by three unstressed ones, while rhythmic irregularity was achieved through the use of proper nouns of different syllable numbers, and common nouns that varied in terms of lexical stress and the number of syllables (for illustration of these properties, see Figure 1). Word frequency for common nouns was counterbalanced across the rhythmically regular and irregular sentence conditions and were not significantly different, $z = 0.13$, $p > 0.1$.

The original 480 sentences were pseudo-randomized and arranged in 32 different written questionnaires to be rated by participants in terms of sentence content, according to a 7-point acceptability rating scale (1 = *unacceptable* and 7 = *highly acceptable*). Sentences with a mean rate of less than 4 points on the acceptability scale were removed from the stimulus material together with their experimental condition counterparts and matching fillers. This resulted in a total of 352 sentences (73.4% from the original sentences), i.e., 44 per condition, with corresponding fillers to be used as final stimulus material in the EEG experiment. The four experimental conditions, as well as their corresponding filler items, are presented in Figure 2.

These 352 final sentences were spoken by a German female professional speaker at a normal speech rate and digitally recorded via a computer with a 16-bit resolution and a sampling rate of 44.1 kHz. In order to prevent participants having access to any prosodic information other than speech rhythm, such as pitch contour variations, sentences were constructed with the application of a cross-splicing procedure.

Proper nouns		Common nouns	
masculine 2 syllables	'Jonas	2 syllables trochaic foot	'Rentner <i>pensioner</i>
feminine 2 syllables	'Jana	3 syllables trochaic foot	'Fußgänger <i>pedestrian</i>
masculine 3 syllables	Ri'cardo	2 syllables iambic foot	Pas'sant <i>pedestrian</i>
feminine 3 syllables	Lo'rena	3 syllables iambic foot	Pensio'när <i>pensioner</i>

The sign ' marks the stressed syllables.

Figure 1. Examples of proper and common nouns used in the stimulus material.

doi:10.1371/journal.pone.0056000.g001

Stimulus material			
Experimental conditions	subject-first	'Roland trifft die 'Diener, die An'tonio mal ge' st ört haben, im 'Park. <i>Roland meets the_{acc.pl.} helpers, who/whom Antonio once bothered_{pl.} in the park.</i> <i>Roland meets the helpers, who once bothered Antonio, in the park.</i>	rhythmically regular
	object-first	'Dora trifft die 'Diener, die Char'lotte mal ge' st ört hat, im Ge'schäft. <i>Dora meets the_{acc.pl.} helpers, who/whom Charlotte once bothered_{sing.} in the store.</i> <i>Dora meets the helpers, whom Charlotte once bothered, in the store.</i>	rhythmically regular
	subject-first	'Bernhard trifft die Ge'hilfen, die Ni'cole mal ge' st ört haben, im 'Park. <i>Bernhard meets the_{acc.pl.} assistants, who/whom Nicole once bothered_{pl.} in the park.</i> <i>Bernhard meets the assistants, who once bothered Nicole, in the park.</i>	rhythmically irregular
	object-first	'Paula trifft die Ge'hilfen, die Jo'el mal ge' st ört hat, im Ge'schäft. <i>Paula meets the_{acc.pl.} assistants, who/whom Joel once bothered_{sing.} in the store.</i> <i>Paula meets the assistants, whom Joel once bothered, in the store.</i>	rhythmically irregular
Sentence fillers	subject-first	'Arno trifft den 'Diener, der Ri'cardo mal ge'stört hat, im Ge'schäft. <i>Arno meets the_{acc.sing.} helper, who Ricardo once bothered_{sing.} in the store.</i> <i>Arno meets the helper, who once bothered Ricardo, in the store.</i>	rhythmically regular
	object-first	'Maren trifft den 'Diener, den Lo'rena mal ge'stört hat, im Ge'schäft. <i>Maren meets the_{acc.sing.} helper, whom Lorena once bothered_{sing.} in the store.</i> <i>Maren meets the helper, whom Lorena once bothered, in the store.</i>	rhythmically regular
	subject-first	'Eckhard trifft den Ge'hilfen, der 'Jana mal ge'stört hat, im Ge'schäft. <i>Eckhard meets the_{acc.sing.} assistant, who Jana once bothered_{sing.} in the store.</i> <i>Eckhard meets the assistant, who once bothered Jana, in the store.</i>	rhythmically irregular
	object-first	'Ellen trifft den Ge'hilfen, den 'Karsten mal ge'stört hat, im Ge'schäft. <i>Ellen meets the_{acc.sing.} assistant, whom Karsten once bothered_{sing.} in the store.</i> <i>Ellen meets the assistant, whom Karsten once bothered, in the store.</i>	rhythmically irregular

The sign ' marks the stressed syllables.

Disambiguating words in the experimental conditions are marked in bold.

Figure 2. Exemplary sentence for each experimental condition and filler sentences.

doi:10.1371/journal.pone.0056000.g002

Cross-splicing. The cross-splicing procedure, i.e., the procedure of replacing an existing sound with another one, was conducted separately for each sentence quadruplet (SFI, SFR, OFI and OFR). Stimuli cross-splicing was accomplished in four steps, using the software Praat (version 5.2.13).

