The final article will be available, upon publication, via its DOI: 10.1037/xhp0000432

Attentional State Modulates the Effect of an Irrelevant Stimulus Dimension on Perception

Björn Herrmann and Ingrid S. Johnsrude
The University of Western Ontario

Paper accepted for publication in *Journal of Experimental Psychology: Human Perception and Performance* (March 17th 2017)

Author Note:
Björn Herrmann, Department of Psychology & The Brain and Mind Institute, The University of Western Ontario; Ingrid S. Johnsrude, Department of Psychology & The Brain and Mind Institute, School of Communication Sciences & Disorders, The University of Western Ontario.

Research was supported by the Canadian Institutes of Health Research (MOP133450 to I.S. Johnsrude), and the Brain and Mind Institute at the University of Western Ontario (postdoctoral fellowship award to B. Herrmann). We thank Jackie Tsang, Kristian McCarthy and Suvarna Moharir for their help during data acquisition. We also thank Alexander J. Billig for his helpful feedback on a previous version of the manuscript.

Correspondence concerning this article should be addressed to Björn Herrmann, Department of Psychology, Brain & Mind Institute, The University of Western Ontario, London, Ontario, N6A 5B7, Canada. E-mail: herrmann.b@gmail.com

Abstract

Co-variations of acoustic features provide redundancy in rapidly changing soundscapes: Hearing one feature enables a listener to infer another if these two features normally co-vary. However, it is unknown whether situational demands affect the degree to which co-variations influence perceptual inferences. We exploited a perceptual interdependency between modulation rate and frequency and examined, in 6 experiments, whether challenging situations would alter the degree to which people rely on frequency information to make decisions about modulation rate. Participants listened to amplitude-modulated (AM) sounds with modulation rates (~5 Hz) either decreasing or increasing over time, and identified the direction of the rate change. Participants were instructed to ignore carrier frequency, which either decreased or increased (~1300 Hz) over time. We observed that participants were more likely to perceive the modulation rate as slowing down when frequency decreased and as speeding up when frequency increased (AM-rate change illusion). The magnitude of the illusion increased when uninformative cues (compared to informative cues) prohibited regulation of attention to sounds, and under distraction introduced by a concurrent visual motion-tracking task. The evidence suggests that the attentional state affects how strongly people rely on featural co-variations to make perceptual inferences.

Keywords: Time-frequency illusion, amplitude-modulation rate, auditory perception, attention, distractor task

Significance Statement

Redundancy in sounds supports auditory perception in complex listening situations. Hearing one sound feature allows inferences about another sound feature if these two features normally correlate in everyday sounds (e.g., speech or music). The current set of experiments tested whether the degree to which listeners rely on learned correlations between sound features depends on cognitive factors and situational demands. Participants judged whether amplitude modulation rate in sounds they heard slowed down or sped up, while ignoring concurrent changes in the sound's frequency. The data show that when the listener is distracted or unable to predict the nature of the upcoming sound, he or she relies more strongly on learned correlations between sound features for sound perception.

Introduction

The sounds common in everyday life, such as speech and music, are acoustically complex but redundant in that acoustic features co-vary. For example, temporal modulation rate and frequency change together in consistent ways; increases in the rate of speech and music are associated with increases in fundamental frequency and note frequency, respectively (Broze & Huron, 2013; Topbas, Orlikoff, & St. Louis, 2012). Extraction of redundancy from acoustic environments is thought to optimize sensory representations and perception (Kluender, Stilp, & Kiefte, 2013; Lewicki, 2002; Smith & Lewicki, 2006). Indeed, perceptual experiences can be shaped by featural co-variations occurring in sounds. For example, speech rate is perceived as faster when vocal frequency or intensity increase (Bond & Feldstein, 1981; Feldstein & Bond, 1981). Similarly, the speed of musical pieces is perceived as faster when note frequency or intensity increase (Boltz, 1998, 2011). And yet, do cognitive states and situational demands alter the degree to which perceptual decisions are shaped by correlated features?

When people make perceptual decisions about a particular stimulus dimension (e.g., modulate-rate change), they sometimes rely upon more information than simply what is present in that dimension. They also rely on other stimulus dimensions that co-vary with the dimension of interest. Several previous investigations have focused on the above-mentioned influence of frequency or intensity on temporal judgments (Alards-Tomalin, Leboe-McGowan, & Mondor, 2013; Boltz, 1998, 2011; Bond & Feldstein, 1981; Crowder & Neath, 1994; Feldstein & Bond, 1981; Henry & McAuley, 2009, 2013; Herrmann, Henry, Grigutsch, & Obleser, 2013; Herrmann, Henry, Scharinger, & Obleser, 2014; Pfeuty & Peretz, 2010; Shigeno, 1986, 1993), observing, for example, that a sound's tempo is perceived as faster when fundamental frequency or intensity increase. The reverse influence has been reported as well, where perceptual decisions about fundamental frequency (heard as pitch) are affected by concurrent changes in the sound duration or the interval between sounds (Henry & McAuley, 2013; Henry, McAuley, & Zaleha, 2009); for example, an interval between sounds is perceived as longer when the sounds differ more in fundamental frequency. Furthermore, perceptual decisions about a sound's loudness are influenced by changes in fundamental frequency; sounds are perceived as louder when frequency increases (Neuhoff, McBeath, & Wanzie, 1999). We use the term "perceptual interdependencies" to refer to instances when perception of a particular stimulus quality is influenced by another, irrelevant but correlated stimulus dimension.

Perceptual interdependencies such as those described here have mainly been attributed to experience with featural co-variations in natural environments: Hearing one acoustic feature allows a listener to infer another if these two features normally co-vary (Boltz, 2011; Bond & Feldstein, 1981; Feldstein & Bond, 1981; Neuhoff, 2004; Neuhoff, et al., 1999; Walsh, 2003). For example, perceptual interdependency between intensity and frequency of sounds might be related to experience with moving sound sources. A source producing a static sound (i.e., stable in frequency, intensity, and spectrum) systematically changes in intensity, frequency, and spectrum at a listener's position when it is moving towards the listener (Neuhoff, 2004). Perceptual interdependency between a sound's temporal aspects (e.g., modulation rate or duration) and frequency as well as intensity might be due to experience with speech and music. Speech is spoken louder and at a higher fundamental frequency when speech rate increases (Black, 1961; Topbas, et al., 2012); musical notes tend to be higher in frequency for pieces with a fast speed (Broze & Huron, 2013); speech utterances, musical compositions, and syntactic boundaries tend to close with decreases in frequency, intensity, and tempo (Pisoni & Luce, 1987); and production of short accented syllables coincides with larger changes in fundamental frequency (Alain, 1993). Experience with emotional vocal expressions might additionally contribute to the perceptual entanglement of correlated acoustic features (Boltz, 2011); screams, for example, are commonly loud and produced at high frequencies (Green, Whitney, & Potegal, 2011).

Especially when listening conditions are ambiguous, such that a perceptual decision about one feature is difficult (on the basis of that feature alone), information from naturally co-varying dimensions might be used to make perceptual decisions in a direction that is consistent with the learned co-variations. But perception might also be biased under conditions when a perceptual decision about one feature is easy and co-occurrence of another feature is consistent with the learned co-variations. For example, even for a clearly discriminable increase in a sound's frequency-modulation rate, a concurrent frequency increase might lead to an overestimation of the perceived modulation-rate change (Herrmann, et al., 2013). It is, however, not well understood whether reliance on correlated dimensions can be altered by task or attentional demands.

The effects of task or attentional demands on perception are commonly studied using a dual-task paradigm (Dalton, Santangelo, & Spence, 2009; Lavie, 2005; Lavie, Beck, & Konstantinou, 2014; Masutomi, Barascud, Kashino, McDermott, & Chait, 2016; Pashler, 1994). Dual-task paradigms utilize two (or more) conditions that manipulate the cognitive or perceptual load (e.g., low- and high-load conditions) during performance of the perceptual task of interest (the primary task). High cognitive/perceptual load compared to low load has been shown to reduce the likelihood of detecting a visual (Macdonald & Lavie, 2008) or auditory stimulus (Macdonald & Lavie, 2011; Raveh & Lavie, 2015), to reduce the effect of foreperiod on reaction times (Vallesi, Arbula, & Bernardis, 2014), to increase auditory target-to-distractor interference (Dalton, et al., 2009), and to increase the influence of lexicality on phoneme categorization (Mattys & Scharenborg, 2014; Mattys & Wiget, 2011). Distraction is commonly associated with reduced perceptual sensitivity for stimuli from which attention is drawn (Konstantinou & Lavie, 2013; Macdonald & Lavie, 2008, 2011; Molloy, Griffiths, Chait, & Lavie, 2015; Raveh & Lavie, 2015), although it remains a debated topic whether attention influences perceptual sensitivity or decision criteria (de Lange, Rahnev, Donner, & Lau, 2013; Rahnev et al., 2011; Schneider, 2011; Schneider & Komlos, 2008).

Critically, when listening situations are demanding (e.g., under distraction) the influence of prior knowledge on a perceptual decision might increase (Mattys, Brooks, & Cooke, 2009; Mattys & Wiget, 2011). For example, Mattys and colleagues (Mattys & Scharenborg, 2014; Mattys & Wiget, 2011) made use of the Ganong effect (Ganong, 1980) – i.e., the tendency to perceive an ambiguous phoneme so that the percept is consistent with the surrounding lexical context – to examine the effects of cognitive load on the interaction between sensory encoding and lexical knowledge for perception. The authors observed that individuals rely more strongly on lexical knowledge for phoneme identification under high compared to low cognitive load, and suggest that this increase is due to degraded encoding of the physical (acoustic) information (Mattys & Scharenborg, 2014; Mattys & Wiget, 2011).

It is, however, unclear for nonlinguistic auditory dimensions (e.g., modulation rate and frequency) whether the reliance on learned featural co-variations depends on concurrent cognitive demands. The perceptual independency of acoustic features is thought to be the result of sensorineural processes (Herrmann, et al., 2013; Shigeno, 1986), whereas the interaction between lexical and auditory processes as investigated by Mattys et al. (Mattys & Scharenborg, 2014; Mattys & Wiget, 2011) likely involves non-sensory neural systems that support lexical representations (Hickok, 2009; Scott & Johnsrude, 2003; Tyler & Marslen-Wilson, 2008).

