Research Report

Entrainment of syntactic processing? ERP-responses to predictable time intervals during syntactic reanalysis

Maren Schmidt-Kassow*, Sonja A. Kotz

Research Group Neurocognition of Rhythm in Communication, Max Planck Institute for Human Cognitive and Brain Sciences, Stephantrasse 1a, 04103 Leipzig, Germany

ARTICLE INFO
Article history:
Accepted 10 June 2008
Available online 16 June 2008

Keywords:
Auditory language processing
Cognitive oscillations
P600
Synchronization
SOA

ABSTRACT
Synchronization of two independent systems is a widely discussed phenomenon observed in many disciplines. Recent studies have shown its relevance to cognitive functions. However, its influence on language perception has not yet been investigated. As successful syntactic processing relies on rules that enable the listener to predict the category of the next incoming element, such prediction can be maximized if the auditory speech input is temporally regular and hence motivates synchronization. For this reason, the present ERP-experiments investigated the influence of successful synchronization in auditory syntactic processing. Our results clearly demonstrate that late syntactic processes (P600) are controlled by a temporally regular input. In particular, the latency of the P600 varies as a function of the duration of a predetermined interval between successive elements. The current data therefore attest the impact of synchronization on higher level cognitive processes such as syntax in language.

© 2008 Elsevier B.V. All rights reserved.

Keywords:
Auditory language processing
Cognitive oscillations
P600
Synchronization
SOA

1. Introduction

Imagine the following situations: a couple sitting in a bar being lost in an intense conversation. Synchronously, they cross their legs or pick up their drinks. Think of two girls walking along a street giggling. As soon as one of them raises her walking tempo the other simultaneously adjusts. Remember listening to your favourite dance school music — you will not be able to resist tapping your foot, or nodding your head along with the underlying beat. Finally, imagine talking to your boss whose speech lets you feel like being in a thunderstorm; intuitively, you will try to accelerate your speech rate as well. These are a few examples of a common phenomenon to be subsumed under the term ‘entrainment’.¹ This physical phenomenon describes a process in which two (or more) autonomous oscillators² synchronize and interact with each other (e.g., Clayton et al., 2005). In this context autonomy means that each of these oscillators keeps oscillating even if they do not interact. However, how does entrainment work?

So-called dynamic attending models (Large and Kolen, 1994; Large and Jones, 1999; McAuley and Jones, 2003) describe entrainment as a highly adaptive process. For instance, whenever successive time frames differ from an oscillator’s current state, the oscillator changes its period, i.e., the time frame between successive attentional peaks. This implies that entrainment is dependent on stimulus-directed attention. It is characterized by expecting “when” a next element will occur. This in turn is critical for an oscillator to adapt. Imagine a

¹ The entrainment principle traces back to the 17th century and has been first described in 1665 by the Dutch physicist Christiaan Huygens. While working on the design of the pendulum clock, Huygens realized that putting two clocks next to each other caused them to swing at the same rate. This is due to their mutual influence.

² Oscillators are generators of periodic vibrations.
horse that gallops faster and faster. Although hoof beats are not isochronous, the listener is capable to predict when a next hoof beat will occur. Analogous to the Dynamic Attending Theory proposed by Large and Jones (1999) the system produces self-sustained oscillations and temporal structured patterns under rhythmic stimulation. Such oscillation is synchronizing with an external rhythmic signal (see Fig. 1) resulting in entrainment.

Successful synchronization modulates the processing of auditory sequences as it enables the listener to predict the onset of the next incoming speech or tonal element, thus, allows the structuring of a complex sequence. This observation gives rise to an empirical investigation of the entrainment phenomenon beyond real life experiences reported above. For example, Cummins and Port (1998) showed that entrainment occurs in speech production. By means of the so-called ‘Speech Cycling Task’ the authors concluded that entrainment is relevant for speech production although speech, at first glance, seems to be completely random. In the ‘Speech Cycling Task’ participants were asked to repeat a short phrase such as “Give the dog a bone” in time with metronome beats. Results indicate that linguistic feet, i.e., a stressed syllable followed by a corresponding unstressed syllable, were entrained with the phrase repetition cycle of the metronome in a 2:1 ratio. This stable ratio has also been reported in motor gestures (e.g., Treffner and Turvey, 1993) and thus speaks to entrainment across production domains. The question arises whether unidirectional4 entrainment has been reported in language entrainment also occurs in speech perception. For example, phrase repetition cycle of the metronome in a 2:1 ratio. This stable ratio has also been reported in motor gestures (e.g., Lakatos et al. 2005, 2008).