Subject-first rhythmically irregular (SFI) sentences from each quadruplet were chosen as “standards”; i.e., their words were used as replacements for equivalent words in the remaining experimental conditions of the quadruplet. This was the case because SFI sentences present the preferred syntactic order in German, i.e., subject-first order, and their rhythm is natural (not experimentally manipulated). Because of this, we could create a more natural stimulus material which is also closer to natural speech. In a first step, the German plural relative pronoun (“die”/the) from the standard sentence (SFI) replaced its equivalents in the other conditions, i.e., SFR, OFI, OFR. Second, we utilized the segment immediately after the proper noun, containing the adverb and the participle of the main verb, from the standard sentence (SFI) to replace its equivalent in the other conditions (SFR, OFI, OFR). Third, the critical item, the auxiliary verb (“haben”/have), from the standard sentence (SFI) was used to replace its equivalent in its counterpart SFR sentence. Fourth, the same procedure as in step three was adopted, but this time, the auxiliary verb (hat/has) in the OFI sentence was used as a replacement for its equivalent in its counterpart OFR sentence. After applying the cross-splicing procedure, sentences were presented to 3 German native speakers and naïve listeners, who evaluated the naturalness of the sentences. None of the listeners reported hearing cuts, co-articulations or unnatural sounds in the sentences.

Procedures

Participants were tested individually in a sound-attenuating booth, seated in a comfortable chair and requested to move as little as possible during the experiment. Participants performed a comprehension task, evaluating if the content of an auditorily presented sentence matched the content of a subsequently presented visual sentence. Prior to the experiment, participants received a short training session with 2 blocks of 16 sentences each (2 per condition and 8 equivalent fillers).

Each trial started with a red asterisk presented on the center of a black computer screen. After 1500 ms, the red asterisk was replaced by a white one and, at the same time, a sentence was presented via loudspeakers. With the offset of the auditory sentence, participants saw a written rephrased version of the previously heard relative clause. Participants were instructed to press the response keys of a button box as quickly and accurately as possible: with the “yes”-key if the content of the auditorily and visually presented sentences matched, or the “no”-key, if this were not the case. If, after 2.5 s participants failed to press any response key, a new trial was presented. The position of the correct-response key (left or right side) was counterbalanced across participants.

Sentences were pseudo-randomized and presented in 8 blocks of about 5.5 min each. Experimental blocks contained either rhythmically regular or irregular sentences and were presented in an alternating fashion. Sentences were presented in blocks of rhythmically regular or irregular sentence context which, in case of regularity, was hypothesized to provide a reliable segmentation cue during the disambiguation of syntactic structures. All participants started with a rhythmically irregular block to prevent possible facilitation/entrainment effects that may result from exposure to rhythmic regularity. After each context block, participants were offered a break. At the end of the session, participants were briefly asked about their perception of the

stimulus material used, namely if they had perceived the use of rhythmic regularity in the spoken sentences. No participant reported having perceived rhythmic regularity in any of the presented sentences.

Electrophysiological Recordings. The EEG signal was recorded from 59 scalp sites by Ag/AgCl electrodes placed in an elastic cap (Electro Cap Inc, Eaton, OH, USA). Bipolar horizontal and vertical electro-oculograms (EOG) were recorded to allow for eye artifact correction. Electrodes were online re-referenced to the left mastoid and offline re-referenced to averaged left and right mastoids. Recording impedance was kept below 5 k. EEG and EOG signals were recorded with a sample frequency of 500 Hz, using an anti-aliasing filter of 140 Hz. Trials affected by artifacts, such as electrode drifting, amplifier blocking and muscular artifact, were excluded from analysis ($M=4.78\%$, $SD=6.23$), while trials containing eye movements were individually corrected, using an algorithm based on saccade and blink prototypes [51]. Trials were averaged separately per condition, i.e., SFI, OFI, SFR and OFR, and per participant (subject-average), and across all participants (grand average). Chosen epochs ranged from the onset of the critical item (i.e., the auxiliary verb and the disambiguating word; “haben”/have and “hat”/has) to 900 ms after its offset (i.e., at the onset of the visually presented sentence), and were calculated with a baseline of -200 to 0 ms. Further, all incorrectly answered trials were excluded from data analysis ($M=9.02\%$, $SD=10.04$). For graphical display only, data were filtered off-line using a 7 Hz low pass filter.

Statistical Analysis. For accuracy rates (correct *vs.* incorrect responses) a logistic regression analysis was conducted using *argument position* (subject-first *vs.* object-first order) and *rhythm* (regular *vs.* irregular rhythm) as predictors.

For the reaction times analysis, a repeated-measures analysis of variance (ANOVA) was conducted using the two experimental factors *argument position* and *rhythm* as within-subject factors. In addition, as rhythmically regular sentences contained, on average, significantly less syllables ($M=9.23$, $SD=0.50$) than their rhythmically irregular counterparts ($M=9.74$, $SD=0.972$), $z=5.56$, $p<0.01$, for *reaction times* analysis the *number of syllables* was used as covariate.

For the ERP data analysis, the time window ranging from 350 to 550 ms was chosen based on visual inspection and previous studies [47,50,52,53]. In these studies an earlier than the classical positivity (P600) was elicited during the processing of case ambiguous subject-object relative clauses. It has been suggested that case ambiguous sentences, i.e., subject-first *vs.* object-first order, lead to a less severe Garden Path [47] for structural reasons [54] as well as for lower processing costs [55]. Consequently, the early latency in the positive response would result from the ease of reanalyzing a case ambiguous sentence [55]. However, some of the previous research also reported a late positivity together with an early one [47,52]. The combined elicitation of two positivities may result from a more complex experimental setting, i.e. half of the sentences have to be disambiguated at the final auxiliary verb (similarly to studies encountering an early positivity) and the other half at an earlier point of the sentence (noun phrase). Thus, it has been suggested that the late positivity may account for a secondary verification of structural adequacy, and more likely occurring in experimental settings containing different types of case ambiguous sentences.