In line with previous work (Boltz, 2011; Herrmann, et al., 2013; Herrmann, et al., 2014) we examined the influence of frequency changes on perception of modulation-rate change. Specifically, we hypothesized that this influence increases in unpredictable and in distracting listening situations. We utilized sounds with linearly changing (decreasing or increasing) amplitude modulation (AM) rate and carrier frequency. The changes in the two dimensions were either consistent (i.e., in the same direction) or inconsistent. Experiment 1 demonstrates that changes in carrier frequency lead to illusory AM-rate change percepts. Experiments 2 and 3 investigates whether prior information about the degree to which a sound changes in modulation rate (Experiment 2) and frequency (Experiment 3) reduces biases in AM-rate change perception produced by concurrent changes in frequency. The purpose of Experiments 4–6 was to investigate whether a secondary distractor task increases the bias in AM-rate change perception produced by concurrent changes in frequency. In these experiments, we elucidate how the weights accorded a stimulus dimension of interest, and a correlated dimension, in making perceptual decisions shift as a function of attentional and task demands, thereby clarifying the processes of auditory perception in complex listening situations.

General Methods

Participants

Participants reported no neurological disease or hearing impairment. In Experiments 1–3, audiograms confirmed normal hearing (≤25 dB HL for frequencies ranging from 0.25 to 4 kHz). In Experiments 4–6, audiograms were not measured because of time constraints, but participants were recruited from the same pool of undergraduates as were recruited for Experiments 1–3. All of the participants had normal or corrected-to-normal vision and were naïve to the purposes of the experiment. Participants gave written informed consent prior to the experiment, and either received course credit or were paid $10 CAD per hour for their participation. The study was conducted in accordance with the Declaration of Helsinki and approved by the local General Research Ethics Board of the University of Western Ontario.

Auditory stimuli

In all experiments, participants listened to amplitude-modulated sounds of 4 s duration. The sound's carrier was a sine wave (see below) and the depth of the amplitude modulation was set to 100%. The modulation rate of the sounds either decreased or increased over time. The sound-initial modulation rate was always 5 Hz and changed linearly to one of 10 (Experiment 1) or 8 levels (Experiments 2–6); in each experiment, half of the final levels were lower than 5 Hz and half were higher. The range of amplitude-modulation changes was determined for each participant individually to account for differences in perceptual sensitivity (see below). We focused on a modulation rate of about 5 Hz because this falls into the region of speech-relevant modulation rates (Elliott & Theunissen, 2009; Greenberg, Carvey, Hitchcock, & Chang, 2003).

Please insert Figure 1 around here

Sounds also changed in carrier frequency (heard as pitch) over time, which participants were instructed to ignore (see below). The sound-initial carrier frequency was always a 1300 Hz sine wave and changed linearly to a carrier frequency up to six semitones higher or lower (for half of the trials the carrier frequency increased, for the other half it decreased). The magnitude of frequency change (i.e., number of semitones) varied from trial to trial, but in each experiment the sound-final carrier frequencies spanned 1300 Hz ±6 semitones (semitones were spaced uniformly). In Experiment 1, 48 different sound-final carrier frequencies were presented for each of the 10 levels of AM-rate change. In Experiments 2 and 3, 28 different sound-final carrier frequencies were presented for each of the 8 levels of AM-rate change, and in Experiments 4–6, 20 different sound-final carrier frequencies were presented at each of the 8 levels of AM-rate change. Spectrograms of sample auditory stimuli are displayed in Figure 1.

Procedure

All experimental procedures were carried out in a sound-attenuating booth. Participants sat in front of a Dell LCD computer screen (~75 cm away; 75 Hz repetition rate; 24 inch diagonal) and sounds were presented via Sennheiser (HD 280 pro) headphones and a Steinberg UR22 (Steinberg Media Technologies) external sound card. Stimulation was controlled by a PC (Windows 7, 64 bit) running Psychtoolbox in Matlab (R2015b). Button presses were recorded via a Cambridge Cognition 2-Button Press Pad (Version 2; Cambridge Cognition Ltd).

For all experiments, participants underwent four or five threshold-estimate and training procedures before the main part of the experiment started. First, the individual hearing threshold was determined for a 1300-Hz sine tone using a method-of-limits procedure (Leek, 2011). Tones of 12 s duration either decreased or increased in intensity by 5.4 dB/s over time (decreasing and increasing trials alternated) and participants indicated when they could no longer hear the tone (intensity decrease) or when they started to hear the tone (intensity increase). The mean sound intensity at the time of the button press was noted for 6 decreasing trials and 6 increasing trials, and these were averaged to determine the individual hearing threshold. In the main experiment, sounds were presented at 55 dB above the individual threshold (sensation level).

In the second procedure (training), participants were familiarized with stimuli that decreased or increased in amplitude-modulation rate, and practiced the task they would perform in the main experiment (discrimination of the direction of AM-rate change). That is, participants were shown visual representations of sequences of tone pips with inter-tone intervals that progressively lengthened and shortened (to clarify the concept of "slowing down" and "speeding up" on discrete events first), followed by visual representations of amplitude-modulated waveforms with decreasing and increasing modulation rate. Participants then listened to sample sounds (N≥6; with large modulation-rate changes and feedback after each trial) followed by 8 trials of the discrimination task in which they indicated with button press whether each stimulus was slowing down or speeding up – again, feedback was given after each trial. The experiment was continued only after the participant indicated that he/she understood the meaning of “slowing down” and “speeding up” in the context of amplitude-modulated sounds. (Note that the perceptual interdependency between modulation-rate change and frequency does not depend on the precise response labels used here. In previous work with German participants, we used the German labels “langsamer” and “schneller”, that is, “slower” and “faster”, and observed a perceptual interdependency between modulation-rate and carrier frequency (Herrmann, et al., 2013; Herrmann, et al., 2014).)

Third, a method-of-constant-stimuli procedure was used to determine each participant’s psychometric function relating AM-rate change to perception. Eighty trials were used, in which modulation rates changed from a 5-Hz starting modulation rate to one of 80 linearly spaced modulation rates ranging between 3 Hz and 7 Hz. That is, 40 sounds decreased and 40 sounds increased in modulation rate over time, varying in the degree of change. This time, no feedback was provided after each discrimination (slowing down or speeding up), and this block lasted about 8 min. A logistic function was fitted to the 80 responses (0 – slowing down; 1 – speeding up) as a function of sound-final modulation rate (ranging from 3 Hz to 7 Hz) using a least-squares routine. Based on the fitted function (and the estimated parameters), 10 (for Experiment 1) or 8 (for Experiments 2–6) levels of modulation-rate change were identified for each participant individually (i.e., normalizing for differences in perceptual sensitivity to modulation-rate changes across participants). The equation for the logistic function used was as follows:

$$p = \frac{1}{1 + e^{-r \cdot (x - x_0)}}$$

where p reflects the proportion of “speeding up” responses, r the growth rate (or slope), x the sound-final modulation rate, and x_0 the inflection point (or point of subjective equality, PSE, the x-value or sound-final modulation rate at which the participant gave a “speeding up” response half the time). Participant-specific modulation-rate-change levels were calculated as follows: The just noticeable difference (JND; i.e., half the difference between the x-values – sound-final modulation rates – corresponding to the 0.25 and 0.75 proportion of “speeding up” responses) was calculated from the function fit. Levels of modulation-rate change for the main experiment were then selected to vary linearly between –4 and 4 times the JND, centered on the participant-specific PSE. For Experiment 1, this resulted in sound-final modulation-rate levels corresponding to 0.012, 0.032, 0.080, 0.188, 0.380, 0.620, 0.812, 0.920, 0.968, 0.988 proportion of “speeding up” responses (10 levels). For Experiments 2–6, the procedure resulted in sound-final modulation-rate levels corresponding to 0.012, 0.042, 0.132, 0.348, 0.652, 0.868, 0.958, 0.988 proportion of “speeding up” responses (8 levels).

The fourth procedure before the main experiment comprised a block during which participants were familiarized with the sounds changing in modulation rate and carrier frequency simultaneously. Participants were instructed to ignore frequency changes and to listen only to whether a sound was slowing down or speeding up. Presentation of sound examples was followed by a 16-trial discrimination task in which participants judged whether each sound was slowing down or speeding up. Feedback was provided. The experiment was continued only after the participant indicated that he/she understood the difference between modulation rate changes (“slowing down”, “speeding up”) and the frequency changes (decreasing, increasing), and that the latter had to be ignored.

Finally, in Experiments 2–6, participants were trained in the cueing paradigm (Experiments 2 and 3), or in the distraction paradigm (Experiments 4–6).

General analysis

Single-trial responses were coded using a zero when the participant pressed the button for “slowing down” and a one when the participant pressed the button for “speeding up”. For each modulation-rate-change level, the proportion of “speeding up” responses was calculated, separately for stimuli in which frequency decreased versus increased. A logistic function was fitted to the proportion of “speeding up” responses as a function of sound-final modulation-rate (i.e., the rate to which the sound changed over time; in JND units), separately for frequency decreases and frequency increases using a least-squares routine. Throughout the current study, we used the estimated slope of the logistic function (r in the equation above) as a measure of perceptual sensitivity to the direction of AM-rate change, and thus as a measure that reflects discrimination difficulty; small slope values reflect poor perceptual sensitivity and difficulty in AM-rate change discrimination. The estimated PSE of the logistic function (in JND units; x_0 in the equation above) was used as a measure of bias in AM-rate change percepts induced by featural co-variations (here carrier frequency) (see also Alards-Tomalin, et al., 2013;

Henry & McAuley, 2009, 2013). Effect sizes are provided as partial eta-squared (η_p^2) when an analysis of variance (ANOVA) is reported and $r_{\text{equivalent}}$ (hereafter simply r_e) when a t-test is reported (Rosenthal & Rubin, 2003). $r_{equivalent}$ is equivalent to a Pearson product-moment correlation for two continuous variables, to a point-biserial correlation for one continuous and one dichotomous variable, and to the square root of partial η^2 for ANOVAs.

Experiment 1: Influence of Frequency on AM-Rate Change Discrimination

The purpose of Experiment 1 was to examine whether changes in the carrier frequency of an AM sound systematically affect perception of AM-rate change. Here we focused on amplitude modulations in the speech-relevant range, i.e., at around 5 Hz.