However, symmetrical entrainment has also been reported in language and speech ontogenesis children respond very sensitively to strongly metrical sequences such as nursery rhymes, counting rhymes or jump-rope-rhymes to which they clap their hands or bounce. In addition the ‘prosodic bootstrapping’ theory offers an interesting link between the building of a grammatical knowledge and the metric pattern of a given language (Morgan, 1986; Jusczyk, 1997; Meherl et al., 1996). According to this proposal (originally formulated by Gleitman and Wanner, 1982) prosodic cues may provide infants with cues to detect syntactic boundaries even before lexical knowledge is available. Thus, external acoustic cues seem to pave the way to syntactic and semantic knowledge. Hence, external oscillators (i.e., the metric pattern of a native language) seem to play a critical role during language development. However, when examining an established language system (i.e., in adults)—how can one relate entrainment to syntactic processing?

Looking more closely, synchronization and syntax are both based on predictability. Successful syntactic processing relies on the build-up of certain structural expectancies, and can be labeled a ‘rule-based’ system (e.g. Jackendoff, 2002). Thus, syntactic sequencing of any utterance enables the listener to predict the order of future events (‘What next?’ Large and Kolen, 1994). Consider the following English example: As soon as one hears an article (Art), the listener predicts that the next incoming element is a nominal phrase (NP). Thus, the rule is modelled as “Art -> NP” that guides a listener’s expectancy. On the other hand, a temporal predictable external stimulus (i.e., an external oscillator) allows predicting when the next element occurs, provided that synchronization takes place. Consequently, synchronization may affect the temporal analysis of a linguistic input and in turn sequencing as well as the build-up of syntactic hierarchies (Large and Kolen, 1994).

Recent ERP-studies support the outlined importance of external oscillators on syntactic processing. Kotz and colleagues (Kotz et al., 2003; Frisch et al., 2003) reported that basal ganglia (BG) patients failed to show a P600 in response to

---

3 It is important to note that Fig. 1 is merely a simplistic illustration of the entrainment phenomenon. In fact, entrainment is highly complex and hierarchically organized as recent oscillatory neural studies evidenced that entrainment might occur in multiple bidirectional (bottom-up and top-down) influences in the brain (e.g., Lakatos et al. 2005, 2008).

4 Perceptual entrainment comes in two forms. Several authors observed ‘symmetrical’ (or ‘bidirectional’) entrainment, i.e., two systems mutually influence each other when subjects are asked to synchronize their leg or index finger movements with that of another person (Schmidt, Carello and Turvey, 1990; Kelso, 1995). However, symmetrical entrainment has also been reported in contexts where participants are not asked to synchronize their leg or index finger movements with that of another person (Shockley et al., 2003; Richardson et al., 2007). Additionally, there is evidence for ‘uni-directional’ entrainment, i.e., a robust external oscillator influences a system, but not vice versa. This has been shown in studies investigating brainwave synchronization following rhythmic stimulation, i.e., entrainment to periodic acoustic stimuli (Will and Berg, 2007).

5 ERP = Event related potentials.

6 The “P600” is as positive deflecting ERP-component that is inter alia connected with the re-analysis of syntactic violations in a given sentence. It is evoked about 600 ms after the onset of the syntactically violated word. However, it is important to note that the P600 seems to be far from being syntact- or language-specific as this component has also been elicited by a series of other domains, e.g., semantic violations, misspellings, mathematical operations, rhythmization of simple tone sequences, etc. (for a review, see Kuperberg, 2007; van Herten et al., 2005). However, in the current study we focus on the syntactically evoked P600.
syntactic violations. However, Kotz et al. (2005) demonstrated that BG patients profit from the presentation of an external metrically regular stimulus (i.e., an abstract sound sample of a march) before presentation of random speech samples. When presented with such “primes” BG-patients are able to re-elicit the P600. We deduct from this result, that entrainment may play a critical role in re-eliciting a P600 in BG-patients. This should result from re-synchronized firing neurons via an external oscillator, i.e., a march rhythm.

To summarize, previous findings suggest that humans automatically synchronize internal and external perceptual oscillations. Thus, uni-directional entrainment can influence cognitive processes in general as it

1. can facilitate a listener’s perception in terms of the timing of an upcoming event. This is based on successful synchronization with an external oscillator. In the case of syntactic processing this should maximize the build-up of structural expectancies7.
2. forces to adjust or to assimilate in those cases in which the build-up of precise temporal expectancies is not possible (Barnes and Jones, 2000).

The present study investigated whether speech perception in adult participants can be influenced by external oscillators. Here, we concentrate on possible entrainment mechanisms in auditory syntactic processing as (i) syntax is a rule-based system that relies on predicting which element occurs next (independent of world-knowledge), and as such may be sensitive to synchronize with external oscillators resulting in enhanced predictability, and (ii) the influence of external oscillators on language perception is maximally verifiable by the use of our paradigm (i.e., presenting phrase structure violations). Investigating phrase structure violations has the advantage that entrainment can be proved exactly because one element is missing in the critical phrase (see Figs. 5 and 6). That is, if entrainment does not take place reanalysis of the structure should be initiated immediately. However, if participants are entrained they should await the externally given period before they reanalyze the syntactic structure.