Furthermore, a repeated-measures ANOVA quantifying the mean amplitude data was conducted using the two experimental factors *argument position* (subject-first *vs.* object-first order) and *rhythm* (regular *vs.* irregular rhythm), and two topographical factors *region* (anterior *vs.* posterior region) and *hemisphere* (left *vs.* right

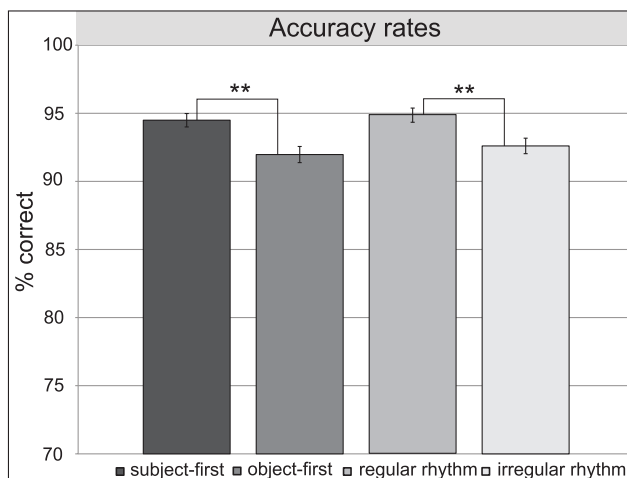
hemisphere) as within-subject factors. *Region* and *hemisphere* comprised four regions of interest (ROIs), constituted by 6 electrodes each: *left anterior* (F1, F3, F5, FC1, FC3, FC5), *right anterior* (F2, F4, F6, FC2, FC4, FC6), *left posterior* (CP1, CP3, CP5, P1, P3, P5) and *right posterior* (CP2, CP4, CP6, P2, P4, P6). To focus on main results, only significant main effects and interactions of critical factors, namely *argument position* (subject-first *vs.* object-first order) and *rhythm* (irregular *vs.* regular rhythm), are reported.

Results

Behavioral Results

Accuracy rates. Overall correct response rates were above 90% ($M_{SFI} = 93.95\%$, $SD = 23.84$, $M_{OFI} = 91.25\%$, $SD = 28.25$, $M_{SFR} = 95.12\%$, $SD = 21.54$ and $M_{OFR} = 92.69\%$, $SD = 26.02$). The full logistic model was significant, indicating that the experimental factors significantly predict participants' scores ($X^2 = 14.99$, $p < 0.001$ with $df = 2$). The Wald criterion revealed that *argument position* ($X^2 = 10.14$, $p < 0.01$) and *rhythm* ($X^2 = 4.88$, $p < 0.05$) made a significant contribution to prediction for participants' scores ($p < .001$). A follow-up analysis indicates that participants had higher scores for subject-first sentences ($M = 94.54\%$, $SD = 22.72$) than for object-first order ($M = 91.97\%$, $SD = 27.16$) and for rhythmically regular sentences ($M = 93.91\%$, $SD = 23.99$) in comparison to rhythmically irregular ones ($M = 92.61\%$, $SD = 27.16$). Table 1 presents the logistic regression analysis of participants' accuracy rates and Figure 3 the accuracy rates for *argument position* and *rhythm* in the comprehension task.

Reaction times. Overall participants' reaction times were faster than 1100 ms ($M_{SFI} = 989.82$ ms, $SD = 487.63$, $M_{OFI} = 1044.86$ ms, $SD = 514.80$, $M_{SFR} = 907.79$ ms, $SD = 451.86$ and $M_{OFR} = 926.35$ ms, $SD = 446.88$). Results revealed a significant main effect of *argument position*, $F(3,75) = 3.14$, $p < 0.05$, with faster responses for subject-first ($M = 963$ ms, $SD = 472$) than for object-first order ($M = 1004$ ms, $SD = 484$). Mean reaction times for subject-first and object-first sentences are presented in Figure 4. Contrary to what we initially expected, no significant effect of *rhythm* and no interaction between the two experimental factors were found.



Error bars indicate the standard error.

Figure 3. Accuracy rates for *argument position* and *rhythm* in the comprehension task.

doi:10.1371/journal.pone.0056000.g003

ERP Data

A repeated-measures ANOVA revealed a significant interaction between *argument position* and *rhythm*, $F(1, 23) = 6.66$, $p < 0.05$. When resolving this interaction for *argument position*, a significant main effect of *rhythm* was found for object-first sentences only, $F(1, 23) = 4.36$, $p < 0.05$, with a smaller P600 mean amplitude in rhythmically regular sentences ($M = 1.10$ μV , $SD = 2.95$) than in their rhythmically irregular counterparts ($M = 2.04$ μV , $SD = 2.06$), corroborating our initial hypothesis. For subject-first sentences, the analysis did not yield statistically significant differences between rhythmically regular and irregular sentences, $p > 0.1$; also in line with what we initially expected. No further significant interactions or main effects for the critical factors were found. Figure 5 depicts ERP responses for experimental conditions in the time window of interest.