Participants

Twenty-three university students took part in Experiment 1 (13 female; mean age: 21.7 years; range: 18–31 years). Due to technical problems, data from two additional participants were not saved and the participants were thus excluded.

Methods

Participants listened to amplitude-modulated sounds that changed in modulation rate (decreasing or increasing) and carrier frequency (decreasing or increasing). The modulation rate changed from a sound-initial rate of 5 Hz to one of 10 sound-final modulation rate levels (5 decreases, 5 increases) determined for each participant individually. For each modulation rate level, the sound additionally changed from a 1300-Hz sound-initial frequency to one of 48 sound-final frequency levels (different trials; 24 decreases, 24 increases) of different magnitude spanning 1300 Hz ±6 semitones (semitones were spaced uniformly). Participants were instructed to judge whether a sound was slowing down or speeding up over time and to ignore changes in frequency. The experiment was divided into 6 blocks, each of which lasted about 8–9 minutes, and the presented sounds were randomly drawn from the different AM-rate change and frequency conditions without replacement. Participants took short breaks between blocks. Paired sample t-tests tested the difference between frequency decreases and frequency increases, separately for the PSE and the slope parameters of the fitted psychometric functions.

In Experiment 1, participants underwent an additional block of stimulation during which they listened to sounds of 4-s duration that changed only in carrier frequency (i.e., sounds were unmodulated). On 80 trials, sounds changed from a 1300-Hz sound-initial frequency to one of 80 different sound-final frequency levels spanning 1300 Hz ±4 semitones (semitones were spaced uniformly). That is, the carrier frequency differed for each trial; 40 sounds decreased and 40 sounds increased in frequency over time, varying in the degree of change. Participants judged whether sounds were decreasing or increasing in pitch (no feedback). A logistic function was then fitted to the 80 responses (0 – falling pitch; 1 – rising pitch) as a function of sound-final frequency using a least-squares routine. Three of the 23 participants were perfect at discriminating between frequency decreases and frequency increases, and the logistic function fit could not be performed properly. These participants were thus not considered in this analysis. The just-noticeable-difference (JND) to discriminate the direction of frequency change was calculated based on the estimated parameters from the function fit. In order to investigate whether some of the between-participant variance in the AM-rate change illusion can be explained by the participants' frequency-change discrimination abilities, the illusion magnitude (i.e., difference between the PSE to frequency decreases and the PSE to frequency increases) was correlated with the frequency-change discrimination JND (Spearman procedure).

Results and Discussion

Figure 2A shows that participants perceived the sounds as speeding up when the modulation rate increased and as slowing down when the modulation rate decreased. Participants were additionally influenced in their AM-rate change percept by changes in carrier frequency. The PSE was significantly larger when frequency decreased compared to when frequency increased ($t(22) = 4.27$, $p < .001$, $r_e = .674$). In other words, when frequency decreased participants tended to perceive the sound as slowing down and when frequency increased they tended to perceive the sound as speeding up. The difference between the PSE for frequency decreases and the PSE for frequency increases indexes the magnitude of the AM-rate change illusion (single bar graph in Figure 2A). No difference between slopes for frequency decreases and frequency increases was observed ($t(22) = 0.90$, $p = .379$, $r_e = .188$).

Please insert Figure 2 around here

As a confirmatory analysis, we split the data further, separating small and large frequency changes. Small frequency changes were defined as carrier frequency changes within ±3 semitones from the sound-initial frequency of 1300 Hz, and large frequency changes were defined as carrier frequency changes greater than 3 semitones from the sound-initial frequency of 1300 Hz. The difference between the PSE for frequency decreases and the PSE for frequency increases (illusion magnitude) was calculated for small and for large frequency changes. The PSE difference was significantly larger than zero confirming the AM-rate change illusion for small ($t(22) = 4.83$, $p < .001$, $r_e = .717$) and for large ($t(22) = 3.84$, $p < .001$, $r_e = .633$) frequency changes. The illusion magnitude was enhanced for large compared to small frequency changes ($t(22) = 2.28$, $p = .033$, $r_e = .438$; Figure 2A, right).

Finally, the ability to discriminate the direction of a frequency change in 4-s sounds was correlated with the magnitude of the AM-rate change illusion ($r = .466$, $p = .040$). That is, participants that had a smaller frequency-change JND tended to be less influenced in their AM-rate change percept by changes in frequency (i.e., smaller illusion magnitude). Frequency-change discrimination ability (JND) explained about 22 % of variance in AM-rate change illusion magnitudes across listeners.

The results of Experiment 1 demonstrate that perception of AM-rate change is systematically biased by changes in carrier frequency. AM sounds that change in modulation rate within a range spanning 8 JNDs around the PSE are perceived as slowing down when frequency decreases and as speeding up when frequency increases. This illusion is particularly strong when frequency changes are large and the modulation-rate change is small. The results presented here are in line with previous work showing similar effects for frequency-modulated sounds (Herrmann, et al., 2013; Herrmann, et al., 2014). It appears that some of the inter-individual variance in the illusion magnitude (~22 %) can be explained by individual frequency-change discrimination abilities.

Experiments 2 & 3: Information about the Difficulty of the Upcoming Trial Decreases the Influence of Frequency Change on Perception of AM-Rate Change

The aim of Experiments 2 and 3 was to investigate whether prior information related to an upcoming sound's acoustic properties affects the influence of changing frequency on AM-rate change perception. We hypothesized that knowledge about whether a trial would be easy or difficult (informative cue) would reduce the magnitude of the AM-rate change illusion compared to when no such information was provided (uninformative cue). In other words, by providing informative cues we aimed to enable participants to regulate their attention to the upcoming sound depending on its predicted degree of difficulty.

Participants

Sixteen university students who were not tested in Experiment 1 participated in Experiment 2 (13 female; mean age: 21.7 years; range: 18–29 years) and another sixteen students participated in Experiment 3 (8 female, 5 male; mean age: 19.5 years; range: 17–28 years; three participants in Experiment 3 did not provide demographic information but were recruited from the same participant pool). One additional participant was excluded because performance was at chance level even for the largest AM-rate changes.

Methods

Participants listened to 4-s amplitude-modulated sounds. The modulation rate changed from a sound-initial rate of 5 Hz to one of 8 sound-final modulation rate levels (4 decreases, 4 increases) determined for each participant individually. For each modulation rate level, the sound changed from a 1300-Hz sound-initial frequency to one of 28 sound-final frequency levels (different trials; 14 decreases, 14 increases) of different magnitude spanning 1300 Hz ±6 semitones (semitones were spaced uniformly). Participants were instructed to judge whether a sound was slowing down or speeding up over time and to ignore changes in frequency.

Critically, a 1-s visual cue was presented prior to each sound (Figure 3A). The cue was either informative or uninformative (each half the time) with respect to the predicted degree of difficulty of the upcoming sound. In each of the two experiments (Exp. 2 & 3), participants listened to each of the 8 (rate levels) × 28 (frequency levels) sounds twice; once preceded by an informative cue and once preceded by the uninformative cue.

The uninformative cue was the letter ‘N’ (neutral) presented in white and was uninformative with respect to the degree of difficulty of the upcoming sound. The informative cue could either be the letter ‘E’ (for easy) presented in green (on half of the informative-cue trials) or the letter ‘D’ (for difficult) presented in red (on the other half). ‘Easy’ and ‘Difficult’ were defined differently in Experiment 2 and Experiment 3 (and a between-group approach was chosen due to time constraints). In Experiment 2, informative cues indexed the degree of modulation-rate change: the ‘Easy’ cue indicated a large upcoming modulation-rate change (i.e., the 2 largest rate decreases and the 2 largest rate increases) and the ‘Difficult’ cue indicated a small upcoming modulation-rate change (i.e., the 2 smallest rate decreases and the 2 smallest rate increases). In Experiment 3, informative cues indexed the degree of frequency change: the ‘Easy’ cue indicated a small upcoming frequency change (i.e., the 2 smallest frequency decreases and the 2 smallest frequency increases) and the ‘Difficult’ cue indicated a large upcoming frequency change (i.e., the 2 largest frequency decreases and the 2 largest frequency increases). Note that in Experiment 3, the participant was still discriminating AM-rate changes, so “Easy” and “Difficult” referred to the listener’s predicted ability to ignore the irrelevant dimension (i.e., easy to ignore if the frequency difference is small; more difficult to ignore if it is larger).

Participants were informed about the meaning of the cues and to which dimension the informative cue referred before the experiment began. Stimulation was conducted in 6 blocks separated by breaks and conditions were presented randomly.

Logistic functions were fitted to the proportion of “speeding up” responses as a function of sound-final modulation rate (in JND units) separately for frequency decreases and frequency increases, and separately for informative and uninformative cues. The PSEs and slopes were separately fed into a mixed ANOVA with the within-subject factors Cue Type (informative vs. uninformative) and Frequency (decrease vs. increase), and the between-subject factor Cue Dimension (modulation rate vs. frequency).

Results and Discussion

Figure 3B shows the proportion of “speeding up” responses for frequency decreases versus increases, and for informative versus uninformative cues. The magnitude of the frequency induced AM-rate change illusion is displayed for informative and uninformative cues.

The ANOVA calculated for PSEs revealed a main effect of Frequency ($F(1, 30) = 17.43$, $p < .001$, $\eta_p^2 = .367$) which was due to the larger PSE value for frequency decreases compared to frequency increases (i.e., indexing the AM-rate change illusion). The main effect of Cue Dimension ($F(1, 30) = 1.08$, $p = .306$, $\eta_p^2 = .035$) and the main effect of Cue Type were not significant ($F(1,30) = 2.35$, $p = .136$, $\eta_p^2 = .073$). Critically, the Cue Type × Frequency interaction was significant ($F(1, 30) = 6.23$, $p = .018$, $\eta_p^2 = .172$): The illusion magnitude (i.e., the difference between the PSE for frequency decreases and the PSE for frequency increases) was smaller for informative compared to uninformative cues. The Frequency × Cue Dimension interaction and the Cue Type × Frequency × Cue Dimension interaction were not significant (for both, $F < 1.6$, $p > .2$, $\eta_p^2 < .05$), which indicates that cueing the degree of modulation-rate change versus frequency change did not differentially affect the illusion magnitude. Slight changes in PSEs independent of frequency changes (i.e., independent of illusory rate-change perception) were revealed by the Cue Type × Cue Dimension interaction ($F(1,30) = 4.81$, $p = .036$, $\eta_p^2 = .138$): The PSE (averaged across decreasing and increases frequencies) for informative cues was slightly shifted (more positive) with respect to uninformative cues, but only for participants cued for modulation-rate changes.

