Intentional switching of auditory attention between long and short sequential tone patterns

Sophie Nolden[1] · **Iring Koch**[1]

Abstract The current study focuses on auditory task switching, more precisely on switching attention between different temporal patterns of the same auditory stimulus. Tone sequences consisting of nine different pitch tones were presented aurally. Three repetitive short 3-tone patterns (local focus) were combined to a long pattern (global focus), and each could be either rising or falling, resulting in congruent or incongruent combinations. Participants were informed by a cue if they had to attend to the short or to the long pattern, and they indicated if the target pattern was rising or falling by pressing one of two keys. In two experiments, we investigated cued switches between the two attentional foci. Switch costs in reaction times and errors were observed when switching from the long to the short pattern but not when switching from the short to the long pattern. These asymmetric switch costs were reduced when participants had more time to prepare for the switch in a condition with a prolonged cue-stimulus interval. In addition, participants made more errors when global and local patterns did not correspond to each other (i.e., in incongruent trials) when attending to either of the patterns, but this congruency effect was not modulated by preparation time. The data suggest that the mechanisms of task goal prioritizing, as indicated by the asymmetric attention switch costs, are dissociable from those underlying stimulus selection, as indicated by the congruency effects.

Keywords Task switching · Auditory selective attention · Cognitive control · Global/local processing

✉ Sophie Nolden
nolden@psych.rwth-aachen.de

[1] Institute of Psychology, RWTH Aachen University, Jaegerstr. 17/19, D-52066 Aachen, Germany

Introduction

Switching between different tasks is an important capacity that allows us to pursue complex goals as well as to adapt our behavior when environmental demands change (Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995; see Kiesel et al., 2010; Vandierendonck, Liefooghe, & Verbruggen, 2010, for reviews). However, it is difficult to find an unambiguous definition of what a task is (Rogers & Monsell, 1995), and a task can comprise motivational factors, attention, perceptual, and motor processes, stimulus-response rules (i.e., a certain stimulus requires a left-hand response and another stimulus requires a right-hand response), or mnemonic processes, or even multiple subtasks (Hirsch, Nolden, & Koch, in press).

A typical finding in task switching in general are so-called switch costs. That is, when participants shift from one task to another, performance, as for example measured in response times or error rates, is worse than when the task remains constant. Two kinds of behavioral indices can be distinguished when switching between tasks. First, performance can be worse when the task changes from one trial to the next trial compared with when the task stays the same ("switch costs," for a review, see Kiesel et al., 2010). Second, in addition to these "transient" switch costs, performance impairments between mixed blocks, in which participants' task can change from trial to trial, and "pure" blocks, in which participants do not need to change the task, are also possible ("mixing costs"). One would typically compare performance in repetition trials of mixed blocks with trials in pure blocks (where all trials are repetition trials by definition). These costs of "sustained" control (Braver, Reynolds, & Donaldson, 2003) can be interpreted in terms of higher working memory load in mixed blocks than in pure blocks because participants have to keep two different

stimulus interpretations active and because participants cannot predict the next task (Kiesel et al., 2010; Liefooghe, Barrouillet, Vandierendonck, & Camos, 2008).

Intentional control of auditory attention in selective listening

One way to implement a sequence of tasks that vary unpredictably is to use a cue at the beginning of each trial. This cue then indicates which task participants should perform in the upcoming trial. This kind of explicit cueing-procedure has been used in several studies to investigate auditory selective attention in a dichotic listening paradigm (Koch & Lawo, 2014, 2015; Koch, Lawo, Fels, & Vorländer, 2011; Lawo & Koch, 2014, 2015; Lawo, Fels, Oberem, & Koch, 2014). In the dichotic listening paradigm, two different auditory stimuli or auditory streams are presented simultaneously, one to each ear. Typically, participants only attend to one ear. Auditory selective attention allows the listener to focus on the relevant auditory stimulus, while distracting stimuli remain in the perceptual background (for reviews, see Bronkhorst, 2015; Lachter, Forster, & Ruthruff, 2004; Schneider, Li, & Daneman, 2007; Shinn-Cunningham, 2008). Early studies on auditory selective attention in a setting with competing stimulus streams have focused on sustained auditory attention and involuntary attention capture (for early examples of dichotic listening, see Broadbent, 1958; Cherry, 1953; for a review see Hugdahl, 2011, for studies on selective attention to a certain part of a complex stimulus, see Mondor, & Bregman, 1994; Mondor, Zatorre, & Terrio, 1998).

Combining task-switching methodology with dichotic listening allowed investigating an important aspect of auditory task switching, namely intentional control of the auditory attentional focus. In one study, Koch and collaborators applied a special case of auditory task switching in which the main focus was on the attentional demands of the task. Participants had to classify one of two dichotically presented number words: one spoken by a female voice and the other one spoken by a male voice (Koch et al., 2011). A visual cue that was presented before the spoken number words informed participants which speaker they had to attend to. The relevant speaker category (female or male) could change from trial to trial (switch trial) or remain the same (repetition trial). Classification rules were held constant because participants had to classify the relevant numbers according to their magnitude in all trials, no matter if the relevant speaker changed or repeated. In addition, stimulus-response mappings were also held constant, because participants had to press the left key for small numbers and the right key for large numbers, no matter if the relevant speaker changed or repeated. Thus, this special variant of auditory task switching allowed investigating the control of the auditory attentional focus while keeping the remaining task-related factors constant. Hence, cued auditory

attention switching can be understood as a special case of task switching where the classification and the response mappings remain constant, but the relevant auditory stimulus selection criterion changes (see also Logan, 2005, for a similar approach to visual attention switching). Koch et al. (2011) found faster responses when the selection criterion repeated than when it switched and thus auditory switch costs.

To investigate if control processes in auditory task switching were proactive, preparation effects were investigated by manipulating the interval between cue and target stimuli. In addition to general preparation benefits, switch-specific preparation effects are an index of proactive control processes. The authors used a 2:1 cue-to-task mapping and focused on trials without immediate cue repetitions to control for perceptual cue priming in task repetitions. However, unlike many studies in the domain of visual task switching (Kiesel et al., 2010, for a review), auditory attention switch costs were not reduced when participants had more time to prepare for a switch (Koch et al., 2011, Experiment 3). Therefore, the factors that determine whether preparation for cued auditory attention shifts is effective still remained unclear.

Intentional control of auditory attention to temporally limited parts of auditory patterns

Notably, attending to one auditory stimulus among distracting stimuli, as applied in the dichotic listening paradigm, represents only one specific situation where auditory attention is deployed. It also is possible to attend to a certain part of an auditory stimulus, for instance when attending to a specific part of binaurally presented tone sequences. Some studies used "hierarchical" auditory stimuli consisting of short repetitive sequential patterns that were combined to a long sequential pattern (Bouvet et al., 2011; Bouvet et al., 2014; Justus & List, 2005; List & Justus, 2007, 2010; List et al., 2007; Ouimet, Foster, & Hyde, 2012; Sanders & Poeppel, 2007; Sanders & Astheimer, 2008). Hence, such studies examine the auditory analogy of hierarchical visual stimuli (Ivry & Robertson, 1998; Kimchi, 1992; Navon, 1977, 1981).

Justus and List (2005) used these kinds of tone sequences to investigate attentional persistence. In addition, they examined whether frequency and time were so-called indispensable attributes in the auditory modality (List & Justus, 2007). Auditory hierarchical stimuli consisted of sequences of nine tones (long pattern) that could be subdivided into three 3-tone sequences (short pattern). Each individual tone lasted for 150 ms. The 9-tone sequence as a whole could be arranged in one of four possible ways, as well as the 3-tone sequence (either two consecutive rising changes, two consecutive falling changes, a rising change followed by a falling change, or a falling change followed by a rising change). Two of these arrangements were targets, the other two served as distractors.

Either the long or short pattern contained one of the targets; the other pattern contained one of the distractors (Justus & List, 2005, Experiment 2). Participants had to detect which of the two targets had been presented in a trial (e.g., two rising changes), but they did not know if the target would occur in the short pattern or in the long pattern before the presentation of the auditory stimulus. They thus needed to attend to both patterns to make a decision.

In this experimental paradigm, participants responded faster when the target was presented in the same temporal range (either short or long) as in the previous trial, suggesting attentional persistence, even when the specific target pattern had changed. In addition, participants made fewer errors when the target was presented in the long pattern than when it was presented in the short pattern. These results suggested, first, that participants showed generally better performance when attending to the long pattern than when attending to the short pattern, and second, that performance was worse when participants needed to change their auditory attentional focus from one trial to another than when it remained constant.

Further studies revealed that the non-attended pattern also influenced the processing of the attended pattern, which became evident in congruency effects that showed up as performance costs when target pattern and distractor pattern are incongruent to each other relative to when they are congruent (List, 2006). Several studies revealed greater congruency effects when attending to the short pattern than when attending to the long pattern (Bouvet et al., 2011; Ouimet et al., 2012; Sanders & Poeppel, 2007, Experiments 1 and 2).

One important property of the auditory hierarchical stimuli as described above is the sequential presentation of the tone patterns. Unlike simultaneously presented visual hierarchical stimuli (Navon, 1977), auditory hierarchical stimuli unfold over time. Consequently, the information necessary to classify the repetitive short pattern is available earlier than the information necessary to classify the long pattern. While this is a necessary prerequisite of these kinds of auditory hierarchical patterns, it makes overall response time (RT) differences between the long and the short pattern somewhat difficult to interpret.