General Discussion

In the current work, we utilized ERPs as well as behavioral measures to investigate the impact of speech rhythm as a segmentation cue during the processing of sentential syntactic ambiguity. We presented participants with syntactically ambiguous sentences embedded in regular and irregular rhythmic contexts. By providing participants with a rhythmically regular context, we expected to see a reduction of processing costs for the less-preferred syntactic structure, i.e., object-first sentences if regular rhythm works as a sentence segmentation device.

Our results partially corroborate the proposition that regular rhythm facilitates the processing of the less-preferred syntactic structure, i.e., object-first sentences. On the one hand, behavioral results, confirm rhythmic facilitation of overall accuracy rates, but independent of sentence structure type. On the other hand, in line with our hypothesis, ERP data confirm a significant rhythmic facilitation effect for the less-preferred syntactic order only (i.e., object-first sentences). This rhythmic facilitation effect is revealed by a significantly reduced P600 mean amplitude response to object-first sentences in rhythmically regular context only.

One possible explanation why behavioral results not to depict an interaction between rhythm and sentence structure type may be due to the fact that behavioral measures may only capture the outcome of the syntactic disambiguation, at the end of sentence processing. If an interaction of *rhythm* and *argument position* occurs as the sentence unfolds, then behavioral measures may not be sensitive enough to reveal such an interaction. In order to depict the complexity of an ongoing process (i.e., the use of rhythm as a sentence segmentation cue), online measures, such as ERPs, may be better suited for detecting the more immediate effects of rhythm. An alternative explanation for the differences between the behavioral and the ERP results could be based on participants' qualitatively different online and task specific responses. While behavioral measure may reflect the decision of whether the auditory and the visual rephrased sentence are the same, ERPs may reflect the response to the encountered ambiguity. Thus different task and non-task related aspects may be reflected in the two measures.

Yet, one may also argue that the use of a constant metric pattern does not occur naturally in spontaneous speech, and therefore our result reflects an artificial consequence of our manipulation. However, this reasoning seems unlikely, because a post-experimental debriefing revealed that participants did not perceive rhythmic regularity in any of the sentences they listened to. This suggests that even though rhythmicity was manipulated, this was done in a natural not obvious (i.e. as spoken by a metronome) fashion.

Table 1. Logistic regression analysis of participants' accuracy rates.

Predictor	B	SE β	Wald's χ^2	df	P	e ^β (odds ratio)
Constant	-2.6869	0.1093	604.6390	1	<0.001	NA
Argument position (subject-first = 0, object-first order = 1)	0.3952	0.1232	10.1383	1	0.0015	1.4850
Rhythm (irregular = 0, regular = 1)	-0.2722	0.1241	4.8825	1	0.0271	0.7620
Test	χ^2		df		P	
Overall model evaluation						
Likelihood ratio test	15.2171		2		0.0005	
Score test	15.1416		2		0.0005	
Wald test	14.9871		2		0.0006	
Goodness-of-fit Test						
Hosmer & Lemeshow	0.1270		2		0.9385	

Kendall's Tau-a = 0.0170; Goodman-Kruskal Gamma = 0.1750; Somers's Dxy = 0.1320; c-statistic = 56.60%. For statistical precision, all statistics here reported use 4 decimal places. NA = not applicable.

doi:10.1371/journal.pone.0056000.t001

Our findings provide new evidence of how prosodic information may affect the disambiguation of syntactic structure during sentence processing. First, while previous research has focused exclusively on the role of *intonation* [6,7,10,15,16,56] on syntactic processing, this is the first study to address the temporal nature of prosody, namely *rhythm*, during the disambiguation of syntactic structures. Second, previous research has investigated the role of intonation, i.e., prosodic breaks, as a local cue which may be used to facilitate syntactic processing [9,14,41]. Here, we addressed the role of rhythm during ongoing sentence processing, that is even before encountering syntactic ambiguity. Hence we investigated a broader scope of how rhythm operates as a segmentation cue during online sentence processing.

Our work is consistent with the idea of an existing prosodic representation available already in early stages of language processing, which interacts with the syntactic parser, guiding it through the processing of syntactic constituents [7,8,16,48,49]. Further, our work is based on the idea that prosodic units, in our

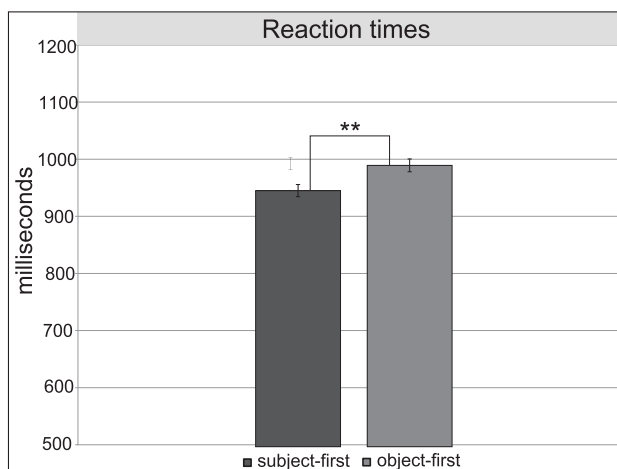
case rhythmic groups, constitute perceptual units [36,38,57], which in turn operate as processing units [6,17], reducing the memory load and facilitating language processing [16,48,49]. Thus, in the current work, we provided participants with a prosodic representations based on rhythmic regularity, which created a reliable segmentation context for the unfolding sentence, reducing the processing costs of the less-preferred syntactic structure, i.e., object-first sentences.