Please insert Figure 3 around here

The ANOVA testing for effects of experimental factors on slopes did not reveal any significant effects or interactions (for all, $p > .10$, $\eta_p^2 < .1$) with the exception of a significant main effect of Frequency ($F(1,30) = 5.94$, $p = .021$, $\eta_p^2 = .165$), caused by shallower slopes for frequency decreases compared to frequency increases.

The results of Experiment 2 and 3 reveal that information about a sound’s acoustic features can help to reduce the influence of frequency changes on perceived AM-rate. Informative cues might have allowed individuals to specifically increase attentional engagement for trials indicated as being difficult. Because sensitivity to AM-rate changes was unaffected by presenting informative cues (i.e., slopes were not modulated by Cue Type), the cues appear to have facilitated suppression of frequency changes rather than improving sensitivity to AM-rate changes, regardless of whether the magnitude of modulation-rate changes or frequency changes was cued. However, the precise cognitive factor driving the reduction in biased AM-rate change percepts cannot be inferred based on Experiments 2 and 3 alone. Experiments 4-6 tested the effects cognitive factors on the illusion magnitude directly.

Experiment 4: Distraction by a Visual Memory Task

The aim of Experiment 4 was to investigate whether cognitive load as induced by a visual memory task increases the influence of changes in carrier frequency on perceived AM-rate changes.

Participants

Fifteen university students who were not tested in Experiments 1-3 participated in Experiment 4 (7 female; mean age: 19.5 years; range: 18–27 years). Data from one additional participant were excluded due to technical problems during recording.

Methods

The amplitude-modulated sounds were similar to those in Experiments 2 and 3 with the exception that 20 different sound-final frequency levels (10 decreases, 10 increases) were presented for each of the 8 modulation-rate levels.

A visual digit-memory task with three levels of difficulty was utilized as a distraction (Figure 4A). In this task, participants saw a series of digits and then had to say whether or not a subsequently presented digit had been seen in the series. Simultaneously with the 4-s amplitude-modulated sound, 8 digits were sequentially presented on the screen (each digit was displayed for 0.44 s interleaved with a 0.069-s blank screen). In the easy 1-digit condition, a single digit was repeated 8 times (e.g., 3 3 3 3 3 3 3 3). In the moderately difficult 4-digit condition, four different digits were each repeated twice in a row (e.g., 4 4 8 8 2 2 1 1). In the difficult 8-digit condition, all of the presented digits were different (e.g., 3 4 1 2 9 7 5 8). A visual cue of 0.7-s duration prior to each trial indicated whether the memory task on that trial would be ‘Easy’ (in green; 1 digit), ‘Moderate’ (in blue; 4 digits), or ‘Difficult’ (in red; 8 digits). The cue aimed to inform individuals about the difficulty of the distraction task on each trial before the AM sound and the sequence of digits started in order to avoid modulations in attention at times when the second or third digit occurred. After the 4 s of stimulation (amplitude-modulated sound and 8 digits) a probe digit was presented either matching one of the digits in the preceding stream (50 % of trials) or mismatching all of the preceding digits (50 % of trials). Participants first indicated whether they heard the amplitude-modulated sound as “slowing down” or “speeding up” and then indicated whether the probe digit was among the stream of digits. For each of the memory conditions (1, 4, 8 digits) all 8 (rate levels) × 20 (frequency levels) = 160 sounds were presented once each. Stimulation was conducted in 6 blocks separated by breaks and conditions were presented randomly.

Data from the distractor task were analyzed using signal detection theory, extracting d′ as a measure of perceptual sensitivity (Macmillan & Creelman, 2004); a hit was defined as a “yes” response when the probe digit matched one of the digits in the stream of digits. A one-way repeated measures ANOVA with the factor Memory Condition (1, 4, 8 digits) was conducted on d′ values.

Data from the auditory AM-rate change discrimination task were analyzed using the extracted PSE and slope from fitted logistic functions. Separately for PSEs and slopes, a two-way repeated measures ANOVA with the factors Frequency (decrease, increase) and Memory Condition (1, 4, 8 digits) was calculated.

Results and Discussion

Distractor task. The number of digits to be remembered modulated d′ ($F(2, 28) = 75.85$, $p < .001$, $\eta_p^2 = .844$). All d′ were significantly different from each other, with sensitivity best for the 1-digit task, and poorest for the 8-digit task (1 vs. 4: $t(14) = 7.16$, $p < .001$, $r_e = .886$; 1 vs. 8: $t(14) = 10.28$, $p < .001$, $r_e = .940$; 4 vs. 8 $t(14) = 6.72$, $p < .001$, $r_e = .874$; Figure 4B).

Please insert Figure 4 around here

Auditory AM-rate change discrimination task. Figure 4C shows the proportion of “speeding up” responses and the magnitude of the frequency-induced AM-rate change illusion for each memory condition. The repeated measures ANOVA using the PSE as the dependent measure revealed no effect of Memory

Condition ($F(2, 28) = 1.49, p = .243, \eta_p^2 = .010$), but did show a main effect of Frequency ($F(1, 14) = 9.13, p = .009, \eta_p^2 = .395$), caused by more positive PSEs for frequency decreases compared to frequency increases (i.e., indexing the AM-rate change illusion). No Memory Condition × Frequency interaction was found ($F(2, 28) = 0.80, p = .459, \eta_p^2 = .054$). The repeated measures ANOVA for slopes revealed a main effect of Memory Condition ($F(2, 28) = 3.86, p = .033, \eta_p^2 = .216$). Slopes were shallower for 8 digits compared to 1 digit ($t(14) = 2.59, p = .022, r_e = .569$) and 4 digits ($t(14) = 2.54, p = .024, r_e = .562$), consistent with sensitivity declining with increasing memory load. The main effect of Frequency was marginally significant ($F(1,14) = 3.61, p = .078, \eta_p^2 = .205$) due to slightly shallower slopes for frequency decreases compared to frequency increases. No Memory Condition × Frequency interaction was found ($F(2, 28) = 0.57, p = .574, \eta_p^2 = .039$).

To summarize Experiment 4, we hypothesized that high memory load (as compared to low load) would enhance the degree to which the carrier frequency influences perception of AM-rate change. The results of Experiment 4 revealed, contrary to our predictions, no influence of cognitive (i.e., memory) load on the magnitude of the illusory AM-rate change percept. The results displayed in Figure 4B indicate that the distraction paradigm successfully manipulated cognitive load: performance was reduced as the number of digits to be remembered increased. One possibility is that distraction in the visual sensory modality does not affect auditory illusory rate-change percepts (Duncan, Martens, & Ward, 1997).

Experiment 5: Distraction by an Auditory Memory Task

Experiment 5 aimed to investigate whether cognitive distraction in the auditory modality (i.e., the same modality as the AM-rate changes) would affect the illusion magnitude.

Participants

Seventeen university students who were not tested in Experiments 1-4 participated in Experiment 5 (12 female; mean age: 19.5 years; range: 18–25 years). Three additional participants took part in the study but were excluded because performance was around chance level for the AM-rate change task as well as for the distractor task.

Methods

Amplitude-modulated sounds were similar to those in Experiment 4 (8 levels of modulation-rate change and 20 levels of frequency change). However, the amplitude-modulated sounds were presented only to the left or right ear during the main part of the experiment (8 participants received right-ear stimulation; 9 left). The auditory distractor task was presented to the opposite ear.

The auditory distractor was a syllable memory task (Figure 5A) with three difficulty levels, in which participants heard a series of syllables and then had to say whether a subsequently presented probe syllable had been heard in the series. Twelve syllables were recorded by a male speaker (/ba/, /da/, /ga/, /ha/, /ja/, /ka/, /la/, /na/, /pa/, /ra/, /ta/, /wa/). On each trial, simultaneously with the 4-s amplitude-modulated sound, 6 syllables were sequentially presented (approximately equally spaced over the 4 s). On 1-syllable trials, 1 syllable was repeated 6 times. On 3-syllable trials, each of 3 syllables was repeated twice in a row (e.g., /ba/, /ba/, /ra/, /ra/, /ta/, /ta/). On 6-syllable trials, 6 different syllables were presented randomly. Participants received a visual cue of 0.7-s duration prior to each trial indicating whether the memory task would be ‘Easy’ (in green; 1 syllable), ‘Moderate’ (in blue; 3 syllables), or ‘Difficult’ (in red; 6 syllables). After the 4 s of sound stimulation (amplitude-modulated sound and 6 syllables) a probe syllable was auditorily presented: this either matched one or more syllables heard in the preceding series (50 % of trials) or did not match any of the preceding syllables (50 % of trials). Participants first indicated whether they had heard the sound on that trial as “slowing down” or “speeding up” and then indicated whether the probe syllable matched one or more of the concurrently presented series of syllables. For each of the memory conditions (1, 3, 6 syllables) all 8 (rate levels) × 20 (frequency levels) = 160 sounds were presented once each. Stimulation was conducted in 6 blocks separated by breaks and conditions were presented randomly.

As in Experiment 4, performance on the distractor task was indexed with d′, analyzed using a one-way repeated measures ANOVA (Memory Condition: 1, 3, 6 syllables). Data from the auditory AM-rate change discrimination task were analyzed using the PSE and slope values in two-way repeated measures ANOVAs (factors Frequency and Memory Condition).

Results and Discussion

Distractor task. Performance (d′) depended on the number of syllables to be remembered (*F*(2, 32) = 170.13, $p < .001$, $\eta_p^2 = .914$). All d′s differed from each other: performance was best on the 1-syllable task and poorest on the 6-syllable task (1 vs. 3: *t*(16) = 10.96, $p < .001$, $r_e = .939$; 1vs. 6: *t*(16) = 19.63, $p < .001$, $r_e = .980$; 3 vs. 6: *t*(16) = 6.66, $p < .001$, $r_e = .857$; Figure 5B).