For example, one could imagine that participants used the minimum information possible to classify the 9-tone sequences, which is earlier in the short-pattern condition than in the long-pattern condition. Notably, such a strategic temporal order bias would effectively lead to a shorter interval between the cue and the critical pattern-discriminating tone in the short-pattern condition than in the long-pattern condition. If so, RT switch costs should be reduced for the long pattern, because preparation would be more advanced with the long pattern, for which there is certainty once the second tone has passed by, than with the short pattern.

Moreover, there could be reduced RT congruency effects in the short pattern condition compared with the long pattern condition, due to the later presentation of interfering information from the long pattern when attending to the short pattern than vice versa. However, so far there is little empirical support for a strong impact of temporal order on auditory congruency effects. Previous studies showed symmetric congruency effects or even greater congruency effects in the short pattern than in the long pattern (Bouvet et al., 2011; Ouimet et al., 2012; Sanders & Poeppel, 2007).

Some authors dealt with the issue of temporal order by reporting RTs from the point in time when the minimum information to solve the task is available (Ouimet et al., 2012). Yet, such an approach suffers from requiring specific assumptions about participants' strategies, which may vary and depend on factors such as musical experience. Therefore, overall RT difference between the long and the short pattern should be interpreted carefully and considered in light of the specific characteristics of the sequential presentation.

Goals of the present study

We examined mechanisms of intentional control of the auditory attentional focus. Indeed, the sequential level-repetition priming effect (Justus & List, 2005) also could be due to passive "inertia" (i.e., persistence) of the previously established auditory attentional focus rather than to an active attentional focus shifting process. Therefore, it is important to examine active preparation for shifts in auditory focus to different temporal patterns. To do this, we adapted Koch et al.'s (2011) cued attention switching approach to investigate if shifting the focus to a certain auditory short vs. long pattern, which is cued before the presentation of an auditory tone sequence, may rely on active and intentional processes. Therefore, we used 9-tone sequences that were similar to the sequences described above (Justus & List, 2005), with the long and the short pattern either rising or falling. Critically, before the stimulus occurred, a cue instructed participants to attend to either the long or the short pattern (Meiran, 1996; Jost et al., 2013, for review). That is, the task did not resemble an auditory search task that required attending to both patterns until a target was detected (Justus & List, 2005), because participants could prepare for the selection of the cued relevant temporal pattern and could completely ignore the irrelevant temporal pattern. Our study thus targeted a novel research question, because it addressed mechanisms of intentional control of the auditory focus to a temporally limited part of a tone sequence.

Notably, attention to an auditory pattern within a sequence of tones differs in several aspects from attention to one of two dichotically presented stimuli (Koch et al., 2011). For instance, Koch et al. (2011) used dichotic listening and presented stimuli that belonged to different categories (such as male

vs. female voices) and are spatially distinct, which is not the case for long or short patterns of the same stimulus. In addition, stimuli are presented simultaneously in dichotic listening situations and require attenuation of the distracter and/or enhancement of the target, which also is not the case when attending to auditory patterns of a sequence. Therefore, the present study examined a different research question, focusing on auditory attention switching of temporal levels in tone sequences.

Overview of the current experiments

Adjusting the attentional auditory focus in auditory patterns could potentially be related to two types of performance impairments, namely mixing costs and switch costs. In addition, we were interested in whether attention switches can be prepared before the presentation of the auditory stimulus. Therefore, we manipulated preparation time (i.e., the time between the cue and the auditory stimulus [CSI]).

In Experiment 1, we investigated auditory mixing costs and switch costs. In previous studies, the processing of long auditory patterns was related to better performance than the processing of short auditory patterns (Justus & List, 2005). We therefore expected asymmetric costs between the two patterns, with two possibilities for the direction of the asymmetry. Either switch costs would be smaller for attending to the long pattern, because this represents the default and should therefore be rather easily processed. Alternatively, if the processing of the (default) long pattern is inhibited when attending to the short pattern, residual inhibition might cause larger switch costs when shifting the focus back to the global pattern (Allport et al., 1994; for reviews see also Koch et al., 2010; Monsell, Yeung, & Azuma, 2000). In both cases, switch costs would be asymmetric, with larger switch costs for either the short or the long pattern. Experiment 1 was aimed at deciding between these opposing scenarios.

In Experiment 2, we were interested in preparatory mechanisms of attention switches and therefore varied the time that participants could use to prepare for the auditory attentional focus of the next trial (i.e., the CSI). We explicitly targeted the active adjustment of the attentional focus, in contrast to potentially passive processes of sequential level repetition priming. In addition, in both experiments, we also examined congruency effects, which reflect the involuntary processing of irrelevant stimulus aspects, so that patterns of asymmetric interference can inform us about processing biases. Importantly, by investigating mixing costs and switch costs as well as congruency effects, we targeted two different important aspects of auditory attention, namely cognitive control of the auditory attentional focus and involuntary processing of irrelevant information.

Experiment 1

The goal of Experiment 1 was to investigate task-switching when participants attended to a specific auditory pattern within the same auditory stimulus. Participants listened to sequences of tones and they attended either to the entire pattern (long) or to the repetitive shorter pattern (short). The auditory attentional focus either varied from trial to trial ("mixed blocks") or remained constant within an experimental block. We used a 1:1 cue-to-task mapping in the current experiment because the objective was to investigate if mixing costs and switch costs could be found at all using the present task requirements.

Method

Participants Twenty-four participants participated in Experiment 1. Three participants with an excessive number of errors (> 40%) in either the long-pattern or the short-pattern condition were replaced by new participants. The final 24 participants had a mean age of 25 years (SD = 5 years, range: 19-36 years), 17 were female, and 22 were right-handed. None of them reported any hearing problems. On average, participants had 9 years (SD = 3) of musical training during their school education. One participant reported that she saw herself as a musician. Participants gave informed consent and received partial course credit or 8 € for their participation.

Stimuli and apparatus Visual cues were presented at the center of a 17-inch monitor with white background. The participants' distance to the screen was about 60 cm. The cues were a blue and an orange asterisk that were 6 mm in width and 6 mm in length.

Auditory stimuli were sequences of 9 tones that were chosen from a set of 30 different tones. The fundamental frequencies of the tones in Hz were 155, 165, 176, 188, 201, 215, 230, 246, 263, 281, 300, 320, 342, 365, 390, 416, 444, 474, 506, 540, 576, 615, 657, 701, 748, 798, 852, 910, 971, and 1037. Tones consisted of three harmonics with decreasing intensity (1/number of harmonics). We chose tones that are not related to western musical scale to avoid associations with implicit or explicit musical knowledge (see also Trehub, Schellenberg, & Kamenetsky, 1999). Each tone lasted for 200 ms, including onset and offset ramps of 10-ms each. Tones were adjusted for subjective loudness. Three-tone patterns were built from these tones, such that there were steps of four tones between adjacent tones. The intertone interval was 0 ms. These patterns could be rising or falling, for example the tones with the frequencies 155 Hz, 201 Hz, and 263 Hz would built a rising pattern. Three of these three-tone patterns were then combined to the nine-tone sequence, such that the three-tone patterns always had the same structure, hence all of them were rising or all of them were falling. The first tones of the individual

three-tone patterns were three steps apart and could be combined in either a rising or a falling way, independently from the direction within the three-tone patterns. Four kinds of nine-tone sequences were constructed this way (Fig. 1 depicts a schematic description of the stimuli). First, the short three-tone patterns and the long nine-tone sequence could be both rising (congruent), which would for example result in a sequence comprising the following frequencies: 155, 201, 263/ 188, 246, 320/230, 300, 390 Hz. Second, the short three tone patterns and the long nine-tone sequence could both be falling (congruent), which would for example result in a sequence comprising the following frequencies: 390, 300, 230/320, 246, 188/263, 201, 155 Hz. Third, the short three-tone patterns could be falling and the long nine-tone sequence rising (incongruent), which would for example result in a sequence comprising the following frequencies: 263, 201, 155/320, 246, 188/390, 300, 230 Hz. Last, the short three-tone patterns could be rising and the long nine-tone sequence falling (incongruent), which would for example result in a sequence comprising the following frequencies: 230, 300, 390/188, 246, 320/155, 201, 263 Hz. Auditory stimuli were presented via headphones (Grundig 38629 DJ Headphones). All stimuli were presented with E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

Participants responded with "c" (left index finger) and "m" (right index finger) on the computer keyboard (QWERTZ). They were asked to indicate if the attended pattern was falling or rising by pressing "c" when the attended pattern was falling and "m" when the pattern was rising, thus the mapping between "rising/falling" was compatible with the spatial position of the response keys.

Procedure Each trial started with the visual cue (an orange or blue asterisk) that remained on the screen until the

participants' response. The color of the cue indicated the auditory attentional focus, such that the mapping of the colors to the auditory attentional foci was counterbalanced over participants. After 500 ms, the auditory stimulus started (cue-stimulus interval, CSI). Participants had maximally 4,400 ms from onset of the auditory stimulus to indicate if the pattern of the relevant auditory attentional focus was rising or falling. The visual cue remained on screen until the response was made. In case of an error, the word Fehler! (German for error) was displayed in red color on the center of the screen for 500 ms. In case of no response after 4,400 ms, the word Schneller! (German for faster) was displayed in red color on the center of the screen for 500 ms. After a blank of 500 ms (response-cue interval, RCI), the next trial started (Fig. 2).