The importance of rhythm for speech segmentation in first language acquisition has already been shown. Studies conducted with preverbal infants reveal that infants rely on rhythmic information from their native language in order to segment speech and encode their first words [58–60]. During this process, they appear to refine their ability to discriminate rhythmic information in their native language [61,62], encoding rhythm as phonological information [63].

Once encoded, rhythm helps the listener to organize sounds and pauses in spoken language in form of a prosodic hierarchy that helps to structure an utterance at several levels and various points in time [33]. Thus, rhythm organizes sounds and pauses in the speech flow into words that can be grouped together in a clitic group (a group of grammatical words presenting one common primary stress only). Clitic groups, in turn, can be combined to create phonological phrases (i.e., clusters of clitic groups), which can be integrated into intonational phrases (a linguistic segment with one complete intonational contour, [34]).

Our results are in line with previous studies suggesting that prosodic units may act as processing units, guiding the syntactic parser through the speech stream [6,7]. Our research corroborates previous findings revealing that prosody, in our case rhythm, facilitates information processing when larger information chunks are provided [64,65]. Thus, keeping all sentential cues constant (i.e., phonological, semantic, syntactic, pragmatic and intonational) rhythm may become a salient segmentation cue, which, in turn, may increase efficiency in sentence processing. Hence, rhythm is used to guide the syntactic parser through the processing of larger information units.

One could also argue that rhythm operates as a sentence segmentation cue regardless of which syntactic structure is being processed. However, its benefit may only become apparent when



Error bars indicate the standard error.

Figure 4. Reaction times for subject-first and object-first sentences in the comprehension task.

doi:10.1371/journal.pone.0056000.g004

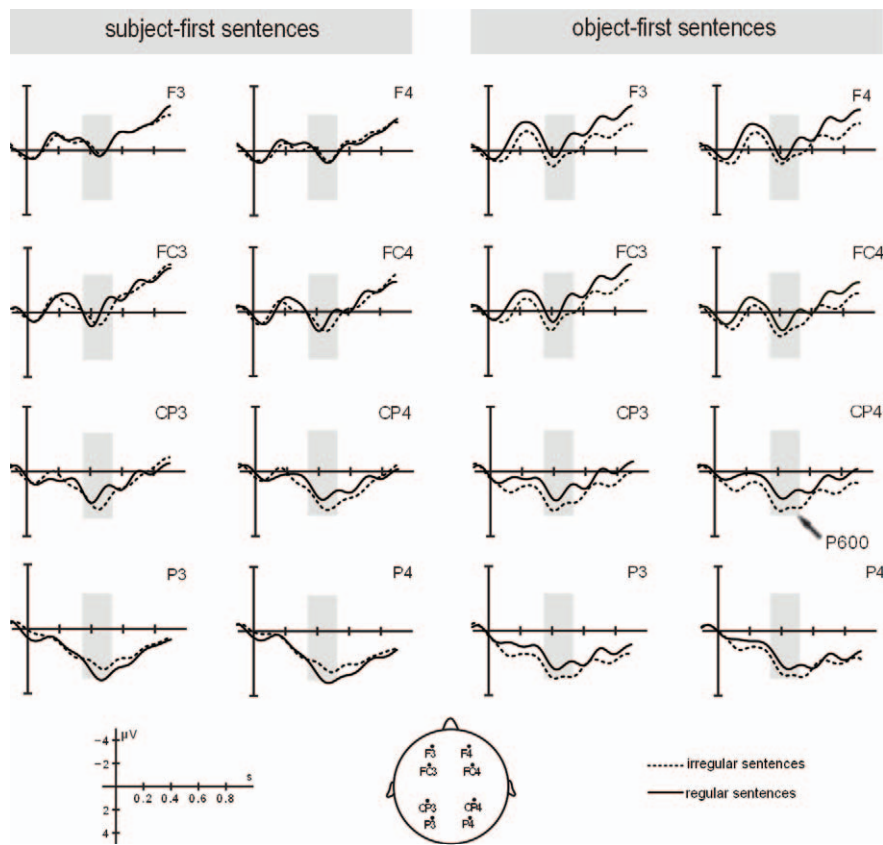


Figure 5. ERP responses for experimental conditions in the time window of interest.
doi:10.1371/journal.pone.0056000.g005

syntactic difficulty increases. Therefore, future studies should investigate the role of rhythm in a broader range of syntactic complexities during sentence processing.

In this sense, the two prosodic facets, i.e., *intonation* and *rhythm*, help to facilitate syntactic processing though in a different manner. On the one hand, *intonation* may provide complementary information to be integrated by the syntactic parser when syntactic ambiguity occurs, and thus facilitates processing [4,7,9,14,15]. On the other hand, as our study reveals, *rhythmic* regularity may already impact sentence segmentation prior to ambiguity resolution, thus facilitating information processing, and consequently reducing the overall processing costs for syntactically ambiguous sentences. Our results provide evidence of the early and continuous use of rhythm by the syntactic parser. This evidence is consistent with language processing models assuming an interactive and incremental use of linguistic information during sentence processing [6–8,16,17,48,49].