Please insert Figure 5 around here

Auditory AM-rate change discrimination task. Figure 5C shows the proportion of "speeding up" responses and the magnitude of the frequency-induced AM-rate change illusion for each memory condition. A repeated measures ANOVA was carried out using the PSE as the dependent measure. No effect of Memory Condition was found (*F*(2, 32) = 1.44, $p = .252$, $\eta_p^2 = .083$), but a main effect of Frequency was observed (*F*(1, 16) = 56.36, $p < .001$, $\eta_p^2 = .779$), caused by a larger PSE for frequency decreases compared to frequency increases (i.e., indexing the AM-rate change illusion). The Memory Condition × Frequency interaction was not significant (*F*(2, 32) = 1.21, $p = .311$, $\eta_p^2 = .070$). The repeated measures ANOVA using the slope as dependent measure showed no effect of Memory Condition (*F*(2, 32) = 0.87, $p = .427$, $\eta_p^2 = .052$), no effect of Frequency (*F*(1, 16) = 1.50, $p = .239$, $\eta_p^2 = .086$), and no interaction (*F*(2, 32) = 0.45, $p = .644$, $\eta_p^2 = .027$).

Similar to Experiment 4, Experiment 5 revealed no effect of cognitive (i.e., memory) load on the magnitude of the AM-rate change illusion (although distraction parametrically modulated memory performance; Figure 5B). That is, neither the visual (Exp. 4) nor the auditory memory task (Exp. 5) led to a modulation in the illusion magnitude.

One possible reason for which we failed to observe modulatory influences of the cognitive distractor tasks in Experiments 4 and 5 might be that participants employed a task-switching strategy. Given the discrete nature of the distractor task – digits/syllables were presented sequentially – participants may have been able to switch their attention back and forth between this digit/syllable stimulus stream and the auditory AM-rate change stimulus, in order to perform both tasks. Indeed, a few participants reported during debriefing that they used such a strategy. To remove the possibility of such attentional switching, we adopted a continuous monitoring task, the multiple-object tracking task, as a distractor task in Experiment 6.

Experiment 6: Distraction by a Multiple-Object Tracking Task

The purpose of Experiment 6 was to investigate whether a distractor task that is demanding throughout a trial affects the influence of carrier frequency on AM-rate change percepts. We chose to utilize a multiple object tracking (MOT; Pylyshyn & Storm, 1988) task that has previously been shown to be attentionally demanding and that requires continuous attentional focus (Alvarez & Franconeri, 2007; Cavanagh & Alvarez, 2005; Scholl, 2009; Tombu & Seiffert, 2008).

Participants

Fifteen university students who did not take part in Experiments 1-5 participated in the experiment (13 female; mean age: 19.1 years; range: 18–23 years). Data from one additional participant were excluded because performance in the distractor task was at chance level even for the easiest condition.

Methods

Auditory stimulation, and analysis of AM-rate change discrimination judgments, were similar to Experiment 4.

The distraction in Experiment 6 was a concurrent MOT task (Figure 6A). Presentation of objects (here dots) was constrained to a display frame of 20.6 cm width (15.6°) and 19.4 cm height (14.7°) centered on the screen and highlighted to the participants by a gray frame on a black background. A yellow fixation square (0.16 cm [0.12°]) was presented at the center of the display frame. Each trial started with a 1-s stationary display of 16 dots (dot diameter: 1.2 cm [0.9°]) of which either 1 or 5 were marked in red. Participants were asked to track the marked dots over time. After 1 s, all dots reverted to white and participants had to track the previously marked dots over the course of 4 s during which the amplitude-modulated sound was presented. The MOT task was presented at three levels of difficulty: Track one stationary dot among 15 additional stationary dots (Stationary1); Track one moving dot among 15 additional moving dots (Moving1); Track five moving dots among 11 additional moving dots (Moving5). Implementation of dot movements in the two moving conditions was adopted from (Wilson, O'Grady, & Rajsic, 2013): dots never moved outside of the display frame and never overlapped during movements; dots moved approximately 3.7 cm/s (2.8°/s). Following the sound presentation, the dot display froze (if it had been moving) and one dot was marked in green. On half of the trials the green dot was one of the ones the participant had been asked to track and on the other half of the trials it was a (randomly selected) dot that had not been marked in red at the beginning of the trial. Participants first responded whether the AM sound had been “slowing down” or “speeding up”, and then indicated whether the green dot was among the ones the participant had been asked to track. For each of the MOT conditions (Stationary1, Moving1, Moving5) all 8 (rate levels) × 20 (frequency levels) = 160 sounds were presented once each. Stimulation was conducted in 6 blocks separated by breaks and conditions were presented randomly.

Data from the distractor task were again analyzed using d′ and a one-way repeated measures ANOVA (MOT Condition: Stationary1, Moving1, Moving5). Data from the auditory AM-rate change task were analyzed using the PSE and slope in two-way repeated measures ANOVAs (factors Frequency and MOT condition).

Results and Discussion

Distractor task. Performance in the visual distractor task was modulated by MOT condition ($F(2, 28) = 47.73$, $p < .001$, $\eta_p^2 = .773$; Figure 6B). All d′s differed from each other: performance was best for the Stationary1 and poorest for the Moving5 condition (Stationary1 vs. Moving1: $t(14) = 3.26$, $p = .006$, $r_e = .657$; Stationary1 vs. Moving5: $t(14) = 9.31$, $p < .001$, $r_e = .928$; Moving1 vs. Moving5: $t(14) = 5.88$, $p < .001$, $r_e = .844$).

Please insert Figure 6 around here

Auditory AM-rate change discrimination task. Figure 6C shows the proportion of “speeding up” responses and the magnitude of the frequency-induced AM-rate change illusion for each MOT condition. For PSEs, the repeated measures ANOVA revealed no effect of MOT condition ($F(2, 28) = 1.82$, $p = .180$, $\eta_p^2 = .115$), but did show a main effect of Frequency ($F(1, 14) = 22.97$, $p < .001$, $\eta_p^2 = .622$) which was due to more positive PSE values for frequency decreases than increases (i.e., indexing the AM-rate change illusion). Critically, the MOT condition × Frequency interaction was significant ($F(2, 28) = 4.89$, $p = .015$, $\eta_p^2 = .259$). That is, the AM-rate change illusion (i.e., PSE difference between frequency decreases and frequency increases) was larger for the Moving5 compared to the Stationary1 condition ($t(14) = 4.27$, $p < .001$, $r_e = .752$; Figure 6C). The difference between Stationary1 and Moving1 ($t(14) = 1.42$, $p = .177$, $r_e = .355$) and the difference between Moving1 and Moving5 ($t(14) = 1.40$, $p = .182$, $r_e = .352$) was not significant. The repeated-measures ANOVA using the slope of the psychometric functions as the dependent measure revealed no effect of MOT condition ($F(2, 28) = 1.33$, $p = .279$, $\eta_p^2 = .087$), no effect of Frequency ($F(1, 14) = 0.04$, $p = .848$, $\eta_p^2 = .003$), and no MOT condition × Frequency interaction ($F(2, 28) = 0.06$, $p = .938$, $\eta_p^2 = .005$).

The results of Experiment 6 demonstrate that increased attentional demands correlate with an increased AM-rate change illusion (i.e., an increased difference between the PSE for frequency decreases and frequency increases, so increased bias) without changing perceptual sensitivity to AM-rate changes (i.e., slopes of the psychometric functions). That is, distraction by a concurrent visual motion tracking task increased the influence of frequency on AM-rate change percepts, but did not affect the ability to discriminate AM-rate changes.

General Discussion

In this series of experiments, we investigated whether prior knowledge of stimulus features, or a concurrent distractor task, alter the influence of a sound’s frequency change on perceived AM-rate change. Experiment 1 demonstrated that AM-rate changes are more likely to be perceived as slowing down when the sound frequency decreases and as speeding up when frequency increases. This illusion is probably the result of extensive experience of co-variation in acoustic rate of change and frequency, common in natural sounds (Broze & Huron, 2013; Topbas, et al., 2012). Reliable cues about the degree to which an upcoming sound would change in modulation rate or frequency reduced the influence of frequency on rate-change perception (Exp. 2 & 3), whereas this perceptual interdependency either did not change (Exp. 4 & 5) or was increased when a concurrent distractor task was performed (Exp. 6). The data suggest that in situations that prohibit regulation of attention (uninformative cues) or in situations that are distracting (MOT) perceptual decisions about sounds might rely strongly on information in stimulus dimensions that are naturally correlated with, but are not identical with, the dimension upon which the perceptual decision is supposed to be based.

Perceptual interdependency of modulation rate and carrier frequency

The discrimination of AM-rate change direction depended not only on the degree to which the modulation rate actually changed, but also on how, the carrier frequency changed, producing an illusory rate-change percept. Our data are in line with previous studies using frequency-modulated sounds (instead of the amplitude-modulated sounds employed here), which found that changes in mean carrier frequency induce illusory changes in perceived frequency-modulation rate (Herrmann, et al., 2013; Herrmann, et al., 2014). The current results are also consistent with other previous observations. For example, musical pieces are perceived as faster when frequency increases (Boltz, 2011), speech is perceived as faster when it is spoken at a higher frequency (Feldstein & Bond, 1981), the interval between two sounds is perceived as shorter when the sounds are played at a higher frequency (Lake, LaBar, & Meck, 2014).

In addition to purely perceptual measures, neural synchronization – that is, the propensity of neural oscillatory activity to synchronize with temporal modulations of sounds over time – is systematically affected by featural co-variations of modulation rate and carrier frequency. For example, magnetoencephalographic recordings from human auditory cortex show that neural synchronization with a sound's frequency modulation increases when simultaneous changes in mean carrier frequency are congruent (compared to incongruent) with the sound's modulation-rate change: Synchronization is highest for sounds that simultaneously increase in modulation rate and in mean carrier frequency, and for sounds that simultaneously decrease in modulation rate and in mean carrier frequency (Herrmann, et al., 2013). These data show that the perceptual interdependency of modulation rate and frequency has a strong neural correlate in auditory cortex.