The question we ask is: does a temporally predictable input guide auditory syntactic processing as a critical uni-directional entraining oscillator?

For this purpose, words or linguistic phrases of a sentence were presented acoustically in a predictable manner, i.e., pauses were inserted into a continuous speech signal. We expected two ERP components in response to phrase structure violations, namely an ELAN8 (as an early correlate of phrase structure violations) and a P600 (as a correlate of syntactic reanalysis). Both ERP components have been established in a number of experiments (e.g., Hahne and Friederici, 1999; Rossi et al., 2006). However, as sketched out above, entrainment is dependent on attentional resources and consequently we focus on the P600, a late posterior ERP component correlated inter alia (see also Footnote 5) with late syntactic integration and/or reanalysis (Friederici, 1999; Friederici and Kotz, 2003; Kaan et al., 2000). This component has also been labeled a “controlled” component due to the fact that the P600 is attention dependent (Coulson et al., 1998). As such the amplitude size of the P600 varies as a function of probability or as a function of task. Thus, the P600 effect is more pronounced if attention is directed to the violation and it is less pronounced dependent on the frequency of a given violation. In a recent report Lakatos et al. (2008) argue that oscillations in the sensory cortices will entrain to events in the attended stream when attention is allocated to one of several rhythmic event streams. Furthermore, if relevant stimuli appear in a rhythmic and predictable pattern, neuronal oscillations should entrain to the structure of the attended stimulus stream and thus serve as instruments of sensory selection. Hence, we predicted that the P600, which is subject to attentional control, is sensitive to external predictable stimulation and in turn may be subject to entrainment as entrainment is a mechanism of attentional selection (Lakatos et al., 2008). This is also supported by the aforementioned data from patients with BG lesions showing that timing competence seems to be indispensable when handling successful syntactic reanalysis.

The present set of studies consists of two experiments: in the first experiment we manipulated the interval between the onsets of successive incoming elements (stimulus onset asynchrony, SOA) whereas in the second experiment, the interval between the off- and the onset of successive incoming elements (inter-stimulus-interval, ISI) was studied. Both experiments were not intended to contribute to the ongoing discussion how the brain’s timing circuitry may work, that is, beat- or interval-based (Keele et al., 1989; Vos et al., 1997; Pashler, 2001). In fact, both of these timing mechanisms (beat- and interval-based) would predict an advantage to synchronize with a constant SOA. Rather, we conducted Experiment II (constant ISI) in order to ensure that the expected effects in Experiment I (constant SOA) were not simply evoked by pre-segmentation of speech, i.e., separating the speech stream in smaller units by inserting pauses, but by synchronization. Thus, we created two different types of predictability in the form of a constant SOA (Exp. I) and a constant ISI (Exp. II). These manipulations should allow to critically test temporal predictability, that is, do stable SOAs and ISIs have similar or dissimilar predictive power? If entrainment occurs during syntactic processing with a constant SOA, participants should entrain regardless of whether the brain works beat- or interval-based. In the case of a constant ISI, however, participants will have to store the duration of a given ISI to use it. Consequently, working memory demands should be enhanced and synchronization should not take place. On the other hand, if pre-segmentation of speech per se causes the expected effects elicited in Experiment I, then those effects are certainly not related to entrainment. If so, we assume that similar effects should occur in Experiment II and Experiment I.

7 Large and Kolen (1994) stated that meter as a succession of strong and weak beats in an auditory sequence is defined as a temporal processing system that allows predicting when future events are likely to occur. Thus, successful processing of beats results in the building of temporal expectancies, i.e., the listener expects the next incoming element to occur at a certain point in time. As such entrainment, attention, and expectancy share one critical aspect, that is facilitation of the processing of temporally repetitive events.

8 The ELAN is an early negativity with a maximum over left anterior electrode sites.
In both experiments three timing conditions were tested: in the random condition sentences were presented as normal speech (without any manipulation), in the so-called isochronous condition intervals of silence were inserted between successive words, and in the chunked condition each chunk\(^9\) (linguistic phrases, e.g. nominal phrase (NP), verbal phrase (VP), prepositional phrase (PP)) was followed by silence (see Figs. 5 and 6).