In view of these results, some questions remain. Is facilitation by means of rhythmic regularity a language-dependent or language-independent phenomenon? Some studies have shown that the perception of speech rhythm and its use as a word segmentation cue is language dependent [66–68]. Other studies investigating the cognitive ability of listeners have provided evidence that rhythm in its function of grouping elements together facilitates syllable and word recall independent of the rhythmic class of a language [69–71]. Therefore, even though rhythm as a device to segment the speech stream may be language specific, perhaps its use beyond the word level, i.e. when grouping words together, may not be.

If the use of rhythm in grouping organizing the speech stream is a universal and language-independent property, second language (L2) learners may also use rhythmic regularity in the L2 to facilitate syntactic processing. Thus, further investigations regarding the perception and the use of rhythmic regularity as a sentence segmentation cue in the context of L2 processing are called for. Such investigations should shed more light on the perception and use of rhythm in a broader sense, i.e., beyond the level of word segmentation, as a potential cross-linguistic or language-dependent phenomenon.

Conclusion

In the current work we investigated the role of rhythm as a sentence segmentation cue during the disambiguation of syntactic structures. Rhythmic regularity was achieved by the use of a constant metric pattern of three unstressed syllables between two stressed ones. Accuracy rates suggest that rhythmic regularity facilitates overall sentence comprehension. ERP results indicate a reduction of the P600 mean amplitude in response to the less-preferred syntactic structure, i.e., object-first sentences, in rhythmically regular context only. Our results suggest that rhythm may be used as a reliable sentence segmentation cue, facilitating the processing of non-preferred syntactic structures, i.e., object-first sentences, and improving sentence comprehension.

Supporting Information

Audio S1 Example of experimental sentence for rhythmically irregular subject-first order.
(WAV)

Audio S2 Example of experimental sentence for rhythmically irregular object-first order.

(WAV)

Audio S3 Example of experimental sentence for rhythmically regular subject-first order.

(WAV)

Audio S4 Example of experimental sentence for rhythmically regular object-first order.

(WAV)

Audio S5 Example of filler sentence for rhythmically irregular subject-first order.

(WAV)

Audio S6 Example of filler sentence for rhythmically irregular object-first order.

(WAV)

Audio S7 Example of filler sentence for rhythmically regular subject-first order.

(WAV)

Audio S8 Example of filler sentence for rhythmically regular object-first order.

(WAV)

Acknowledgments

The authors would like to thank Cornelia Schmidt, Ingmar Brilmayer for their help in the data acquisition, and Jaap J. A. Denissen for his helpful comments and thoughts on an early draft of this paper.

Author Contributions

Conceived and designed the experiments: MPRD MSK SAK. Performed the experiments: MPRD. Analyzed the data: MPRD. Wrote the paper: MPRD.