Participants completed 12 experimental blocks of 40 trials each. Eight blocks were mixed blocks with the auditory attentional focus varying randomly from trial to trial (as indicated by the color cue). During 2 blocks, participants were instructed to respond to the long pattern only, and during the remaining 2 blocks, participants were instructed to respond to the short pattern only (pure blocks). We chose twice as many trials in mixed blocks as in pure blocks to compare conditions with equal number of trials in the mixing-costs contrast, in which we would compare pure blocks with the repetition trials of the mixing blocks only. The order of the blocks was counterbalanced over participants with two mixed blocks alternating with one pure block. Half of the participants started with a pure block, and the other half started with a mixed block. Half of the participants attended to the long pattern in their first pure block, and the other half attended to the short pattern in their first pure block. Before the experimental blocks, participants completed four practice blocks with eight trials each. Two practice blocks were mixed blocks, two were pure blocks, and the order was counterbalanced over participants.

Participants reported demographic data and musical expertise before the experiment and were asked about strategies after the experiment. Participants were instructed orally and
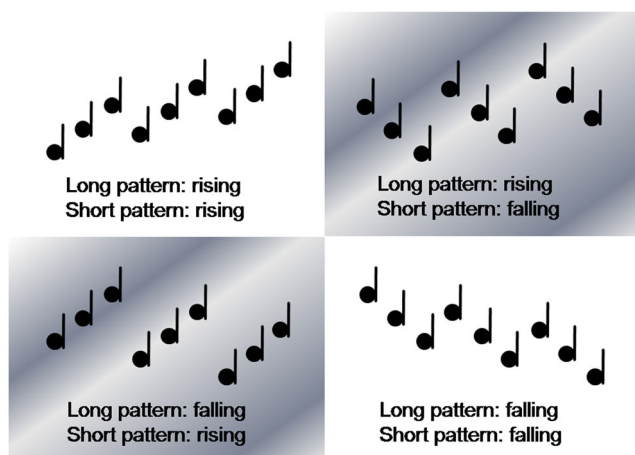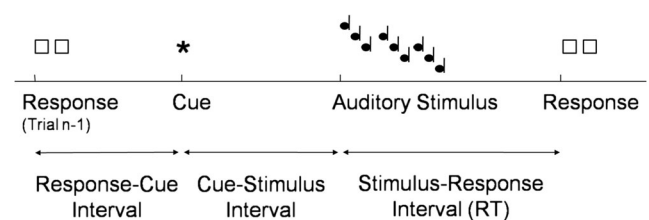

Fig. 1 Schematic depiction of the four stimulus types. Both patterns could be rising, the long pattern could be rising and the short pattern falling, the long pattern could be falling and the short pattern rising, or both patterns could be falling. The sequences with grey background were incongruent; the sequences with white background were congruent.

Fig. 2 Trial procedure. A visual cue instructed participants to either attend to the long or to the short pattern. The visual cue was presented on the screen until the participants' response. After a cue-stimulus interval of 500 ms in Experiment 1 (100 ms or 1,000 ms in Experiment 2), the auditory stimulus was presented. Participants indicated if the attended pattern was rising or falling by pressing one of two buttons. After a response-cue interval of 500 ms in Experiment 1 (1,000 ms or 100 ms in Experiment 2), the cue of the next trial was presented.

with written instructions on the computer screen. The total experiment lasted about 45 minutes.

Design Independent variables were auditory attentional focus (long pattern, short pattern), transition (pure, repetition, switch), and congruency (congruent, incongruent). Dependent variables were RTs and errors.

We analyzed two non-orthogonal contrasts on RTs and errors. First, we analyzed the *mixing-costs contrast* with the independent variables auditory attentional focus (long pattern, short pattern), transition (pure, repetition), and congruency (congruent, incongruent). Only trials from pure blocks and trials from mixed blocks with an immediate repetition of the auditory attentional focus were used for this contrast (Kiesel et al., 2010).

Second, the *switch-costs contrast* was analyzed. Only trials from mixed blocks were used for the switch costs contrast, including switch trials and repetition trials. Independent variables were auditory attentional focus (long pattern, short pattern), transition (repetition, switch), and congruency (congruent, incongruent).

Results

Practice trials, the first trial of each block, error trials, and trials following errors were excluded from the analysis of the RTs, as well as outliers (RT ± 3 SD from the mean of each condition). Practice trials, the first trial of each block, and trials following errors were excluded from the analysis of the error rates.

Mixing costs contrast

Reaction times We conducted a 2x2x2 ANOVA with the within-subject variables auditory attentional focus (long pattern, short pattern), transition (pure, repetition), and congruency (congruent, incongruent) on RTs. The short pattern could be classified at the onset of the second tone (200 ms after stimulus onset) and the long pattern at the onset of the fourth tone (600 ms after stimulus onset). Figure 3a depicts the uncorrected RTs, and Fig. 3b depicts the RTs corrected by 200 ms or 600 ms, respectively (all analyses are based on the uncorrected RTs). The ANOVA revealed a main effect of auditory attentional focus, $F(1, 23) = 56.12$, $MSE = 222977$, $p < 0.001$, $\eta_p^2 = 0.71$, indicating that RTs were 510 ms slower for the long pattern than for the short pattern (1,491 ms vs. 981 ms). The main effect of transition was not significant, $F(1, 23) = 1.15$, $MSE = 17,304$, $p > 0.29$, $\eta_p^2 = 0.05$, suggesting that there were no overall mixing costs. The interaction of transition and auditory attentional focus was a not significant either, $F(1, 23) = 3.26$, $MSE = 24,642$, $p > 0.08$, $\eta_p^2 = 0.12$, even though mixing costs
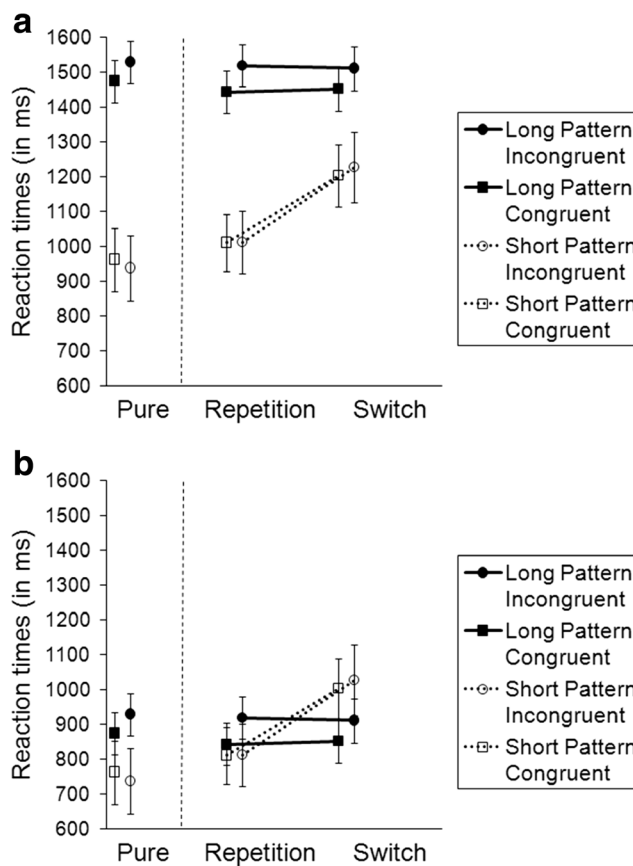

Fig. 3 Response times of Experiment 1. Only in the short-pattern condition, participants responded more slowly in switch trials than in repetition trials. The error bars represent standard error. **a** Response times are measured from the onset of the sequence. **b** Response times are measured from the onset of the second tone for the short-pattern condition and from the onset of the fourth tone for the long-pattern condition.

were somewhat greater for the short-pattern condition than for the long-pattern condition.

The main effect of congruency was not significant, $F(1, 23) = 2.18$, $MSE = 15,622$, $p > 0.15$, $\eta_p^2 = 0.09$, but the interaction of congruency and auditory attentional focus was significant, $F(1, 23) = 8.35$, $MSE = 8,525$, $p < 0.01$, $\eta_p^2 = 0.27$, indicating larger congruency effects in the long-pattern condition than in the short pattern condition. Indeed, only in the long-pattern condition did participants respond faster in congruent trials than in incongruent trials, 1,459 ms vs. 1,524 ms; congruency effect of 65 ms, $t(23) = -3.62$, $p < 0.01$, whereas in the short-pattern condition there was no significant congruency effect, 987 ms vs. 975 ms, $t(23) = 0.45$, $p > 0.66$. The interaction of transition and congruency, $F(1, 23) = 1.30$, $MSE = 5340$, $p > 0.26$, $\eta_p^2 = 0.05$, and the three-way interaction were not significant, $F < 1$.