Our hypotheses were as follows: (i) concerning the ELAN, we predicted that this component is not influenced by entrainment as it is known to be automatic and independent of cognitive control and attention. Therefore, this component should not be elicited in the chunked and the isochronous conditions as previous evidence has shown that the ELAN is sensitive to long inter-stimulus-intervals (Kotz, von Cramon, and Friederici, 2005; Neville et al., 1991). (ii) If temporally predictable onsets of words/chunks (constant SOA) in a sentence serve as external oscillators that in turn entrain internal oscillators coupled with syntax, such entrainment should affect syntactic processing as the acoustic input is maximally predictable. Thus, the P600 amplitude size, linked to syntactic reanalysis/repair (Friederici and Kotz, 2003) or syntactic integration (Kaan et al., 2000), should differ in the isochronous and chunked conditions compared to the random condition, as the predictable onset of successive phrases should facilitate sentence processing. In that context “differ” means that the P600 effect could either increase or decrease. On the one hand, an amplitude decrease may indicate facilitated reanalysis. On the other hand, an amplitude increase of the P600 is also possible as the syntactic violation is perceived as more salient, as it is attended. (iii) However, the main focus of this study was the P600 in the chunked condition ("Der Kuchen – wurde – im Ofen – gebacken" versus "Der Kuchen – wurde – im – gebacken"). In ungrammatical sentences (⁎) the critical chunk (prepositional phrase) consists of a preposition followed by a pause (– im –). In grammatical sentences the critical chunk is complete (a preposition followed by an NP ‘– im Ofen –’). Thus, after a few sentences participants should be able to detect the syntactic violation as soon as they hear a preposition followed by silence (recognition point 1, Fig. 5). However, if internal oscillators are entrained by temporally regular presentation of stimuli, syntactic information should not be processed until the given period (interval between two following chunks) has ended (recognition point 2, Fig. 5). Thus, we predict a P600 at the position of the participle. If entrainment does not take place, participants will already have...

\(^9\) Chunking is what most people call grouping. This refers to the segmentation of a sequence into smaller sub-sequences (Palmer, 1997), i.e. into meaningful structural units (Large, 2001; Lehrlah and Jackendoff, 1983). This segmentation is done to facilitate the processing of events. For example, telephone numbers are much easier to handle if they were split into groups of two or three digits. In speech, such chunks should be linguistic phrases as they build up a hierarchical unit, that means a VP, NP or PP serves as chunk. In the current study pauses were inserted after each linguistic phrase in order to build chunks.
initiated syntactic re-analysis at the beginning of the pause following the preposition. This should result in a P600 in response to the preposition followed by a pause. Therefore, the chunked condition serves as a critical test whether external oscillators such as predictable stimulus onsets can entrain syntactic processing.

2. Results experiment I

2.1. Performance

Participants performed a correctness judgment task. To analyze error rates we conducted an omnibus ANOVA that included the within-subject-factors condition (correct, incorrect) and presentation (isochronous, chunk, random). There was a main effect of condition \( F(1,35)=31.76, p<.01 \) with lower error rates in the incorrect (1%) than the correct condition (3.68%).

2.2. ERPs

As hypothesized syntactic phrase structure violations evoked a biphasic pattern consisting of an early negativity (ELAN) and a late positivity (P600) in the random speech condition (see Fig. 2). The averaged time window started 100 ms prior to the onset of the critical item which was the ‘ge’- participle. On the basis of visual inspection, two time windows were statistically analyzed including the window 120–300 ms (ELAN) and 600–1000 ms (P600) after the onset of the critical item. In the following analyses different subsets of electrodes were analyzed as regions of interest (ROI) and labeled as follows: left anterior (FP1, AF7, AF3, F9, F7, F5, F3, FT9, FT7, FC5, FC3), right anterior (FP2, AF8, AF4, F10, F8, F6, F4, FT10, FT8, FC6, FC4), right posterior (CP4, CP6, TP8, TP10, P4, P6, P8, P10, PO4, PO8, O2), and left posterior (TP9, TP7, CP5, CP3, P9, P7, P5, P3, PO7, PO3, O1).

To evaluate effect sizes we computed omega-square (\( \Omega^2 \)), i.e., the coefficient of determination, which represents the proportion of variance in the dependent variable accounted for by the independent variable (interpreted in a similar manner as \( r^2 \)). As we have used a within-subject-design \( \Omega^2 \) - values greater than 0.26 are defined as large effects, \( \Omega^2 \) - values from 0.048 to 0.26 are defined as medium effects and \( \Omega^2 \) values from 0.019 to 0.048 are small effects (cf. Cohen, 1992). The Greenhouse–Geisser-Correction (Greenhouse and Geisser, 1959) was applied when evaluating effects with more than one degree of freedom. Only those significant results that correlate with a critical factor are reported in the following.

2.2.1. ELAN

To evaluate the early negativity an ANOVA with four within-subject factors was conducted, namely condition (correct/incorrect), presentation (random/isochronous/chunk), hemisphere (left/right) and region (anterior/posterior). The omnibus ANOVA revealed a main effect of presentation \( F(2,70)=92.31, p<.001 \) and an interaction between the factors presentation and condition \( F(2,70)=14.90, p<.001 \), as well as a three-way interaction between the factors condition, presentation, and hemisphere \( F(2,70)=4.10, p<.05 \). Resolving both interactions by factor presentation showed a significant condition effect for random \( F(1,35)=18.74, p<.01; \Omega^2=0.33 \) and isochronous presentation rate \( F(1,35)=5.42, p<.05; \Omega^2=0.11 \) as well as a marginal condition effect for the chunked presentation \( F(1,35)=3.58, p<.08; \Omega^2=0.07 \). However, note that isochronous and chunked condition resulted in positive deflection effects, whereas only the random violation condition evokes an early negative effect. Moreover, we observed a marginal interaction of condition by hemisphere for the random presentation \( F(1,35)=3.98, p<.06 \). Resolving the interaction by the factor hemisphere resulted in a significant condition-effect in both hemispheres (left: \( F(1,35)=21.20, p<.001, \Omega^2=0.22 \); right: \( F(1,35)=1.39, p<.01, \Omega^2=0.12 \), with the left hemispheric effect being larger than the right hemispheric effect. Hence, an ELAN-effect was only evoked in the random presentation.