References

- Lehiste I (1973) Phonetic Disambiguation of Syntactic Ambiguity. Available: <http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ087996>. Accessed 31 July 2012.
- Price PJ, Ostendorf M, Shattuck-Hufnagel S, Fong C (1991) The use of prosody in syntactic disambiguation. *J Acoust Soc Am* 90: 2956–2970.
- Warren P (1985) The temporal organisation and perception of speech. [Ph.D.]. University of Cambridge. Available: <http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.355053>. Accessed 5 October 2012.
- Marslen-Wilson WD, Tyler LK, Warren P, Grenier P, Lee CS (1992) Prosodic Effects in Minimal Attachment. *The Quarterly Journal of Experimental Psychology Section A* 45: 73–87. doi:10.1080/14640749208401316.
- Watt SM, Murray WS (1996) Prosodic form and parsing commitments. *Journal of Psycholinguistic Research* 25: 291–318. doi:10.1007/BF01708575.
- Carroll PJ, Slowiaczek ML (1987) Modes and modules: Multiple pathways to the language processor. In: Garfield JL, editor. *Modularity in knowledge representation and natural-language understanding*. Cambridge, MA, , US: The MIT Press. pp. 221–247.
- Kjelgaard MM, Speer SR (1999) Prosodic Facilitation and Interference in the Resolution of Temporary Syntactic Closure Ambiguity. *Journal of Memory and Language* 40: 153–194. doi:10.1006/jmla.1998.2620.
- Speer SR, Kjelgaard MM, Dobroth KM (1996) The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research* 25: 249–271. doi:10.1007/BF01708573.
- Bögels S, Schriefers H, Vonk W, Chwilla D, Kerkhofs R (2009) The Interplay between Prosody and Syntax in Sentence Processing: The Case of Subject- and Object-control Verbs. *Journal of Cognitive Neuroscience* 22: 1036–1053. doi:10.1162/jocn.2009.21269.
- Schafer A, Carter J, Clifton C, Frazier L (1996) Focus in relative clause control. *Language and Cognitive Processes* 13: 135–166.
- Streeter LA (1978) Acoustic determinants of phrase boundary location. *Journal of the Acoustical Society of America*: 1582–1592.
- Swets B, Desmet T, Hambrick DZ, Ferreira F (2007) The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*. *Journal of Experimental Psychology: General* 136: 64–81. doi:10.1037/0096-3445.136.1.64.
- Beach CM (1991) The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language* 30: 644–663. doi:10.1016/0749-596X(91)90030-N.
- Steinhauer K, Alter K, Friederici AD (1999) Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience* 2: 191–196.
- Warren P, Grabe E, Nolan F (1995) Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes* 10: 457–486. doi:10.1080/01690969508407112.
- Schafer AJ (1997) Prosodic parsing: The role of prosody in sentence comprehension University of Massachusetts - Amherst. Available: <http://scholarworks.umass.edu/dissertations/AAI9809396>. Accessed 3 October 2012.
- Slowiaczek M (1981) Prosodic Units as Language Processing Units. University of Massachusetts.
- Gussenhoven C (1984) On the Grammar and Semantics of Sentence Accents. Walter de Gruyter. 368 p.
- Shriberg E, Stolcke A, Hakkani-Tür D, Tür G (2000) Prosody-based automatic segmentation of speech into sentences and topics. *Speech Communication* 32: 127–154. doi:10.1016/S0167-6393(00)00028-5.
- Rothermich K, Schmidt-Kassow M, Kotz SA (2012) Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia* 50: 232–244. doi:10.1016/j.neuropsychologia.2011.10.025.
- Schmidt-Kassow M, Kotz SA (2008) Event-related Brain Potentials Suggest a Late Interaction of Meter and Syntax in the P600. *Journal of Cognitive Neuroscience* 21: 1693–1708. doi:10.1162/jocn.2008.21153.
- Eckstein K, Friederici AD (2006) It's Early: Event-related Potential Evidence for Initial Interaction of Syntax and Prosody in Speech Comprehension. *Journal of Cognitive Neuroscience* 18: 1696–1711. doi:10.1162/jocn.2006.18.10.1696.
- Lehiste I (1970) Suprasegmentals. Available: <http://mitpress.mit.edu/catalog/item/default.asp?tid=5623&tttype=2>. Accessed 22 August 2012.
- Hayes B (1995) *Metrical stress theory: principles and case studies*. University of Chicago Press. 476 p.
- Nootboom S (1997) The prosody of speech: melody and rhythm. In: ; in Hardcastle L, editor. *The handbook of phonetic sciences*. Oxford: Blackwell. pp. 640–673.
- Magne C, Aramaki M, Astesano C, Gordon RL, Ystad S (2004) Comparison of rhythmic processing in language and music: An interdisciplinary approach. - Google Scholar. *Journal of Music and Meaning*. Available: http://scholar.google.de/scholar?hl=de&q=MAGNE%2C+C.%2C+M.+Aramaki%2C+C.+Astesano%2C+R.L.+Gordon%2C+S.+Ystad.+%282004%29.+Comparison+of+rhythmic+processing+in+language+and+music%3A+An+interdisciplinary+approach.+&btnG=Suche&lr=&as_ylo=&as_vis=0. Accessed 30 January 2012.
- Patel AD (2008) *Music, Language and the Brain*. New York: Oxford University Press. p.
- Dilley LC, McAuley JD (2008) Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language* 59: 294–311. doi:10.1016/j.jml.2008.06.006.
- Niebuhr O (2009) F₀-Based Rhythm Effects on the Perception of Local Syllable Prominence. *Phonetica* 66: 95–112. doi:10.1159/000208933.
- Robinson GM (1977) Rhythmic organization in speech processing. *Journal of Experimental Psychology: Human Perception and Performance* 3: 83–91. doi:10.1037/0096-1523.3.1.83.
- Magne C, Astésano C, Aramaki M, Ystad S, Kronland-Martinet R, et al. (2007) Influence of Syllabic Lengthening on Semantic Processing in Spoken French: Behavioral and Electrophysiological Evidence. *Cereb Cortex* 17: 2659–2668. doi:10.1093/cercor/bhl174.
- Dooling DJ (1974) Rhythm and syntax in sentence perception. *Journal of Verbal Learning and Verbal Behavior* 13: 255–264. doi:10.1016/S0022-5371(74)80062-9.
- Nespor M, Vogel I (1986) *Prosodic Phonology*. Dordrecht: Foris Publications. 364 p.
- Hayes B (1989) The prosodic hierarchy in meter. In: Kiparsky P, Youmans G, editors. *Rhythm and Meter*. Orlando, FL: Academic Press.
- Johnson RE (1970) Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior* 9: 12–20. doi:10.1016/S0022-5371(70)80003-2.
- Tyler LK, Warren P (1987) Local and global structure in spoken language comprehension. *Journal of Memory and Language* 26: 638–657. doi:10.1016/0749-596X(87)90107-0.
- Jusczyk PW, Hirsh-Pasek K, Kemler Nelson DG, Kennedy LJ, Woodward A, et al. (1992) Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology* 24: 252–293. doi:10.1016/0010-0285(92)90009-Q.
- Morgan JL (1996) Prosody and the Roots of Parsing. *Language and Cognitive Processes* 11: 69–106. doi:10.1080/016909696387222.
- Frazier L, Carlson K, Clifton Jr C (2006) Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences* 10: 244–249. doi:10.1016/j.tics.2006.04.002.
- Handy TC (2004) *Event-Related Potentials: A Methods Handbook*. MIT Press. 430 p.