Errors We conducted the same ANOVA on error rates (Fig. 4). The ANOVA revealed a main effect of auditory attentional focus, $F(1, 23) = 6.00$, $MSE = 0.021$, $p < 0.03$, $\eta_p^2 = 0.20$, indicating
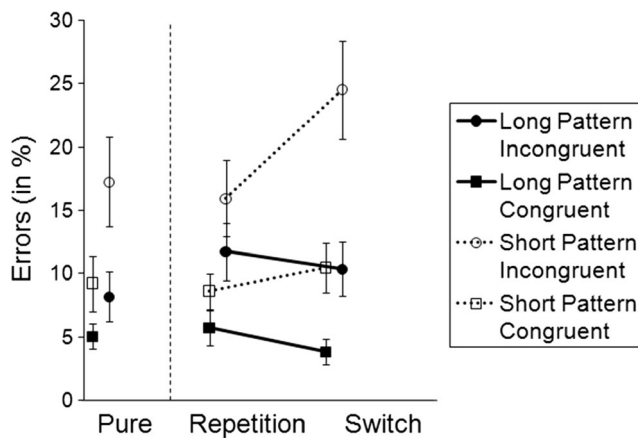
Fig. 4 Error rates of Experiment 1. Participants made more errors in the short-pattern condition than in the long pattern condition. Only in the short-pattern condition, participants made more errors in switch trials than in repetition trials. Participants made more errors in incongruent trials than in congruent trials. The error bars represent standard error.

that error rates were smaller for the long pattern than for the short pattern: 7.7% vs. 12.8%. Note that the smaller error rates and the slower responses in the long-pattern condition compared with the short-pattern condition might suggest a speed-accuracy trade-off. However, as explained in the introduction, it is possible that the slower responses in the RTs to the long pattern might be partly attributable to the sequential presentation of the sequences (see also Experiment 2).[1]

The main effect of transition was not significant, $F < 1$, showing no evidence for mixing costs. The interaction of transition and auditory attentional focus was not significant either, $F(1, 23) = 3.67$, $MSE = 0.003$, $p > 0.06$, $\eta_p^2 = 0.14$, even though mixing costs were somewhat larger for the long-pattern condition than for the short-pattern condition. Note that this nonsignificant trend was in the opposite direction of the non-significant trend in the RTs, and both failed to reach the significance threshold, so that they may not represent robust findings. Mixing costs in RTs and errors were thus small and non-significant for both auditory attentional foci.

In addition, the main effect of congruency was significant, $F(1, 23) = 10.73$, $MSE = 0.017$, $p < 0.01$, $\eta_p^2 = 0.32$, indicating that participants made fewer errors in congruent trials than in incongruent trials (7.2% vs. 13.3%). The interaction of auditory attentional focus and congruency was not significant, $F(1, 23) = 1.39$, $MSE = 0.008$, $p > .25$, $\eta^2 = 0.06$, just as all other effects, $Fs < 1$.

<hr />

[1] If there was a speed-accuracy trade off, RTs and error rates should be negatively correlated. However, RTs and error rates were positively correlated or not correlated at all, for the overall RTs and errors, $r = 0.71$, $p < 0.001$, as well as for both the long pattern, $r = 0.27$, $p > 0.19$, and for the short pattern, $r = 0.73$, $p < 0.001$.

Switch costs contrast

Reaction times We conducted a 2x2x2 ANOVA with the within-subject variables auditory attentional focus (long pattern, short pattern), transition (repetition, switch), and congruency (congruent, incongruent) on RTs (Fig. 3). The ANOVA revealed a main effect of auditory attentional focus, $F(1, 23) = 37.75$, $MSE = 172320$, $p < 0.001$, $\eta_p^2 = 0.62$, indicating that RTs were 368 ms slower for the long pattern than for the short pattern (1,481 ms vs. 1,113 ms). The main effect of transition was significant, $F(1, 23) = 39.75$, $MSE = 12,505$, $p < 0.001$, $\eta_p^2 = 0.63$. Importantly, the interaction of transition and auditory attentional focus also was significant, $F(1, 23) = 45.14$, $MSE = 10,962$, $p < 0.001$, $\eta_p^2 = 0.66$. Indeed, in the short-pattern condition, RTs were 203 ms higher in switch trials than in repetition trials, 1,215 ms vs. 1,012 ms, $t(23) = -7.52$, $p < 0.001$, whereas in the long-pattern condition, there were no switch costs at all, 1,481 ms vs. 1,481 ms, $t(23) = -0.02$, $p < 0.98$.

In addition, the main effect of congruency was significant, $F(1, 23) = 6.56$, $MSE = 11735$, $p < 0.02$, $\eta_p^2 = 0.22$, indicating that participants responded faster in congruent trials than in incongruent trials (1,277 ms vs. 1,317 ms). The interaction of auditory attentional focus and congruency: $F(1, 23) = 4.24$, $MSE = 8,479$, $p > 0.05$, $\eta_p^2 = 0.16$ was not significant but suggested slightly greater congruency effects for the long-pattern condition than for the short-pattern condition. The interaction of transition and congruency was not significant, $F < 1$. The three-way interaction was not significant either, $F(1, 23) = 1.29$, $MSE = 3,814$, $p > 0.27$, $\eta_p^2 = 0.05$.

Errors We conducted the same ANOVA on error rates (Fig. 4). The ANOVA revealed a main effect of auditory attentional focus, $F(1, 23) = 15.71$, $MSE = 0.015$, $p < 0.001$, $\eta_p^2 = 0.41$, indicating that error rates were smaller for the long pattern than for the short pattern (7.9% vs. 14.9%). The main effect of transition was not significant, $F(1, 23) = 3.92$, $MSE = 0.004$, $p > 0.05$, $\eta_p^2 = 0.15$, but the interaction of auditory attentional focus and transition was significant, $F(1, 23) = 13.39$, $MSE = 0.004$, $p < 0.001$, $\eta_p^2 = 0.37$, indicating that switch costs were smaller in the long-pattern condition than in the short-pattern condition. Indeed, in the long-pattern condition, there was no significant difference of repetition trials and switch trials, 8.7% vs. 7.1%, $t(23) = 1.67$, $p > 0.10$, whereas in the short-pattern condition, participants made less errors in repetition trials than in switch trials, 12.3% vs. 17.5%, $t(23) = -3.37$, $p < 0.01$. Note that the asymmetric switch costs in the errors rates confirm the pattern in the RTs (see *Discussion*).

In addition, the main effect of congruency was significant, $F(1, 23) = 19.22$, $MSE = 0.018$, $p < 0.001$, $\eta_p^2 = 0.46$, indicating that participants made less errors in congruent trials than in incongruent trials (7.2% vs. 15.6%). The interaction of transition and congruency was not significant, $F(1, 23) =$

4.13, $MSE = 0.004$, $p > 0.05$, $\eta_p^2 = 0.15$, with slightly smaller congruency effects in repetition trails than in switch trials. The interaction of auditory attentional focus and congruency was not significant, $F(1, 23) = 2.25$, $MSE = 0.011$, $p > 0.14$, $\eta_p^2 = 0.09$. The three-way interaction was not significant either, $F(1, 23) = 1.88$, $MSE = 0.006$, $p > 0.18$, $\eta_p^2 = 0.06$.

Discussion

In this experiment, we examined mixing costs and switch costs as empirical markers for intentional control of attentional focus on short vs. long auditory patterns. In addition, we examined congruency effects as an empirical marker for involuntary processing of task-irrelevant (i.e., noninstructed) information. Overall, we found no significant mixing costs when using performance in experimental blocks with constant attentional focus as baseline, suggesting that working-memory load in repetition trials plays a minor role in auditory selection of tone patterns. However, we observed clear switch costs in the mixed blocks, which were markedly asymmetric, with large switch costs when switching to the short pattern and basically no switch costs at all when switching to the long pattern. We observed this asymmetry in both RT and error rates, confirming each other. Hence, the asymmetric switch costs in RTs and errors suggest more efficient adjustment of auditory attention when attending to the long pattern than when attending to the short pattern, possibly indicating that attending to the longer pattern represents the default processing mode in situations with short and long patterns.

However, as the 9-tone sequences unfolded over time with the short pattern being faster to be identified than the long pattern, the question arises if the asymmetric switch costs in the RTs could be related to the temporal structure of the sequences. Hence, it is possible that participants used the minimum information possible, which is the second tone in the short pattern but the fourth tone in the long pattern. This would lead to a shorter interval between the cue and the informative tone in the short-pattern than in the long pattern, so that there would be effectively a longer preparation time for a switch to the long pattern. If preparation reduces the switch costs, then we would expect smaller switch costs for the long pattern than for the short pattern. Yet, we propose that the observed switch-cost asymmetry is at least partly, or even largely, due to attending to the long pattern representing the default processing mode rather than on critical differences in time-based, for two reasons.

First, the pattern of congruency effects is not in line with such an account. Specifically, if participants based their performance on the minimal information required to discriminate the short and the long auditory pattern, then participants should be able to select their responses much earlier (by about 400 ms) when attending to the short pattern. If so, then the response should be selected in many cases even before the

discriminating information for the long pattern (i.e., the fourth tone) becomes available, which should result in clearly asymmetric congruency effects. However, the pattern of congruency effects does not clearly support this account. In the mixing costs contrast, there is no such asymmetry in the error rates, even though it is present in the RT data. Note though that we did not find mixing costs in the first place, for both attentional foci. Moreover, in the switch costs contrast, this asymmetry of the congruency effect was nonsignificant both in the RT data and the error rates, suggesting that congruency effects occurred for both the short and the long pattern. This pattern of results appears to be in line with the idea that participants actually often tend to wait until the information about the identity of the long pattern is available, which would be expected by the idea that attending to the long pattern represents the default processing mode.

Second, the possible assumption that participants are simply better prepared for a switch to the long pattern because they have relatively more time for preparing such a switch is post hoc. In Experiment 2, we explicitly test the influence of preparation on switch costs by manipulating the preparation interval (i.e., the CSI). To foreshadow the results of Experiment 2, the pattern of preparation effects does not conform to predictions derived from the idea that participants in Experiment 1 were simply better prepared for a switch to the long pattern.