2.2.2. P600

To evaluate the P600 component, an ANOVA over posterior electrodes was conducted. It included the within-subject-factors condition (correct/incorrect), presentation (random, isochronous, chunk) and hemisphere (left, right). In the 600 to 1000 ms latency window the omnibus ANOVA resulted in a main effect of condition \( F(1,35)=6.29, p<.05 \), a main effect of presentation \( F(2,70)=11.56, p<.001 \) and an interaction effect between condition and hemisphere \( F(1,35)=7.58, p<.05 \). Step-down-analyses showed a significant effect of condition only for the left hemisphere \( F(1,35)=9.86, p<.01 \). Post-hoc comparisons of the different presentation rates resulted in significant differences between the isochronous and the chunked presentation rate \( F(1,35)=11.84, p<.01 \), as well as the random and the isochronous presentation rate \( F(1,35)=13.72, p<.001 \), but not between random and chunked presentation \( p=.1 \). These effects varied as the isochronously presented sentences overall resulted in a more positive amplitude (mean: 3.72 \( \mu \)V) than the random (mean: 2.21 \( \mu \)V) and chunked (mean: 2.57 \( \mu \)V) sentences. Thus, the P600 was evoked in all of the three timing conditions after the onset of the participle.

2.3. Interim summary

We investigated whether the entrainment phenomenon that has been reported in multiple domains also holds for auditory language comprehension. In particular, we focussed on syntactic processing if an SOA is kept constant. We presented syntactically correct and incorrect sentences in three different presentation modes, namely a random speech presentation (random), a word-by-word-presentation (isochronous) and a phrase-by-phrase-presentation (chunked). As hypothesized, two ERP-components were evoked in response to the syntactic violation, namely, an early anterior negativity (ELAN), and a late posterior positivity (P600). As expected, the ELAN was only elicited in the random presentation mode, and absent in the isochronous as well as in the chunked presentation mode. This is consistent with the results of previous studies that reported no ELAN when the inter-stimulus interval exceeded a critical time-window. This finding supports the assumption that a relatively quick input is necessary to successfully build initial phrase structures (Kotz et al., 2005).

The P600 was evoked in response to the participle in each of the presentation modes. This is astonishing as the syntactic violation in the chunked manipulation was already detectable at the offset of the preposition. Thus, the process underlying
the P600 was synchronized with the external oscillator. Due to the constant SOA the onset of every phrase was maximally predictable and was maximally anticipated at a certain time point.

Nonetheless, one could argue that pre-segmentation and thus slowing down of the incoming speech stream influences a participants’ processing speed. As recently shown by Mueller and colleagues (Mueller et al., 2008), pauses in an acoustic input can crucially influence cognitive operations. The authors provided evidence that in newly learned sequences a P600 effect in response to rule-violations was only elicited if the acoustic input included pauses. Thus, it remains open whether the constant SOA in fact served as an entraining oscillator (resulting in synchronization) or if induced pauses per se have an impact on the human processing system. To look more closely at this possibility, we conducted a second experiment. Here, pauses between successive words/phrases were equally long whereas the interval between onsets of words/chunks differed. Thus, in contrast to the previous SOA manipulation, we now manipulated the inter-stimulus interval (ISI). The prediction was that if entrainment takes place as a result of a constant SOA during auditory syntactic processing then a constant ISI should not lead to synchronization as the onset of successive word/chunks is not predictable in the latter condition.

3. Experiment II

Experiment II was conducted to ensure that possible entrainment effects observed in Experiment I were not just due to pauses serving as entraining oscillators. Thus, we manipulated the inter-stimulus interval (ISI), i.e., the interval between the offset and the onset of successive words or phrases by inserting constant pauses of 400 ms.

The following hypotheses were formulated:

1. Concerning the ELAN, the same results as in Experiment I were expected, i.e., the ELAN should not be elicited in the chunked and the isochronous conditions due to its temporal sensitivity.
2. It is a constant SOA that serves as external oscillator leading to entrained syntactic reanalysis in chunks. A constant ISI does not lead to entrained syntactic processing. Therefore, the P600 should only be elicited at the preposition in the chunked condition.