41. Kerkhofs R, Vonk W, Schriefers H, Chwilla DJ (2007) Discourse, Syntax, and Prosody: The Brain Reveals an Immediate Interaction. *Journal of Cognitive Neuroscience* 19: 1421–1434. doi:10.1162/jocn.2007.19.9.1421.
42. Bornkessel I, McElree B, Schlesewsky M, Friederici AD (2004) Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language* 51: 495–522. doi:10.1016/j.jml.2004.06.011.
43. Bornkessel I, Schlesewsky M (2006) The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113: 787–821. doi:10.1037/0033-295X.113.4.787.
44. Haupt FS, Schlesewsky M, Roehm D, Friederici AD, Bornkessel-Schlesewsky I (2008) The status of subject–object reanalyses in the language comprehension architecture. *Journal of Memory and Language* 59: 54–96. doi:10.1016/j.jml.2008.02.003.
45. Vos SH, Gunter TC, Schriefers H, Friederici AD (2001) Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes* 16: 65–103. doi:10.1080/01690960042000085.
46. Frisch S, Schlesewsky M, Saddy D, Alpermann A (2002) The P600 as an indicator of syntactic ambiguity. *Cognition* 85: B83–B92. doi:10.1016/S0010-0277(02)00126-9.
47. Friederici AD, Mecklinger A, Spencer KM, Steinhauer K, Donchin E (2001) Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research* 11: 305–323. doi:10.1016/S0926-6410(00)00065-3.
48. Kennedy A, Murray WS, Jennings F, Reid C (1989) Parsing complements: Comments on the generality of the principle of minimal attachment. *Language and Cognitive Processes* 4: S151–S176. doi:10.1080/01690968908406363.
49. Murray WS, Watt S, Kennedy A (1998) Parsing ambiguities: Modality, processing options and the garden path. [unpublished manuscript].
50. Mecklinger A, Schriefers H, Steinhauer K, Friederici AD (1995) Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition* 23: 477–494. doi:10.3758/BF03197249.
51. Croft RJ, Barry RJ (2000) Removal of ocular artifact from the EEG: a review. *Neurophysiologie Clinique/Clinical Neurophysiology* 30: 5–19. doi:10.1016/S0987-7053(00)00055-1.
52. Friederici AD, Steinhauer K, Mecklinger A, Meyer M (1998) Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology* 47: 193–221. doi:10.1016/S0301-0511(97)00033-1.
53. Steinhauer K, Mecklinger A, Friederici AD, Meyer M (1997) Wahrscheinlichkeit und strategie: Eine EKP-Studie zur verarbeitung syntaktischer anomalien. [Probability and strategy: An ERP study on the processing of syntactic anomalies.]. *Zeitschrift für Experimentelle Psychologie* 44: 305–331.
54. Gorrell P (2000) The subject-before-object preference in German clauses. In: Hemforth B, Koniczny L, editors. *Cognitive Parsing in German*. Dordrecht: Kluwer. pp. 25–63.
55. Friederici AD, Mecklinger A (1996) Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research* 25: 157–176. doi:10.1007/BF01708424.
56. Stirling L (1996) Does Prosody Support or Direct Sentence Processing? *Language and Cognitive Processes* 11: 193–212. doi:10.1080/016909696387268.
57. Martin JG (1967) Hesitations in the speaker's production and listener's reproduction of utterances. *Journal of Verbal Learning and Verbal Behavior* 6: 903–909. doi:10.1016/S0022-5371(67)80157-9.
58. Höhle B, Bijeljac-Babic R, Herold B, Weissenborn J, Nazzi T (2009) Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development* 32: 262–274. doi:10.1016/j.infbeh.2009.03.004.
59. Nazzi T, Ramus F (2003) Perception and acquisition of linguistic rhythm by infants. *Speech Communication* 41: 233–243. doi:10.1016/S0167-6393(02)00106-1.
60. Ramus F (2002) Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition* 2: 85–115. doi:10.1075/arla.2.05ram.
61. Jusczyk PW (1999) How infants begin to extract words from speech. *Trends in Cognitive Sciences* 3: 323–328. doi:10.1016/S1364-6613(99)01363-7.
62. Jusczyk PW (2002) How Infants Adapt Speech-Processing Capacities to Native-Language Structure. *Current Directions in Psychological Science* 11: 15–18. doi:10.1111/1467-8721.00159.
63. Gerken LA (1996) Phonological and distributional information in syntax acquisition. In: Morgan JL, Demuth K, editors. *Signal to Syntax: Bootstrapping From Speech To Grammar in Early Acquisition*. Routledge.
64. Bor D, Duncan J, Wiseman RJ, Owen AM (2003) Encoding Strategies Dissociate Prefrontal Activity from Working Memory Demand. *Neuron* 37: 361–367. doi:10.1016/S0896-6273(02)01171-6.
65. Carpenter PA, Just MA (1989) The role of working memory in language comprehension. In: Klahr D, Kotovsky K, editors. *Complex Information Processing: The Impact of Herbert A. Simon*. Routledge.
66. Cutler A (1994) The Perception of rhythm in language. *Cognition*: 79–81.
67. Cutler A (1994) Segmentation problems, rhythmic solutions. *Lingua* 92: 81–104. doi:10.1016/0024-3841(94)90338-7.
68. Mattys SL, Jusczyk PW, Luce PA, Morgan JL (1999) Phonotactic and Prosodic Effects on Word Segmentation in Infants. *Cognitive Psychology* 38: 465–494. doi:10.1006/cogp.1999.0721.
69. Boucher VJ (2006) On the Function of Stress Rhythms in Speech: Evidence of a Link with Grouping Effects on Serial Memory. *Language and Speech* 49: 495–519. doi:10.1177/00238309060490040301.
70. Henson RN., Burgess N, Frith C. (2000) Recoding, storage, rehearsal and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia* 38: 426–440. doi:10.1016/S0028-3932(99)00098-6.
71. Hitch GJ (1996) Temporal Grouping Effects in Immediate Recall: A Working Memory Analysis. *The Quarterly Journal of Experimental Psychology Section A* 49: 116–139. doi:10.1080/713755609.