In sum, in Experiment 1 we found asymmetric switch costs of the two auditory attentional foci. This finding suggests the idea that switch costs reflect primarily an attentional shifting process, and that attending to the long pattern represents the default processing mode and therefore shows only little (if any) switch costs. In Experiment 2, we investigated the question if switch costs could be reduced with increased preparation time for the attention switch. The preparatory reduction of switch costs would give further support to the notion of active attention shifting after the cue presentation and before the presentation of the tone sequence, which would further differentiate the results of the present study from carryover effects in terms of sequential-level priming (Justus & List, 2005; List & Justus, 2007).

Experiment 2

The goal of Experiment 2 was to enhance our understanding of the asymmetric switch costs that we found in Experiment 1. We wanted to examine if the switch costs in the short-pattern condition could be reduced if participants had more time to prepare for the attention switch. Because we did not observe switch costs for the long pattern in Experiment 1, we thus predicted a three-way interaction of attentional focus, transition, and CSI. Importantly, this interaction would strongly

point to an active process of auditory attentional shifting that takes place before the onset of the tone sequence.

In Experiment 2, the design was similar to the mixed blocks of Experiment 1. We did not use pure blocks anymore, because we did not find any mixing costs in Experiment 1. We used two different CSIs in Experiment 2 (100 ms and 1,000 ms) to increase or reduce the time to prepare for a switch relative to the cuing interval in Experiment 1 (CSI = 500 ms). We used a 1:1 cue-to-task mapping again because the results of Experiment 1 did not suggest any role of visual cue-priming effects, as such cue-priming benefits also should have been observed in the long-pattern condition, for which we did not find any effect of switch vs. repetition of the cue (hence ruling out the presence of general cue repetition priming).

Method

Participants Twenty-four new participants participated in Experiment 2. One participant with an excessive number of errors (> 40%) in the short-pattern condition was replaced by a new participant. One of the final 24 participants did not report her age, the remaining 23 had a mean age of 22 years (SD = 4 years, range: 19-35 years), 17 of 24 were female, and 21 of 24 were right-handed. None of them reported any hearing problems. On average, participants had 9 years (SD = 3) of musical training during their school education. Eight participants reported that they saw themselves as musicians. Participants gave informed consent and received partial course credit or 8 € for their participation.

Stimuli, apparatus, and procedure All stimuli, apparatus, and procedures were identical to Experiment 1, except that the CSI varied randomly from trial to trial (100 ms vs. 1000 ms) instead of being 500 ms (constant CSI in Experiment 1). As a consequence, the RCI was varied inversely, so that it was 1000 ms when the CSI was 100 ms, and it was 100 ms when the CSI was 1000 ms, thus resulting in a constant response-stimulus interval of 1100 ms.

Participants completed 10 experimental blocks of 64 trials each. All blocks were mixed blocks with the auditory attentional focus varying randomly from trial to trial. Before the experimental blocks, participants completed two practice blocks with 16 trials each. The total experiment lasted approximately 45 minutes.

Design Independent variables were auditory attentional focus (long pattern, short pattern), transition (repetition, switch), congruency (congruent, incongruent), and CSI (100 ms, 1000 ms). Dependent variables were reaction times (RTs) and errors.

Results

Practice trials, the first trial of each block, error trials, and trials following errors were excluded from the analysis of the RTs, as well as outliers (RT ± 3 SD from the mean of each condition). Practice trials, the first trial of each block, and trials following errors were excluded from the analysis of the error rates.

Reaction times We conducted a 2x2x2x2 ANOVA with the within-subject variables auditory attentional focus (long pattern, short pattern), transition (repetition, switch), congruency (congruent, incongruent), and CSI (100 ms, 1000 ms) on RTs. Figure 5a depicts the uncorrected RTs, and Fig. 5b depicts the RTs corrected by 200 ms or 600 ms, respectively (all analyses are based on the uncorrected RTs).

The ANOVA revealed a main effect of auditory attentional focus, $F(1, 23) = 28.36$, $MSE = 26,8554$, $p < 0.001$, $\eta_p^2 = 0.55$, indicating that RTs were 382 ms slower for the long pattern than for the short pattern (1,463 ms vs. 1,181 ms). The main effect of transition was significant, $F(1, 23) = 83.11$, $MSE = 11,172$, $p < 0.001$, $\eta_p^2 = 0.78$. Importantly, the interaction of auditory attentional focus and transition also was significant, $F(1, 23) = 92.93$, $MSE = 11,826$, $p < 0.001$, $\eta_p^2 = 0.80$, showing that there were switch costs of 205 ms in the short-pattern condition (1,284 ms vs. 1,079 ms), $t(23) = -9.97$, $p < 0.001$, but no significant switch costs (−8 ms) ms in the long-pattern condition (1,467 ms vs. 1,459 ms), $t(23) = 1.17$, $p > 0.26$.

In addition, there was a main effect of congruency, $F(1, 23) = 14.18$, $MSE = 20,698$, $p < 0.01$, $\eta_p^2 = 0.38$. This indicates that participants responded faster in congruent trials than in incongruent trials (1,295 ms vs. 1,350 ms). Congruency effects were slightly greater in switch trials than in repetition trials in the short-pattern condition only, but the corresponding three-way interaction of auditory attentional focus, transition, and congruency was not significant, $F(1, 23) = 3.44$, $MSE = 7,052$, $p > 0.07$, $\eta_p^2 = 0.13$.

So far, the aforementioned pattern of results is similar to the data pattern of Experiment 1. Now we turn to the influence of the CSI. The main effect of CSI was significant, $F(1, 23) = 47.64$, $MSE = 10,799$, $p < 0.001$, $\eta_p^2 = 0.67$, as well as the interaction of auditory attentional focus and CSI, $F(1, 23) = 84.96$, $MSE = 5,406$, $p < 0.001$, $\eta_p^2 = 0.79$, indicating that there was a larger reduction of overall RTs with more preparation time in the short-pattern condition than in the long-pattern condition. Indeed, in the short-pattern condition, participants responded 143 ms faster when the CSI was 1000 ms than when it was 100 ms, 1,110 ms vs. 1,253 ms; $t(23) = -8.42$, $p < 0.001$, but in the long-pattern condition there was no significant effect of CSI: 1,461 ms vs. 1,465 ms; $t(23) = -0.56$, $p > 0.57$. Moreover, the interaction of transition and CSI was not significant, $F(1, 23) = 3.68$, $MSE = 6,941$, $p >$
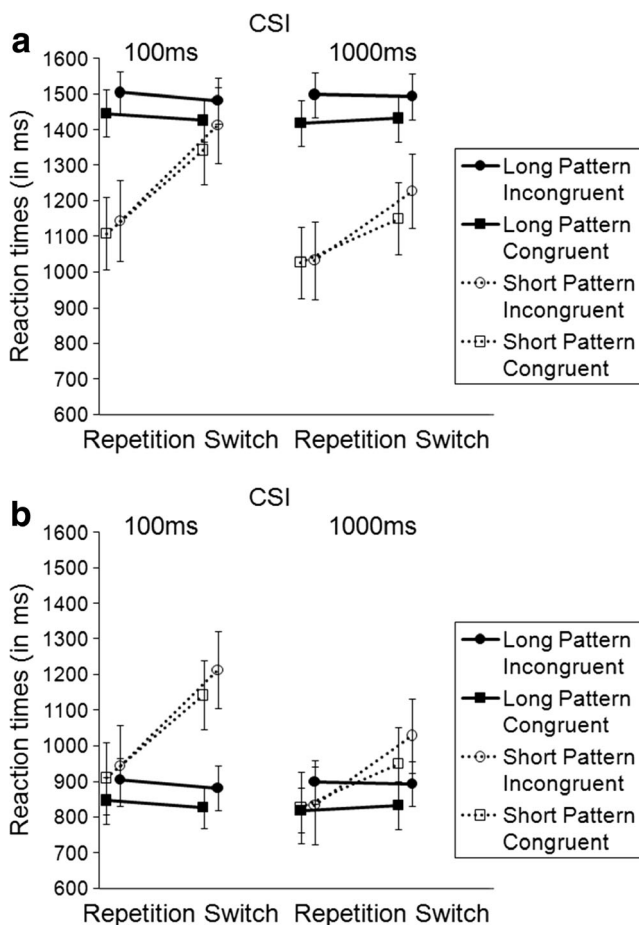
Fig. 5 Response times of Experiment 2. Only in the short-pattern condition, participants responded more slowly when the CSI was 100 ms than when it was 1000 ms. In addition, in the short-pattern condition, participants responded more slowly in switch trials than in repetition trials, and even more so when the CSI was 100 ms than when the CSI was 1,000 ms. The error bars represent standard error. **a** Response times are measured from the onset of the sequence. **b** Response times are measured from the onset of the second tone for the short-pattern condition and from the onset of the fourth tone for the long-pattern condition.

0.06, $\eta_p^2 = 0.14$, but the expected three-way interaction of auditory attentional focus, transition, and CSI was significant, $F(1, 23) = 8.15$, $MSE = 10242$, $p < 0.01$, $\eta_p^2 = 0.26$. All other effects were not significant, interaction of transition and congruency, $F(1, 23) = 1.90$, $MSE = 5,858$, $p > 0.18$, $\eta_p^2 = 0.08$; all other $Fs < 1$.

To decompose the significant three-way interaction of auditory attentional focus, transition, and CSI, we analyzed the long-pattern condition and the short-pattern condition separately. We conducted 2x2 ANOVAs with the variables transition and CSI. Figure 6 shows a summary of the switch costs in both Experiment 1 and Experiment 2.