4. Results experiment II

4.1. Performance

To evaluate differences in error rates, we conducted a repeated-measures ANOVA with the within-subject-factors condition (correct, incorrect) and presentation (random, isochronous, chunk). This analysis revealed a significant effect of condition \( F(1,35) = 14.04, P < .01 \) with the lower error rates in the incorrect (1.57% correct) than the correct condition (3.62% correct).

4.2. ERPs

As expected, we observed a biphasic ELAN-P600 pattern at the participle in the random condition, a P600 at the participle in the isochronous condition (see Fig. 3), and a P600 at the preposition in the chunked condition (see Fig. 4). As ERPs had slightly different latencies compared to Experiment I different time windows were defined for statistical analyses. These were 120–250 ms (ELAN), 500–900 ms (P600) after the onset of the critical item (participle) and additionally for the chunked condition 800–1200 ms after the onset of the preposition. The same ROIs as in Experiment I were used.

4.2.1. ELAN

To evaluate the early negativity an ANOVA with four within-subject factors was conducted, namely condition (correct/incorrect), presentation (random/isochronous/chunk), hemisphere (left/right) and region (anterior/posterior). The omnibus ANOVA revealed a main effect for condition \( F(2,70) = 10.26, P < .001 \) and an interaction between condition and presentation \( F(2,70) = 13.48, P < .001 \). Resolving the interaction revealed a significant effect for condition in all of the three timing conditions, but only the random condition resulted in a negativity, whereas chunk and isochronous conditions resulted in a statistically significant positivity (chunk: \( F(1,35) = 24.56, P < .001; \quad \Omega^2 = 0.25 \); isochronous: \( F(1,35) = 5.13, P < .05; \quad \Omega^2 = 0.05 \); random: \( F(1,35) = 4.75, P < .05; \quad \Omega^2 = 0.05 \)).

Thus, as initially predicted only the random presentation of syntactically violated sentences resulted in an early negativity.

4.2.2. P600

Comparable to Experiment I we analyzed posterior electrodes to evaluate the P600 effect. The omnibus ANOVA included the within-subject-factors condition, presentation (random/isochronous/chunked) and hemisphere (left/right) and yielded in a main effect for condition \( F(1,35) = 6.87, P < .05 \), an interaction condition × hemisphere \( F(1,35) = 5.15, P < .05 \), and an interaction condition × presentation \( F(2,70) = 6.94, P < .01 \). Step-down-analyses revealed significant effects for condition in the isochronous \( F(1,35) = 4.40, P < .05; \quad \Omega^2 = 0.04 \) and random presentation \( F(1,35) = 13.50, P < .001; \quad \Omega^2 = 0.15 \), only. Thus, consistent with our hypotheses, no P600 was evoked in response to the participle in the chunked condition. Resolving the interaction condition × hemisphere by hemisphere revealed a significant effect for condition only in the left hemisphere \( F(1,35) = 10.30, P < .01; \quad \Omega^2 = 0.11 \).

To evaluate the P600 in the chunk condition a separate ANOVA was conducted in a time-window of 800–1200 ms (see Fig. 4) after the onset of the preposition. In this ANOVA a condition × effect turned out to be significant \( F(1,35) = 37.43, P < .001 \) indicating that syntactic reanalysis has already taken place at the offset of the preposition.

5. Discussion

To summarize the reported results, each of our experiments clearly leads to very different results in the critical chunked
condition depending on the respective timing of the chunks (SOA, ISI). While we found a P600 effect in response to the participle (e.g., "gebacken") in Experiment I (constant SOA), the P600 was elicited at the offset of the preposition (e.g., "im") in Experiment II (constant ISI). Concerning the other timing conditions (random and isochronous) Experiment I and II revealed similar results: In both experiments we found a biphasic pattern in the random condition (ELAN followed by a P600) and a P600 effect in response to the participle in the isochronous condition.

The aim of the current experimental series was twofold. Firstly, by means of an isochronous condition compared to a random speech condition in both manipulations (constant ISI/constant SOA), we wanted to investigate whether temporally predictable incoming words facilitate syntactic reanalysis. Secondly and most importantly, the chunked condition served as critical test whether entrainment takes place in language perception ("Der Kuchen – wurde – im – gebacken instead of Der Kuchen – wurde – im Ofen – gebacken"). If syntactic processing can be entrained by either a constant ISI or a constant SOA, we predicted a P600 in response to the participle although detection of the syntactic violation is already possible at the beginning of the preceding pause.

Concerning the first aim the answer is ‘No’. There were no significant differences in P600 amplitude size between the presentation modes in both experiments. If temporally predictable presentation would facilitate the re-analysis of a syntactic structure, we expected that the P600 amplitude in the isochronous condition would be significantly reduced compared to the random condition. Otherwise, if syntactic violations are perceived as more salient due to the attentional...
focus, the P600 should be enhanced in the isochronous compared to the random condition. However, neither one nor the other was the case in both experiments. This result confirms the assumption that healthy participants do not benefit from an external oscillator.