The correlation between modulation rate and frequency is only one instance of featural co-variations influencing perception among many others reported in the literature. For example, numerical magnitudes, sound intensity, sound frequency, visual disk size, visual color saturation, and visual space all influence judgments of time-interval duration (Alards-Tomalin, Leboe-McGowan, Shaw, & Leboe-McGowan, 2014; Alards-Tomalin, et al., 2013; Henry & McAuley, 2009; Huang & Jones, 1982; Jones & Huang, 1982; Shigeno, 1986). Furthermore, pitch perception is influenced by manipulation of temporal intervals (Henry, et al., 2009; Shigeno, 1986, 1993) and visual color (Melara, 1989), and fundamental frequency (heard as pitch) in turn influences perception of loudness (Neuhoff, et al., 1999), timbre (Melara & Marks, 1990), and visual color (Melara, 1989). Perception in natural environments thus appears to be based on evidence arising from many different correlated features.

There is a general consensus that perceptual interdependencies arise from exposure to featural co-variations in natural environments (Alards-Tomalin, et al., 2014; Boltz, 1998, 2011; Neuhoff, 2004; Neuhoff, et al., 1999; Walsh, 2003). Auditory perceptual interdependencies, such as the relation between modulation rate and frequency that was studied here, might be the result of exposure to featural co-variations in speech, music, and vocal emotional expressions as we have reviewed in the introduction. Perception of a stimulus feature is biased by the co-occurrence of another, for the perceptual decision irrelevant feature in a manner that is consistent with the learned featural co-variations. Critically, as we discuss in the following, reliance on featural co-variations in sounds for perceptual decisions appears to depend on cognitive factors and situational demands.

The influence of frequency on AM-rate change discrimination increases when the listener is not optimally attentive

Providing individuals with informative cues about the degree to which a subsequent sound would change in modulation rate (Experiment 2) or carrier frequency (Experiment 3) reduced the magnitude of the frequency-induced bias in AM-rate change perception. The informative cue reliably predicted the degree of change in the primary (to-be-judged) dimension (Experiment 2) or in the secondary (to-be-ignored) dimension (Experiment 3). This manipulation was intended to modulate attention by controlling the relevance of the signal dimensions for perceptual decisions (Summerfield & Egner, 2016), such that individuals could actively regulate the degree to which they attentively listened to the sounds. Informative cues appeared to have enabled participants to more successfully suppress changes in carrier frequency (compared to uninformative cues), because no frequency-induced modulation in AM-rate change sensitivity (i.e., in slopes) accompanied the reduction in biased AM-rate change percepts.

In order to examine more directly the influence of cognitive factors on the perceptual interdependency of AM-rate change and carrier frequency, a memory distractor task was utilized in Experiments 4 and 5 that aimed to modulate cognitive load during presentation of the amplitude-modulated sounds. A higher memory load was expected to increase the influence of carrier frequency on AM-rate change perception. Consistent with previous studies, increasing the number of items that had to be hold in memory led to a reduction in performance (Hester & Garavon, 2005; Intaitė, Koivisto, & Castelo-Branco, 2014; Sternberg, 1966; Zäske, Perlich, & Schweinberger, 2016). However, manipulation of item number had no effect on the frequency-induced bias in AM-rate change perception. We can only speculate why no such modulation was observed. Previous studies often employed stimulation during the retention phase of a memory distractor task instead of presenting the items that had to be hold in memory concurrently with the stimulation, as was done here (Dalton, et al., 2009; de Fockert, Rees, Frith, & Lavie, 2001; Lavie & de Fockert, 2005; but see Macken, Tremblay, Houghton, Nicolls, & Jones, 2003). Furthermore, we manipulated memory load by increasing the number of items to be remembered and asked participants whether a probe item had been presented in the preceding stream of items. Other studies have instead manipulated memory load by changing the order of items (e.g., digits) and probed participants on the item order (Dalton, et al., 2009; de Fockert, et al., 2001; Lavie & de Fockert, 2005). A third possibility for the absence of modulatory influences of memory load on biased AM-rate change perception might be related to the sequential nature of the memory task (digits/syllables were presented sequentially), which might have allowed participants to switch between the memory task and the auditory AM-rate change discrimination task.

In order to eliminate the possibility of task switching, we utilized a multiple-object tracking (MOT) task that is attentionally demanding and requires continuous attentional focus over time (Alvarez & Franconeri, 2007; Masutomi, et al., 2016; Tombu & Seiffert, 2008). Although MOT is not commonly used as a distractor task, it might be useful when distraction from attending to auditory stimuli over multiple seconds is essential (Masutomi, et al., 2016). Indeed, we observed that MOT distraction increased the influence of carrier frequency on AM-rate change percepts. Individuals appeared to be less able to ignore carrier-frequency changes for perceptual decisions about the direction of AM-rate change in sounds when they tracked multiple moving dots compared to when they tracked one dot on a stationary display of dots. Our tentative conclusion from this is that, under high load, individuals relied more on the known co-variation of modulation rate and frequency in natural sounds and weighted frequency information more heavily in their perceptual decision.

The MOT task relies on mechanisms related to attentional capacity, likely increasing the number of attentional foci under high load (Cavanagh & Alvarez, 2005). Our data are thus in line with work reporting cross-modal attentional modulations (Helbig & Ernst, 2008; Macdonald & Lavie, 2011; Molloy, et al., 2015; Raveh & Lavie, 2015), although attentional resources might not always be shared between the auditory and the visual modality (Duncan, et al., 1997; Keitel, Maess, Schröger, & Müller, 2013; Rees, Frith, & Lavie, 2001). Taking the results from Experiments 2, 3 and 6 together, the current study indicates that when regulation of attention to sounds is prohibited or listening situations are distracting human auditory perception might be strongly shaped by learned featural co-variations in natural sounds.

Changes in perceptual bias versus sensitivity

In the current study, the effects of predictability and MOT distraction only affected the degree to which AM-rate change perception was biased by changes in frequency, but did not affect perceptual sensitivity to AM-rate changes: The interaction of Frequency × Cue Type (Experiments 2 & 3) and Frequency × MOT condition (Experiment 6) was significant only for the PSEs, but not for the slopes of the psychometric functions.

Previous studies using cueing or dual-task distraction paradigms have mostly used reaction times or the proportion of correct responses as dependent measures; perceptual sensitivity and bias cannot be distinguished using such measures. Although signal-detection methods allow the distinction between perceptual sensitivity (e.g., slope or d′) and bias (e.g., PSE or *c*) (Macmillan & Creelman, 2004), assessment of whether modulations of bias reflect effects of perception may depend on the employed task (Witt, Taylor, Sugovic, & Wixted, 2015). Theoretical considerations as well as simulations suggest that perceptual effects in Yes/No tasks – that is, when the presence or absence of a stimulus is judged – are captured by d′ and not by bias (Witt, et al., 2015). In discrimination tasks, such as the one employed here, perceptual effects might either be captured by measures of perceptual sensitivity or by measures of bias (Witt, et al., 2015). In the current study, we used the slope of the psychometric function (i.e., perceptual sensitivity) as a measure of discrimination difficulty in the AM-rate change task. The PSE (i.e., bias) of the psychometric function was used to investigate the influence of carrier frequency on perceptual inferences about AM-rate changes. In particular for featural co-variations, effects of perception are commonly captured by measures of bias (Alards-Tomalin, et al., 2013; Henry & McAuley, 2009, 2013).

Studies that have investigated the effects of distraction on perception mainly used Yes/No tasks and static (i.e., non-changing) stimuli, and have reported a decrease in perceptual sensitivity without significant shifts in criterion (bias) for visual and auditory stimuli when perceptual load in the distractor task is increased (Konstantinou & Lavie, 2013; Lavie, et al., 2014; Macdonald & Lavie, 2008, 2011; Molloy, et al., 2015). We utilized a discrimination task and dynamic (i.e., changing) stimuli, and did not observe any modulations of perceptual sensitivity induced by informative cues or load in a MOT distractor task. The difficulty of the current AM-rate change task thus appeared unaffected by manipulations of attention. Instead, our results show that presentation of cues (enabling participants to regulate the level of attention) as well as distraction only modulate the degree to which the task-irrelevant and potentially distracting stimulus feature (frequency change) influences perceptual inferences about the task-relevant feature (modulation-rate change). Consistent with our observations, a few studies using cueing paradigms report attention-related modulations of bias (de Lange, et al., 2013; Rahnev, et al., 2011; Schneider, 2011; Schneider & Komlos, 2008).

Work by Mattys and colleagues (Mattys & Scharenborg, 2014; Mattys & Wiget, 2011) is consistent with the idea that distraction might increase the degree to which listeners rely on prior knowledge (experience) for perceptual decisions. These authors observe that the tendency to perceive an ambiguous phoneme as belonging to a real word (e.g., hearing /g#/, where # is an ambiguous sound intermediate between /iss/ and /ift/, as ‘gift’ and not ‘giss’) increases under perceptual distraction (effectively a shift in PSE). In these studies, however, visual inspection of the figures showing the load effect (on the lexicality-induced phoneme identification bias) also suggests a decrease in phoneme identification *sensitivity* under high load (shallower slopes; for a more detailed test see Mattys, Seymour, Attwood, & Munafò, 2013); this was also shown in a separate discrimination task. The influences of a task-irrelevant stimulus dimension on perceptual inferences usually increases when the task-relevant dimension is ambiguous (based on that dimension alone) and sensitivity is reduced (Ganong, 1980; Garner, 1976). Hence, if there were a reduction in sensitivity to the primary dimension under high-load conditions, this by itself would potentially increase the interference by the secondary (to-be-ignored) dimension, and thus the bias. Here we observed only a modulation of the PSE (bias), but not of the slope (sensitivity), and our data thus suggest that for nonlinguistic auditory dimensions reliance on featural co-variations for perceptual decisions indeed increases in distracting listening situations.

Potential mechanisms of attentional modulations

The load theory of attention suggests that perceptual sensitivity to the task-irrelevant stimulus is reduced under high compared to low perceptual load (Lavie, 2005; Lavie, et al., 2014). The current results are somewhat inconsistent with the load theory framework as we observed a larger influence of the irrelevant stimulus dimension on perception of the primary stimulus dimension under high load. However, the current study differs substantially from previous work in support of the load theory (Macdonald & Lavie, 2008, 2011; Molloy, et al., 2015; Raveh & Lavie, 2015). For example, in the current study, the task-irrelevant aspect of the experimental design was a stimulus dimension presented at a supra-threshold level rather than an isolated near-threshold stimulus as has been used in previous studies. The current works thus does not challenge load theory, but demonstrates that under specific conditions, an irrelevant aspect of the environment might paradoxically exert a stronger influence on perceptual inferences about the task-relevant stimulus dimension under high load.