In the short-pattern condition, there was a significant main effect of transition, $F(1, 23) = 99.41$, $MSE = 10,179$, $p < 0.001$, $\eta_p^2 = 0.81$, a significant main effect of CSI, $F(1, 23) = 70.86$, $MSE = 6,866$, $p < 0.001$, $\eta_p^2 = 0.76$, and a significant interaction of transition and CSI, $F(1, 23) = 6.54$, $MSE =$
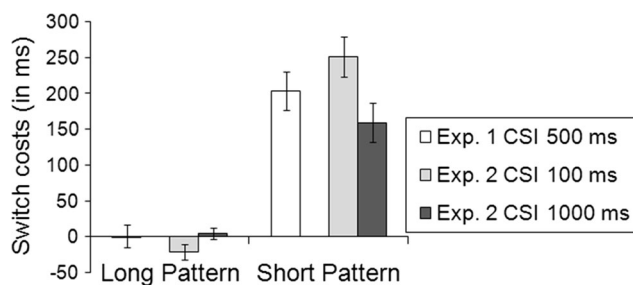

Fig. 6 Switch costs of Experiment 1 and Experiment 2 in RTs. In the short-pattern condition, participants responded faster in repetition trials than in switch trials (Experiment 1, mixed blocks, and Experiment 2). This difference was reduced when participants had more time to prepare for the switch (Experiment 2). In the long-pattern condition, there were no significant switch costs in either condition. The error bars represent standard error.

7,696, $p < .02$, $\eta_p^2 = 0.22$, indicating that switch costs were smaller when the CSI was 1000 ms than when the CSI was 100 ms (160 ms vs. 251 ms).

In contrast, in the long-pattern condition, neither the main effect of transition, $F(1, 23) = 1.36$, $MSE = 1,319$, $p > 0.25$, $\eta_p^2 = 0.06$, nor of CSI, $F < 1$, was significant. The interaction of transition and CSI, $F(1, 23) = 4.64$, $MSE = 896$, $p < 0.05$, $\eta_p^2 = 0.17$, was significant, but this was due to an opposing pattern of switch effects in the two CSI conditions. However, switch costs were not significantly different from 0 ms in either condition, neither when the CSI was 100 ms, −22 ms (inversed switch costs), $t(23) = −1.93$, $p > 0.06$, nor when the CSI was 1,000 ms (4 ms, $t < 1$).

Errors We conducted the same ANOVA on error rates (Fig. 7). The ANOVA revealed a main effect of auditory attentional focus, $F(1, 23) = 18.15$, $MSE = 0.027$, $p < 0.001$, $\eta_p^2 = 0.44$, indicating that error rates were smaller for the long pattern than for the short pattern (5.4% vs. 12.6%).

The main effect of transition was significant, $F(1, 23) = 7.40$, $MSE = 0.003$, $p < 0.02$, $\eta_p^2 = 0.24$. Importantly, the interaction of transition and auditory attentional focus was also significant, $F(1, 23) = 15.86$, $MSE = 0.004$, $p < 0.001$, $\eta_p^2 = 0.41$, indicating that switch costs were smaller in the long-pattern condition than in the short pattern condition. Indeed, in the long-pattern condition, there was no significant difference between repetition trials and switch trials, 5.9% vs. 5.0%, $t(23) = 1.50$, $p > 0.14$, whereas in the short-pattern condition participants made significantly less errors in repetition trials than in switch trials, 10.6% vs. 14.6%, $t(23) = −3.98$, $p < 0.001$. Figure 8 shows a summary of the switch costs in both Experiment 1 and Experiment 2.

The main effect of congruency was significant, $F(1, 23) = 50.34$, $MSE = 0.010$, $p < 0.001$, $\eta_p^2 = 0.69$, indicating smaller error rates in congruent trials than in incongruent trials (5.4% vs. 12.6%). The interaction of auditory attentional focus and congruency was significant as well, $F(1, 23) = 8.02$, $MSE = 0.009$, $p < 0.01$, $\eta_p^2 = 0.26$, indicating larger congruency
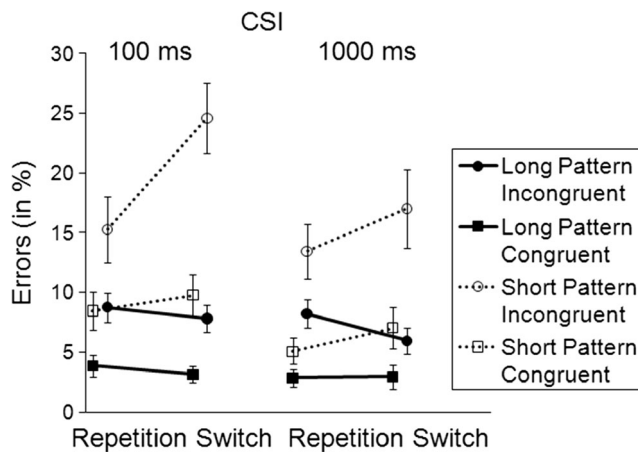
Fig. 7 Error rates of Experiment 2. Participants made more errors in the short-pattern condition than in the long pattern condition. Participants made more errors in incongruent trials than in congruent trials, especially in the short-pattern condition. Only in the short-pattern condition, participants made more errors in switch trials than in repetition trials. In addition, in the short-pattern condition, participants made more errors when the CSI was 100 ms than when it was 1000 ms. The error bars represent standard error.

effects in the short-pattern condition, 10.0% (7.6% vs. 17.6%) than in the long-pattern condition, 4.5% (3.2% vs. 7.7%).

In addition, there was a significant three-way interaction of auditory attentional focus, transition, and congruency, $F(1, 23) = 10.20$, $MSE = 0.002$, $p < 0.01$, $\eta_p^2 = 0.31$. To decompose this three-way interaction, we analyzed the long-pattern condition and the short-pattern condition separately by conducting 2x2 ANOVAs with the variables transition and congruency. In the long-pattern condition, there was a significant main effect of congruency, $F(1, 23) = 50.53$, $MSE = 0.001$, $p < .001$, $_p^2 = .69$. The main effect of transition, $F(1, 23) = 2.23$, $MSE = 0.001$, $p > 0.14$, $\eta_p^2 = 0.09$, as well as the interaction of transition and congruency, $F(1, 23) = 1.86$, $MSE = 0.001$, $p > 0.18$, $\eta_p^2 = 0.08$, were not significant. In contrast, in the short-pattern condition, there was a significant main effect of transition, $F(1, 23) = 15.87$, $MSE = 0.002$, $p < 0.001$, $\eta_p^2 = 0.41$, a significant main effect of congruency, $F(1, 23) = 27.96$, $MSE = 0.009$, $p < 0.001$, $\eta_p^2 = 0.55$, as well
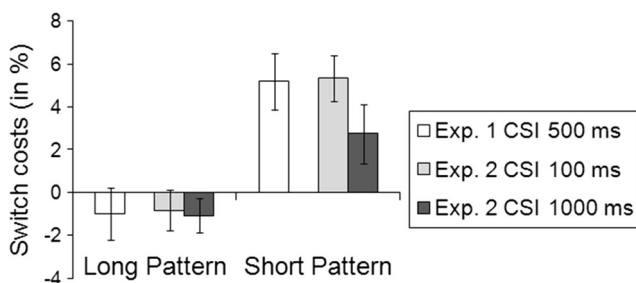
as a significant interaction of transition and congruency, $F(1, 23) = 7.32$, $MSE = 0.002$, $p < 0.02$, $\eta_p^2 = 0.24$, indicating larger congruency effects in switch trials (12.4%) than in repetition trials (7.6%). Note that we did not find this influence of attention switching on the congruency effect consistently in the RT and error rates of Experiment 1 nor in the RT of Experiment 2, so that it should not be overemphasized. Yet, this influence is not highly consistent in more typical task-switching studies. In task-switching studies, the influence of task switching on the congruency effect is typically attributed to an increased vulnerability to distracting information in case of a task switch. This is because the new task is not yet fully prepared (or primed by its actual repetition, as in repetition trials; see Kiesel et al., 2010, for a review).

Now we turn to the influence of preparation (i.e., of the CSI). The main effect of CSI was significant, $F(1, 23) = 29.67$, $MSE = 0.002$, $p < 0.001$, $\eta_p^2 = 0.56$, as well as the interaction of auditory attentional focus and CSI, $F(1, 23) = 10.02$, $MSE = 0.002$, $p < 0.01$, $\eta_p^2 = 0.30$, indicating that there was a larger reduction of error rates with more preparation time in the short-pattern condition than in the long-pattern condition. Indeed, in the short-pattern condition participants made less errors when the CSI was 1000 ms than when the CSI was 100 ms, 10.7% vs. 14.5%, $t(23) = -5.35$, $p < 0.001$, whereas in the long-pattern condition there was no significant difference between CSI 1,000 ms and CSI 100 ms, 5.0% vs. 5.9%, $t(23) = -1.63$, $p > 0.11$.