However, the constant SOA in Experiment I impacted the processing of syntactic violations in healthy participants as evidenced in the chunked condition. Statistical analysis revealed that a P600 was evoked in response to the participle in each of the presentation conditions. This is astonishing as the syntactic violation in the chunked condition was already noticeable at the offset of the preposition. Thus, the process underlying the P600 is synchronized with the external oscillator that clearly reflects an entrainment phenomenon. Due to the constant SOA, the onset of every phrase is maximally predictable and maximally anticipated at a certain point in time. Data from a production study conducted by Cummins and Port (Cummins and Port, 1998) suggested that synchronization with an external event is relatively stable. This is in line with the Dynamic Attending Theory (Large and Jones, 1999). According to this theory the phase and the period of the internal attending rhythm’s oscillations become coupled to an external rhythm (via entrainment) when the individual attends to this rhythm. This should create stable attractor states. Here, the period of the oscillation should reflect the rhythmic rate or overall tempo, while the phase relationship between the coupled external and internal attending rhythms should express the listener’s expectation about when an (external) onset or beat should happen (cf., Hawkins and Smith, 2001). Thus, the constant SOA in the stimulus material tested in Experiment I served as such an external rhythm and entrained the period and phase of internal oscillations resulting in stable attractors states at certain sentence positions. Consequently, processing speed was adapted to the expectancy when the next item was going to be presented. Participants obviously ‘awaited’ a given period (in our study: 1260 ms) before initiating syntactic re-analysis. Thus, entrainment evolved, i.e., period and phase of internal oscillations coupled with syntactic processing, and were synchronized with period and

Fig. 5 – Experiment I: Experimental conditions.
phase of externally given oscillations. However, the present results also show that only the process underlying the P600 can be entrained by external rhythmic stimulation, while the ELAN is not. This fits nicely the observation that the P600 is dependent on attention (Coulson et al., 1998) and as such sensitive to external oscillators and consequently to entrainment as a mechanism of attentional selection (Lakatos et al., 2008). On the other hand, the ELAN is thought to be an automatically evoked component (e.g., Hahne and Friederici, 2002) and independent of attentional control which is in line with our observation that the ELAN fails to be subject of entrainment.

Furthermore, based on the present data we can exclude the possibility that the effect found in Experiment I was simply due to a slow down of the incoming language input as the P600 was unaffected by external input when the interval between offset and onset of successive words/chunks (ISI) was held constant as shown in Experiment II. As a constant ISI likewise allows predicting the onset of the next incoming element, it cannot be taken for granted that predictability automatically leads to changes in syntactic processing. Experiment I therefore evidenced that language comprehension, which, without doubt, demands high timing capabilities is also sensitive to entrainment mechanisms. The current data may also provide interesting suggestions with respect to therapeutic application as has already been proposed by Kotz et al. (Kotz et al., 2005) As external oscillations can influence the time window of syntactic reanalysis in healthy populations such external oscillations may serve as compensatory mechanism in patients that suffer from a syntactic integration deficit, such as Broca aphasics (Wassenaar and Hagoort, 2005), basal ganglia patients (Frisch et al., 2003; Kotz et al., 2003) or stutterers (Alm, 2004). It may be that the application of external oscillators entrains internal oscillators coupled to syntactic processing, thus, resulting in a successful syntactic reanalysis (reflected by the P600) in those patient populations. Previous data on temporal entrainment of cognitive functions has already revealed evidence that temporal structures in music may serve as external oscillations resulting in a triggering of cortical networks involved in learning and memory in patients with multiple sclerosis (Thaut et al., 2005). Consequently the compensatory mechanism of entraining oscillators seems to be a promising approach.
One further aspect concerns possible consequences for visually conducted experiments and thus, their compatibility with acoustic studies. The present results could show that a constant SOA results in a synchronization of the syntactic processing system with an externally given predictable time frame. Thus, if experimenters keep the ISI constant they automatically provide a constant SOA as the presentation time of each single word is usually the same across stimuli. This may result in an entrainment where none is intended and consequently may cause modality-effects that have nothing to do with underlying cognitive processing differences, but are solely induced by timing differences.

However, on the basis of the reported results we can say that the period of cognitive oscillations responsible for at least attentionally controlled syntactic processing is sensitive to entrainment supporting generalization of the entrainment concept not only for motor (production), but also for cognitive processes (comprehension).

6. Experimental procedures

6.1. Experiment I

6.1.1. Participants
Thirty-six right-handed native speakers of German with unimpaired hearing (18 female) aged 20–29 years (mean = 25.1; SD = 2.7) participated. They were paid a small compensatory fee. All had normal or corrected-to-normal vision and none of them had any neurological impairment. The whole study was approved by an ethical committee.