The interaction of transition, congruency, and CSI was not significant, $F(1, 23) = 3.16$, $MSE = 0.004$, $p > 0.08$, $\eta_p^2 = 0.12$, but, numerically, the difference of congruency effects in switch and repetitions trials was somewhat larger with the short CSI than with the long CSI. Likewise, the interaction of auditory attentional focus, transition, and CSI did not reach significance, $F(1, 23) = 2.06$, $MSE = 0.002$, $p > 0.16$, $\eta_p^2 = 0.08$, but the data show a similar pattern as in the RTs, with decreased switch costs with increased CSI in the short-pattern condition. All other effects were not significant: interaction of transition and congruency, $F(1, 23) = 2.77$, $MSE = 0.003$, $p > 0.10$, $\eta_p^2 = 0.10$; interaction of transition and CSI, $F(1, 23) = 2.00$, $MSE = 0.002$, $p > 0.17$, $\eta_p^2 = 0.08$; interaction of congruency and CSI, $F(1, 23) = 1.22$, $MSE = 0.002$, $p > 0.28$, $\eta_p^2 = 0.05$; interaction of auditory attentional focus, congruency, and CSI, $F < 1$; four-way interaction, $F(1, 23) = 1.73$, $MSE = 0.002$, $p > 0.20$, $\eta_p^2 = 0.07$.

Discussion

In Experiment 2, we found the predicted influence of preparation on the switch costs, which was greater for the short-pattern condition than for the long-pattern condition. As discussed for Experiment 1, due to the sequential presentation of the auditory stimuli, participants might have used the minimum information possible to classify the patterns (i.e., the

Fig. 8 Switch costs of Experiment 1 and Experiment 2 in error rates. In the short-pattern condition, participants made less error in repetition trials than in switch trials (Experiment 1, mixed blocks, and Experiment 2). In the long-pattern condition, there were no significant switch costs in either condition. The error bars represent standard error.

first and the second tone in the short-pattern condition and the first and the fourth tone in the long-pattern sequence). The switch costs in the long-pattern condition thus could have been smaller than in the short-pattern condition because the delay in the availability of information in the long-pattern condition might have been used for preparation for the attention switch. The manipulation of the CSI in Experiment 2 provided additional information to assess the role of this putative strategy use for the asymmetric switch costs in the RTs.

Overall, the pattern in the RTs suggests that specific strategy use related to the sequential presentation of the tones alone cannot explain the observed asymmetry of the switch costs and the reduction of switch costs. Specifically, the timing of the sequence and the duration of the CSI created four different intervals between cue onset and the onset of the critical tone (second tone vs. fourth tone). In the short-pattern condition, the minimum time needed (i.e., the interval between the cue and the 2nd tone) when the CSI is short is 300 ms (100 ms + 200 ms) and it is 1,200 ms (1000 ms + 200 ms) when the CSI is long. In comparison, in the long-pattern condition the minimum time needed (the interval between the cue and the 4th tone) when the CSI is short is 700 ms (100 ms + 600 ms) and it is 1,600 ms (1,000 ms + 600 ms) when the CSI is long. However, in the short-pattern condition there were still substantial "residual" switch costs (Meiran, 2000; Rogers & Monsell, 1995) of 160 ms even when the CSI was long (i.e., 1,200 ms total time for preparation). On the other hand, in the short CSI condition of the long-pattern the total time for preparation was actually much shorter (700 ms), but there were still no switch costs in the long-pattern condition. Taken together, this differential pattern of preparation effects clearly speaks against an account that would attribute the observed switch-cost asymmetries for the short vs. long pattern condition to a strategic processing bias that is due to the sequential presentation of the patterns, which inevitably results in earlier presentation of the short pattern. Instead, the data suggest that attending to the long pattern represents the default processing mode, so that no substantial costs occur when participants switch back to this mode.

Moreover, manipulation of the CSI showed that switching focus to attending to the short pattern can be achieved effectively, at least to some degree (considering the substantial residual switch costs with long CSI) if there is sufficient time for advance preparation. That is, the data suggest that the reduction of switch costs in the short-pattern condition with more preparation time indicates an active process of attention shifting that starts before the presentation of the tone sequence.

Finally, the overall larger error rates in the short-pattern condition as well as the asymmetric congruency effects are in line with previous studies, which suggest an attentional bias in favor of the long pattern (Bouvet et al., 2011; Justus & List, 2005; Ouimet et al., 2012; Sanders & Poeppel, 2007). Attention to the short pattern thus seems to be less efficient (and possibly also more prone to interference) than attention to the long pattern.

General discussion

The goal of the present study was to investigate task switching in sequential auditory tone patterns. We were especially interested in intentional control of the auditory attentional focus within an auditory sequence. Participants attended to a long or to a short tone pattern and the auditory attentional focus could vary from trial to trial. We targeted mechanisms of cognitive control of the auditory attentional focus, as assessed with switch costs and preparation effects, and on stimulus-driven involuntary attention shifting, as assessed with congruency effects.

Major findings

Two experiments revealed asymmetric switch costs in RTs and errors. Specifically, switch costs were only present when switching from the long pattern to the short pattern and not the other way round. In addition, Experiment 2 showed that the switch costs were reduced when participants had more time to prepare for the attentional switch to the short pattern. Notably, in Experiment 1 we did not observe mixing costs.

The current experiments also revealed congruency effects when attending to both the short and the long pattern. Yet, the pattern of congruency effects was less consistent, showing no clear asymmetry with respect to the short vs. long pattern across experiments (Experiments 1 and 2), contrasts (mixing costs and switch costs), and dependent measures (RT and error rates).

Attending to temporal patterns within a sequence of tones vs. attending to one of two simultaneously presented stimuli

In the present study, we investigated auditory attention switching to temporal patterns within the same tone sequence. Recently, intentional control of auditory attention has been investigated in other studies as well, but these previous studies focused on shifting between two simultaneously presented auditory stimuli (spoken number words) and used dichotic listening (Koch & Lawo, 2014, 2015; Koch et al., 2011; Lawo & Koch, 2014, 2015; Lawo et al., 2014). Dichotic listening studies did not observe mixing costs consistently. For instance, Koch and Lawo (2015) observed mixing costs in dichotic listening only when participants selected speakers by ear, not by gender. This inconsistent pattern of mixing costs is in line with our Experiment 1, in which we did not observe any mixing costs. The absence of mixing costs suggests that switching-related performance impairments were less due to general processes related to working memory load (i.e., maintenance of attentional focus). Note that our experimental procedure was designed to isolate active shifting of the auditory attentional focus,

whereas other components such as stimulus-response mappings or classification rules were held constant. This should have decreased memory load compared to other task-switching experiments.

In addition, we observed switch costs in the short-pattern condition only, whereas previous studies using dichotic listening have usually observed general switch costs in RTs (Koch et al., 2011). Note that in the present experiments, cue transitions when using a 1:1 cue-to-attentional focus mapping do not seem to have any impact on attentional switch costs because perceptual cue repetition priming should have caused switch costs (i.e., cue repetition benefits) even for the long pattern, which we clearly did not observe (Logan & Bundesen, 2003; see Jost et al., 2013, for a review). Therefore, it is particularly important that we found a clear reduction of switch costs in the short-pattern condition with more preparation time before a switch. This suggests successful active preparation before the onset of the stimulus, which has not systematically been found with dichotically presented tones, at least not when a 2:1 cue-to-task mapping was used to avoid visual priming effects due to immediate cue repetitions (Koch et al., 2011; Lawo et al., 2014; Lawo & Koch, 2015).

Please note that the sequential presentation of the tones made a specific preparatory strategy possible. Participants may have used the first and second tone to identify the short sequence and the first and fourth tone to identify the long sequence. This strategy use may have increased the asymmetry in the switch costs, as the interval between the cue and the last attended tone was shorter in the short-pattern condition than in the long-pattern condition. However, this account is not supported by the pattern of congruency effects, and it is rather invalidated by the pattern of switch costs in Experiment 2. In Experiment 2, we found clear residual switch costs for the short pattern under temporal conditions that would allow even more preparation time than in the short CSI conditions in the long pattern condition (where there were no switch costs), suggesting that the apparent absence of switch costs in the long-pattern condition does not simply reflect the benefit of increased preparation time but rather the benefit of the fact that attending to the long pattern represents the default processing mode (without being able to explain the entire pattern on its own, as discussed above, especially in the error rates). Thus, the relative contribution of differential attentional efficiency to the asymmetric switch-costs in the RTs is not completely resolved.

Task-goal setting and stimulus selection

The results of the current study revealed a partly dissociative pattern between switch costs and congruency effects. Congruency effects were observed in both temporal patterns. Finding clear congruency effects in the short pattern condition suggests that participants most often do not respond before they have identified the long pattern, enabling response crosstalk and thus congruency effects. Likewise, the finding of clear

congruency effects in the long pattern suggests that even for the default mode of processing it is hard to ignore simultaneously available information that is mapped to the same set of responses. In the present study, participants needed to extract structural information from a sequence of tones in which each tone is part of both temporal patterns, and interference from the irrelevant pattern seems to be hard to avoid, even when participants know beforehand which pattern they have to attend to. Thus, stimulus selection did not seem to be perfect, and the preponderance of congruency effects in the error rates might suggest that participants sometimes select the wrong pattern and then they do not correct the selection within the same trial.

Switch costs, on the other hand, were strongly asymmetric and subject to a highly consistent influence of preparation time. These results seem to be related to active preparation of the attentional focus on the short temporal pattern. In comparison, in a search task without explicit attentional cues, no asymmetry in sequential-level priming effects was observed both for the long pattern and the short pattern (Justus & List, 2005). This suggests that intentional and flexible attention shifting between temporal patterns of a tone sequence is at least partly independent from involuntary attention capture (as reflected in congruency effects, see also Lawo & Koch, 2014). Thus, the present results add to the existing evidence for dissociative components of task sets by showing a partial dissociation of switch costs and congruency effects (see Regev & Meiran, 2016, for a recent study that aimed to disentangle the mechanisms of distinct task-set components).

Attending to the long auditory pattern or to the short auditory pattern

In addition to cognitive control in auditory task switching, the results of the present study suggested differences between attending to the long pattern or to the short pattern. In general, our findings were in favor of a relative advantage in attending to the long pattern compared with attending to the short pattern. First, participants made more errors when attending to the short pattern than when attending to the long pattern. This also was the case in congruent trials where errors were not due to the selection of the wrong attentional focus, suggesting that identifying the direction of the pattern was more difficult in the short-pattern condition than in the long pattern condition. We would thus like to argue that structural information is by default extracted more easily from the long pattern than from the short pattern, which might even be similar to global-local processing in vision (Bouvet et al., 2011). However, because participants sometimes responded before the end of the tone sequence when attending to the short pattern, the long pattern was most likely not processed *before* the short pattern, as first suggested for the processing of simultaneously presented global and local visual stimuli (Navon, 1977).

Regarding dynamic shifting between auditory attentional foci, the current results suggest that attention to the long auditory pattern is more efficiently done than attending to the short auditory pattern. The attentional focus of the previous trial or preparation time did not seem to influence attention to the long pattern, whereas switching from the long pattern to the short pattern was related to substantial switch costs. This corroborates the notion that attending to the long pattern can occur in a more efficient way than attending to the short pattern.

Note, however, that some studies in the task-switching literature found greater switch costs for the more dominant task (Allport et al., 1994; for reviews see Koch et al., 2010; Monsell et al., 2000). To explain such switch-cost asymmetries, it has been argued that persisting inhibition of the dominant task is the reason for this asymmetry. However, those studies examined task switches that include changes in the stimulus-response mappings. In the present set of auditory tasks (i.e., attentional foci), persisting inhibition seems less plausible because, according to previous research (Bouvet et al., 2011; Justus & List, 2005; Ouimet et al., 2012; Sanders & Poeppel, 2007), there should actually be an attentional bias in favor of the long pattern. It rather seems that attending to the long pattern is not impaired by the previous auditory attentional focus; however, it becomes even more difficult to attend to the short pattern after an attention switch.

Importantly, attention shifting to the short pattern improved with more preparation time. This could be due to participants preparing for the temporal structure of the sequence that helped to extract the information on the short pattern (Astheimer & Sanders, 2009; Sanders & Astheimer, 2008). Another possible explanation is that participants used the time in between cue and stimulus to build a mental template of the short pattern (Cusack & Carlyon, 2003; Cusack et al., 2004, see also Bregman, 1990).

It would be interesting to see if theoretical notions like "biased competition" (Desimone & Duncan, 1995) or "attentional weighting" (Meiran et al., 2008) could be adapted to the present sequential selection situation. Relational information was needed to perform the task, and some tones were more informative than others for either pattern, even though every tone of the sequence was part of both the short and the long pattern. If the most informative tones were enhanced and/or the less informative ones suppressed, one would not expect the strong asymmetric pattern of results for the long and the short pattern that our data revealed. Temporal structure and processing prevalence therefore must be considered in theoretical accounts on attending to relational information in sequences.

Conclusions

The current study was designed to investigate auditory task switching with the emphasis on intentional control of the attentional focus on a cued sequential auditory pattern in tone sequences. Extracting information and flexible attention shifting depend on the temporal structure of the relevant part of the auditory sequence. In general, extracting information from a short pattern is more error-prone and less efficient than integrating information over a long pattern. However, because there is a preparation benefit for the short pattern, attending to a part of the sequence can be more efficient if one knows beforehand how the stimulus is structured and which part contains the information that is relevant to us, especially when we need to shift attention. This suggests distinct mechanisms for attending to a certain part of a sequential auditory stimulus and for attending to one auditory stimulus amongst distractors. Future research is needed to further clarify these mechanisms.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Hillsdale, NJ: Erlbaum.

Astheimer, L. B., & Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological Psychology, 80*(1), 23–34.

Bouvet, L., Rousset, S., Valdois, S., & Donnadieu, S. (2011). Global precedence effect in vision and in audition: Evidence for cognitive styles across modalities. *Acta Psychologica, 138*(2), 329–335.

Bouvet, L., Simard-Meilleur, A. A., Paignon, A., Mottron, L., & Donnadieu, S. (2014). Auditory local bias and reduced global interference in autism. *Cognition, 131*(3), 367–72.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39,* 713–726.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Broadbent, D. E. (1958). *Perception and communication*. Oxford: Pergamon Press.

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics, 77,* 1465–1487.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25,* 975–979.

Cusack, R., & Carlyon, R. P. (2003). Perceptual asymmetries in audition. *Journal of Experimental Psychology: Human Perception and Performance, 29*(3), 713–725.

Cusack, R., Deeks, J., Aikman, G., & Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analyses. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 613–656.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

Hirsch, P., Nolden, S., & Koch, I. (in press). Higher-order cognitive control in dual tasks: Evidence from task-pair switching. *Journal of Experimental Psychology: Human Perception and Performance.*

Hugdahl, K. (2011). Fifty years of dichotic listening research- Still going and going and… *Brain and Cognition [Dichotic Listening Anniversary Special Issue], 76*(2), 211-213.

Ivry, R., & Robertson, L. C. (1998). *The two sides of perception.* Cambridge: MIT Press.

Jost, K., De Baene, W., Koch, I., & Brass, M. (2013). A review of the role of cue processing in task switching. *Journal of Psychology/ Zeitschrift fuer Psychologie, 221*(1), 5–14.

Justus, T., & List, A. (2005). Auditory attention to frequency and time: An analogy to visual local-global stimuli. *Cognition, 98*(1), 31–51.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching - a review. *Psychological Bulletin, 136*(5), 849–874.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin, 112*(1), 24–38.

Koch, I., & Lawo, V. (2014). Exploring temporal dissipation of attention settings in auditory task switching. *Attention, Perception & Psychophysics, 76*(1), 73–80.

Koch, I., & Lawo, V. (2015). The flip side of the auditory spatial selection benefit: Larger attentional mixing costs for target selection by ear than by gender in auditory task switching. *Experimental Psychology, 62*(1), 66–74.

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review, 17*(1), 1–14.

Koch, I., Lawo, V., Fels, J., & Vorländer, M. (2011). Switching in the cocktail party: Exploring intentional control of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1140–1147.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review, 111*(4), 880–913.

Lawo, V., & Koch, I. (2014). Dissociable effects of auditory attention switching and stimulus-response compatibility. *Psychological Research [Special Issue Auditory Attention], 78,* 379–386.

Lawo, V., & Koch, I. (2015). Attention and action: The role of response mappings in auditory attention switching. *Journal of Cognitive Psychology, 27*(2), 194–206.

Lawo, V., Fels, J., Oberem, J., & Koch, I. (2014). Intentional attention switching in dichotic listening: Exploring the efficiency of nonspatial and spatial selection. *The Quarterly Journal of Experimental Psychology, 67,* 2010–2024.

Liefooghe, B., Barrouillet, P., Vandierendonck, A., & Camos, V. (2008). Working memory costs of task switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(3), 478–494.

List, A. (2006). Investigating auditory local-global processing using visual research as a model (Doctoral dissertation). University of California, Berkeley, Berkeley, CA, USA.

List, A., & Justus, T. (2007). Auditory priming of frequency and temporal information: Effects of lateralised presentation. *Laterality, 12*(6), 507–35.

List, A., & Justus, T. (2010). Relative priming of temporal local-global levels in auditory hierarchical stimuli. *Attention, Perception, & Psychophysics, 72*(1), 193–208.

List, A., Justus, T., Robertson, L. C., & Bentin, S. (2007). A mismatch negativity study of local-global auditory processing. *Brain Research, 1153,* 122–33.

Logan, G. D. (2005). The time it takes to switch attention. *Psychonomic Bulletin and Review, 12*(4), 647–653.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance, 29*(3), 575–599.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1423–1442.

Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research, 63,* 234–249.

Meiran, N., Kessler, Y., & Adi-Japha, E. (2008). Control by action representation and input selection (CARIS): A theoretical framework for task switching. *Psychological Research, 72*(5), 473–500.

Mondor, T. A., & Bregman, A. S. (1994). Allocating attention to frequency regions. *Perception & Psychophysics, 56*(3), 268–276.

Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 66–79.

Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research, 63,* 250–264.

Navon, D. (1977). Forest before trees - precedence of global features in visual-perception. *Cognitive Psychology, 9,* 353–383.

Navon, D. (1981). The forest revisited: More on global precedence. *Psychological Research, 43*(1), 1–32.

Ouimet, T., Foster, N. E., & Hyde, K. L. (2012). Auditory global-local processing: Effects of attention and musical experience. *Journal of the Acoustical Society of America, 132,* 2536–2544.

Regev, S., & Meiran, N. (2016). Cue response dissociates inhibitory processes: Task identity information is related to backward inhibition but not competitor rule suppression. *Psychological Research, 80*(4), 625–639.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

Sanders, L. D., & Astheimer, L. B. (2008). Temporally selective attention modulates early perceptual processing: Event-related potential evidence. *Perception & Psychophysics, 70*(4), 732–742.

Sanders, L. D., & Poeppel, D. (2007). Local and global auditory processing: Behavioral and ERP evidence. *Neuropsychologia, 45*(6), 1172–1186.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology, 18*(7), 559–572.

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences, 12*(5), 182–186.

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 965–975.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin, 136*(4), 601–626.