6.1.2. Materials
As already described in the introduction, sentences with phrase structure violations were constructed and presented in a random (normal speech), isochronous, and chunked condition. Furthermore, semantically violated as well as short correct sentences served as fillers, whereas long correct sentences with a prepositional phrase (PP) served as a contrast condition for syntactic phrase structure violations (see examples above).

Examples:

a) Correct with PP: *Der Kuchen wurde im Ofen gebacken.* ‘The cake was in the oven baked.’

b) Syntactic violation: *Das Kuchen wurde im gebacken.* ‘The cake was in the baked’

c) Filler correct: *Die Pizza wurde gebacken.* ‘The pizza was baked.’

d) Filler semantic violation: *Der Tisch wurde gebacken.* ‘The table was baked.’

6.1.3. Production of the stimuli
All sentences were spoken by a female native speaker of German at a normal speech rate, recorded onto digital audiotape, and digitized at a sampling rate of 44.1 kHz. Ungrammatical sentences (‘*Der Kuchen wurde im gebacken*’) were constructed by an extensive splicing procedure to avoid co-articulatory deviations. In a next step timing manipulations were carried out.

In the random condition all sentences remained as they were spoken by the professional speaker; the chunked and isochronous sentences were manipulated as follows:

1. Isochrony: The SOA corresponds to the duration of the longest of all words (lasting 800 ms) plus 400 ms silence resulting in an interval of 1200 ms between words. If, for example, a word takes only 100 ms, the silent-interval was 1100 ms.

2. Chunking: The SOA corresponds to the duration of the longest chunk of the whole material (lasting 860 ms) plus 400 ms silence resulting in an interval of 1260 ms between successive chunks. If a chunk takes only 250 ms the silent-interval was 1010 ms (see Fig. 5). Consequently the stimulation rate was about 0.8 Hz.

All insertions were done by carefully controlling formant transitions. In order to ensure a precise time locking of the ERPs to individual sentences, the onsets of the critical words (participle as well as preposition) were marked by way of careful inspection of the auditory and visual (spectrogram) signal. All manipulations were realized using the speech editor PRAAT (PRAAT version 4.3.07, © Paul Boersma and David Weenink).

6.1.4. Procedure
The stimulus material was divided into three lists with each list being presented in a different timing condition. The order of lists was counterbalanced across participants. Thus, all participants heard sentences in a random, chunk and isochronous condition, each in a different list×timing version. To fully counterbalance lists, 36 participants were measured (3×2-design, within-subjects) in each condition. All sentences were presented auditory via loud speakers in a randomized order. The experimental trials were presented in six blocks (2 per timing condition, counterbalanced across subjects) of approximately 12 min each. Before running the two blocks in the particular timing condition participants had a training session to familiarize with the respective timing condition. Each sentence was introduced by a visual cue in the centre of a computer screen. 2000 ms after the offset of the sentence participants were asked to perform a correctness judgement. The next trial started, 2000 ms after participant’s button press.

6.1.5. Electrophysiological recording
The EEG was recorded from 59 scalp sites by means of Ag/AgCl electrodes mounted in an elastic cap (Electro-Cap Inc., n.d.) according to the 10–20 International System (Pivik et al., 1993). The Sternum served as ground, the left mastoid as on-line reference (recordings were re-referenced to averaged mastoids off-line). Electrode impedances were kept below 3 kΩ. In order to control for eye movements, a horizontal and a vertical EOG was recorded. Each EEG and EOG channel was amplified with a band pass from DC to 30 Hz with a digitization rate of 500 Hz.

6.1.6. Data analyses
Individual EEG recordings were scanned for artefacts such as electrode drifting, amplifier blocking, muscle artefacts, eye movements or blinks by means of a rejection algorithm as...
well as on basis of visual inspection 100 ms before onset of the critical item (the participle and the preposition respectively) up to 2500 ms after the critical item. All contaminated trials as well as incorrectly answered trials were rejected, thus 21% of the trials were excluded. The remaining trials (about 26 per condition and timing manipulation) were averaged per participant, condition, and electrode site, using a baseline from −100 to 100 ms. For graphical display only, data were filtered offline with a 7 Hz low pass filter. All statistical evaluations were carried out with unfiltered ERP data.

6.2. Experiment II

6.2.1. Participants
Comparable to Experiment I, we tested thirty-six right-handed native speakers of German with unimpaired hearing (18 female) aged 18–28 years (mean=25.1, SD=2.7) applying the same criteria.

6.2.2. Materials
Materials are nearly identical to the one presented in Experiment I. However, in the second experiment 400 ms of silence were inserted between consecutive words in the isochronous condition whereas in the chunked condition 400 ms of silence were inserted after each chunk. This resulted in a constant inter-stimulus-interval of 400 ms (see Fig. 6).

Acknowledgments

The authors would like to thank two anonymous reviewers, Amelie Mahlstedt and Kathrin Rothermich for helpful comments on an earlier draft of this paper, Heike Boethel for help during data acquisition, and Kerstin Flake and Andrea Gast-Sandmann for graphic support.

References