Our data are more consistent with the feature-integration theory of attention (Treisman & Gelade, 1980), which suggests that individual stimulus features are processed early and in parallel, and that those features become automatically integrated – that is, without selective attention – when they are perceptually interdependent (or integral; Garner, 1976) such as modulation rate and frequency appear to be. The current data suggest that focused attention counteracts automatic integration and enables (partial) separability of acoustic features. That is, when a listener is not optimally attentive, integration of the two stimulus dimensions increases, but specifically such that perceptual inferences are biased in a way that is consistent with the learned featural co-variations in natural sounds.

The neural mechanisms that support modulation of perceptual weighting of features by attentional state are unknown. Attention to a stimulus feature might lead to sharpening of neural tuning, increases in response gain, or threshold shifts (Reynolds & Heeger, 2009; Treue, 2001). However, many studies that have investigated attention-related neural response modulations have focused on perceptual sensitivity in detection or discrimination tasks in the absence of featural co-variations (e.g., Reynolds & Desimone, 2003; Reynolds, Pasternak, & Desimone, 2000; Treue & Martínez Trujillo, 1999; Williford & Maunsell, 2006). We have observed that attentional state reweights the correlated dimension in a way that influences bias, but not perceptual sensitivity. Whether changes in tuning, response gain, and/or threshold shifts underlie attention-related modulations of perceptual interdependencies is unknown. Perhaps surprisingly, our data indicate that attention suppresses the influence of a correlated, task-irrelevant feature, rather than modulating sensitivity to a task-relevant feature.

Conclusions

The current study investigated whether co-variations of acoustic features in sounds are utilized in a stable manner or whether their role in perception depends on the degree to which the listener is attending to the sounds. We made use of the fact that temporal and spectral features are correlated in natural sounds (Broze & Huron, 2013; Topbas, et al., 2012) and modeled this in artificial narrowband sounds with imposed amplitude modulation that increased or decreased in rate over time, and with concurrent increases or decreases in frequency. Listeners were asked to judge the direction of AM-rate change, and were biased in their decisions by the direction of a concurrent, but task-irrelevant frequency change – they appeared to be relying on frequency information to make judgments about AM-rate change direction. Critically, the influence of the correlated (but irrelevant) dimension decreased when listeners were cued beforehand to the magnitude of change in either feature dimension (so they could regulate their level of attention), and increased when a concurrent, distracting, multiple-object tracking task (Alvarez & Franconeri, 2007; Pylyshyn & Storm, 1988) had to be simultaneously performed (thereby sharply reducing their ability to attend to the sounds). These data suggest that when the listener is not optimally attentive, he or she might rely strongly on learned featural co-variations for sound perception.

References

Alain, C. (1993). The relation among fundamental frequency, intensity, and duration varies with accentuation. *The Journal of the Acoustical Society of America, 94*, 2434-2436.

Alards-Tomalin, D., Leboe-McGowan, J. P., Shaw, J. D., & Leboe-McGowan, L. C. (2014). The Effects of Numerical Magnitude, Size, and Color Saturation on Perceived Interval Duration. *Journal of Experimental Psychology: Learning, Memory \& Cognition, 40*, 555-566.

Alards-Tomalin, D., Leboe-McGowan, L. C., & Mondor, T. A. (2013). Examining auditory kappa effects through manipulating intensity differences between sequential tones. *Psychological Research, 77*, 480-491.

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision, 7*, 1-10.

Black, J. W. (1961). Relationships among fundamental frequency, vocal sound pressure, and rate of speaking. *Language and Speech, 4*, 196-199.

Boltz, M. G. (1998). Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure. *Perception & Psychophysics, 60*, 1357-1373.

Boltz, M. G. (2011). Illusory tempo changes due to musical characteristics. *Music Perception, 28*, 367-386.

Bond, R. N., & Feldstein, S. (1981). Acoustical Correlates of the Perception of Speech Rate: An Experimental Investigation. *Journal of Psycholinguistic Research, 11*, 539-557.

Broze, Y., & Huron, D. (2013). Is higher music faster? Pitch-speed relationships in Western compositions. *Music Perception, 31*, 19-31.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*, 349-354.

Crowder, R. G., & Neath, I. (1994). The influence of pitch on time perception in short melodies. *Music Perception, 12*, 379-386.

Dalton, P., Santangelo, V., & Spence, C. (2009). The role of working memory in auditory selective attention. *The Quarterly Journal of Experimental Psychology, 62*, 2126-2132.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The Role of Working Memory in Visual Selective Attention. *Science, 291*, 1803-1806.

de Lange, F. P., Rahnev, D. A., Donner, T. H., & Lau, H. (2013). Prestimulus Oscillatory Activity over Motor Cortex Reflects Perceptual Expectations. *The Journal of Neuroscience, 33*, 1400-1410.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387*, 808-810.

Elliott, T. M., & Theunissen, F. E. (2009). The Modulation Transfer Function for Speech Intelligibility. *PLoS Computational Biology, 5*, e1000302.

Feldstein, S., & Bond, R. N. (1981). Perception of speech rate as a function of vocal intensity and frequency. *Language and Speech, 24*, 387-394.

Ganong, W. F. (1980). Phonetic Categorization in Auditory Word Perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 110-125.

Garner, W. R. (1976). Interaction of Stimulus Dimensions in Concept and Choice Processes. *Cognitive Psychology, 8*, 98-123.

Green, J. A., Whitney, P. G., & Potegal, M. (2011). Screaming, Yelling, Whining and Crying: Categorical and intensity differences in Vocal Expressions of Anger and Sadness in Children's Tantrums. *Emotion, 11*, 1124-1133.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech—a syllable-centric perspective. *Journal of Phonetics, 31*, 465-485.

Helbig, H. B., & Ernst, M. O. (2008). Visual-haptic cue weighting is independent of modality-specific attention. *Journal of Vison, 8*, 21.

Henry, M. J., & McAuley, J. D. (2009). Evaluation of an Imputed Pitch Velocity Model of the Auditory Kappa Effect. *Journal of Experimental Psychology: General, 35*, 551-564.

Henry, M. J., & McAuley, J. D. (2013). Perceptual Distortions in Pitch and Time Reveal Active Prediction and Support for an Auditory Pitch-Motion Hypothesis. *PLoS ONE, 8*, e70646.

Henry, M. J., McAuley, J. D., & Zaleha, M. (2009). Evaluation of an imputed pitch velocity model of the auditory tau effect. *Attention, Perception, & Psychophysics, 71*, 1399-1413.

Herrmann, B., Henry, M. J., Grigutsch, M., & Obleser, J. (2013). Oscillatory Phase Dynamics in Neural Entrainment Underpin Illusory Percepts of Time. *The Journal of Neuroscience, 33*, 15799-15809.

Herrmann, B., Henry, M. J., Scharinger, M., & Obleser, J. (2014). Supplementary motor area activations predict individual differences in temporal-change sensitivity and its illusory distortions. *NeuroImage, 101*, 370-379.

Hester, R., & Garavon, H. (2005). Working memory and executive function: The influence of content and load on the control of attention. *Memory & Cognition, 33*, 221-233.

Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews, 6*, 121-143.

Huang, Y. L., & Jones, B. (1982). On the interdependence of temporal and spatial judgments. *Perception & Psychophysics, 32*, 7-14.

Intaitė, M., Koivisto, M., & Castelo-Branco, M. (2014). The linear impact of concurrent working memory load on dynamics of Necker cube perceptual reversals. *Journal of Vision, 14*, 13.

Jones, B., & Huang, Y. L. (1982). Space-Time Dependencies in Psychophysical Judgment of Extent and Duration: Algebraic Models of the Tau and Kappa Effects. *Psychological Bulletin, 91*, 128-142.

Keitel, C., Maess, B., Schröger, E., & Müller, M. M. (2013). Early visual and auditory processing rely on modality-specific attentional resources. *NeuroImage, 70*, 240-249.

Kluender, K. R., Stilp, C. E., & Kiefte, M. (2013). Perception of Vowel Sounds Within a Biologically Realistic Model of Efficient Coding. In G. S. Morrison & P. F. Assmann (Eds.), *Vowel Inherent*

Spectral Change (pp. 117-151). Berlin, Heidelberg: Springer.

Konstantinou, N., & Lavie, N. (2013). Dissociable Roles of Different Types of Working Memory Load in Visual Detection. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 919-924.

Lake, J. I., LaBar, K. S., & Meck, W. H. (2014). Hear it playing low and slow: How pitch level differentially influences time perception. *Acta Psychologica, 149*, 169-177.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences, 9*, 75-82.

Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: attention, awareness and the role of perceptual load. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*, 20130205.

Lavie, N., & de Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review, 12*, 669-674.

Leek, M. R. (2011). Adaptive procedures in psychophysical research. *Perception & Psychophysics, 63*, 1279-1292.

Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nature, 5*, 356-363.

Macdonald, J. S. P., & Lavie, N. (2008). Load Induced Blindness. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1078-1091.

Macdonald, J. S. P., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics, 73*, 1780-1789.

Macken, W. J., Tremblay, S., Houghton, R. J., Nicolls, A. P., & Jones, D. M. (2003). Does Auditory Streaming Require Attention? Evidence From Attentional Selectivity in Short-Term Memory. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 43-51.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide*. USA: Psychology Press.

Masutomi, K., Barascud, N., Kashino, M., McDermott, J. H., & Chait, M. (2016). Sound Segregation via Embedded Repetition Is Robust to Inattention. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 386-400.

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology, 59*, 203-243.

Mattys, S. L., & Scharenborg, O. (2014). Phoneme Categorization and Discrimination in Younger and Older Adults: A Comparative Analysis of Perceptual, Lexical, and Attentional Factors. *Psychology and Aging, 29*, 150-162.

Mattys, S. L., Seymour, F., Attwood, A. S., & Munafò, M. R. (2013). Effects of Acute Anxiety Induction on Speech Perception: Are Anxious Listeners Distracted Listeners? *Psychological Science, 24*, 1606-1608.

Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language, 65*, 145-160.

Melara, R. D. (1989). Dimensional Interaction Between Color and Pitch. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 69-79.

Melara, R. D., & Marks, L. E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics, 48*, 169-178.

Molloy, K., Griffiths, T. D., Chait, M., & Lavie, N. (2015). Inattentional Deafness: Visual Load Leads to Time-Specific Suppression of Auditory Evoked Responses. *The Journal of Neuroscience, 35*, 16046-16054.

Neuhoff, J. G. (2004). Interacting Perceptual Dimensions. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 249-269). New York: Elsevier Academic Press.

Neuhoff, J. G., McBeath, M. K., & Wanzie, W. C. (1999). Dynamic Frequency Change Influences Loudness Perception: A Central, Analytic Process. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1050-1059.

Pashler, H. (1994). Dual-Task Interference in Simple Tasks: Data and Theory. *Psychological Bulletin, 116*, 220-244.

Pfeuty, M., & Peretz, I. (2010). Abnormal pitch–time interference in congenital amusia: Evidence from an implicit test. *Attention, Perception, & Psychophysics, 72*, 763-774.

Pisoni, D. B., & Luce, P. A. (1987). Trading Relations, Acoustic Cue Integration, and Context Effects in Speech Perception. In M. E. Schouten (Ed.), *The Psychophysics of Speech Perception* (pp. 155-172). Netherlands: Springer.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 179-197.

Rahnev, D., Maniscalco, B., Graves, T., Huang, E., de Lange, F. P., & Lau, H. (2011). Attention induces conservative subjective biases in visual perception. *Nature Neuroscience, 14*, 1513-1515.

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics, 77*, 483-492.

Rees, G., Frith, C., & Lavie, N. (2001). Processing of irrelevant visual motion during performance of an auditory attention task. *Neuropsychologia, 39*, 937-949.

Reynolds, J. H., & Desimone, R. (2003). Interacting Roles of Attention and Visual Salience in V4. *Neuron, 37*, 853-863.

Reynolds, J. H., & Heeger, D. J. (2009). The Normalization Model of Attention. *Neuron, 61*, 168-185.

Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention Increases Sensitivity of V4 Neurons. *Neuron, 26*, 703-714.

Rosenthal, R., & Rubin, D. B. (2003). requivalent: A Simple Effect Size Indicator. *Psychological Methods, 8*, 492-496.

Schneider, K. A. (2011). Attention alters decision criteria but not appearance: A reanalysis of Anton-Erxleben, Abrams, and Carrasco (2010). *Journal of Vision, 11*, 7.

Schneider, K. A., & Komlos, M. (2008). Attention biases decisions but does not alter appearance. *Journal of Vision, 10*, 3.

Scholl, B. J. (2009). What have we learned about attention from multiple object tracking (and vice versa)? In D. Dedrick & L. Trick (Eds.), *Computation, cognition, and Pylyshyn* (pp. 49-78). Cambridge, MA, USA: MIT Press.

Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences, 26*, 100-107.

Shigeno, S. (1986). The auditory tau and kappa effects for speech and nonspeech stimuli. *Perception & Psychophysics, 40*, 9-19.

Shigeno, S. (1993). The interdependence of pitch and temporal judgments by absolute pitch possessors. *Perception & Psychophysics, 54*, 682-692.

Smith, E. C., & Lewicki, M. S. (2006). Efficient auditory coding. *Nature, 439*, 978-982.

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science, 153*, 652-654.

Tombu, M., & Seiffert, A. E. (2008). Attentional costs in multiple-object tracking. *Cognition, 108*, 1-25.

Topbas, O., Orlikoff, R. F., & St. Louis, K. O. (2012). The effect of syllable repetition rate on vocal characteristics. *Journal of Communication Disorders, 45*, 173-180.

Treisman, A. M., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology, 12*, 97-136.

Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences, 24*, 295-300.

Treue, S., & Martínez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399*, 575-579.

Tyler, L. K., & Marslen-Wilson, W. D. (2008). Fronto-temporal brain systems supporting spoken language comprehension. *Philosophical Transactions of the Royal Society B, 363*, 1037-1054.

Vallesi, A., Arbula, S., & Bernardis, P. (2014). Functional dissociations in temporal preparation: Evidence from dual-task performance. *Cognition, 130*, 141-151.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*, 483-488.

Williford, T., & Maunsell, J. H. R. (2006). Effects of Spatial Attention on Contrast Response Functions in Macaque Area V4. *Journal of Neurophysiology, 96*, 40-54.

Wilson, D., O'Grady, M., & Rajsic, J. (2013). Multiple Object Tracking: Support for Hemispheric Independence. *Journal of Vision, 13*, 1295-1295.

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception, 44*, 289-300.

Zäske, R., Perlich, M.-C., & Schweinberger, S. R. (2016). To hear or not to hear: Voice processing under visual load. *Attention, Perception, & Psychophysics, 78*, 1488-1495.

Figure Captions

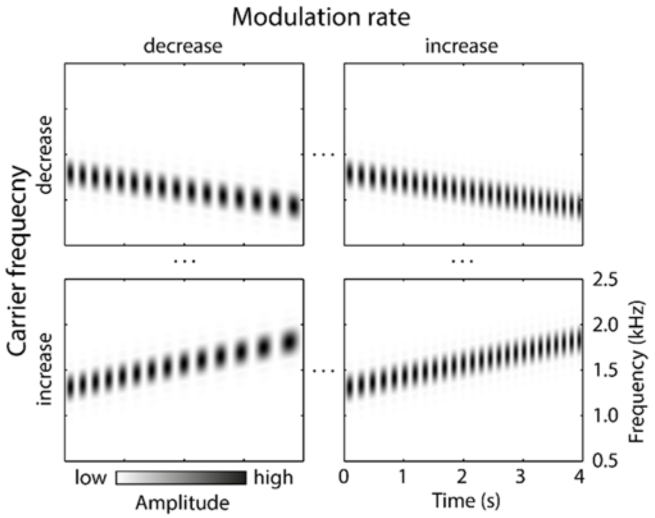

Figure 1: Frequency spectra of sample auditory stimuli. Stimuli were amplitude-modulated sounds changing in modulation rate and carrier frequency over time. In the spectrograms, darker colors represent greater stimulus energy. For visualization, modulation-rate changes are slightly exaggerated here compared to those used in the experiments.

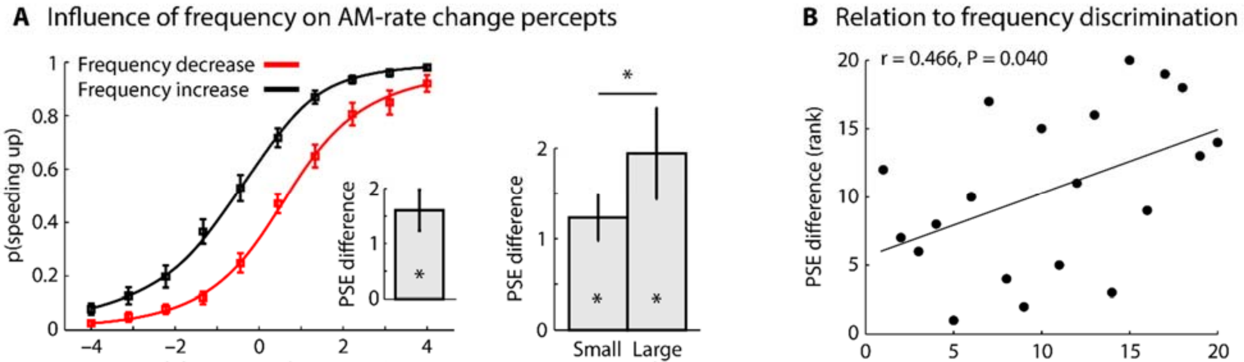

Figure 2: Results for Experiment 1. A) Behavioral performance for the auditory AM-rate change judgment task. Psychometric functions are displayed separately for frequency decreases and frequency increases. The difference in PSE for stimuli with decreasing and increasing frequency is shown in the single bar graph (i.e., indexing the AM-rate change illusion). On the right, the illusion magnitude is displayed for small versus large frequency changes. Asterisks within the bars indicate a significant difference from zero. Error bars reflect the standard error of the mean (SEM). *$p < .05$. **B)** Spearman correlation between frequency-discrimination JND and PSE difference (illusion magnitude). Data on the x- and y-axis reflect the rank-transformed original data.

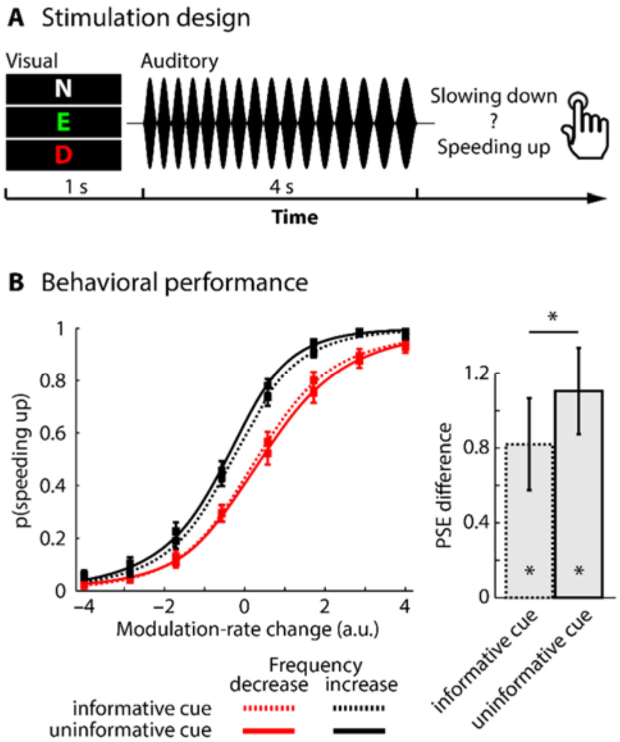

Figure 3: Experimental design and results for Experiments 2 and 3. A) Participants were presented with an amplitude-modulated sound changing in modulation rate and carrier frequency. Each sound presentation was preceded by either an informative visual cue (‘E’ or ‘D’) or an uninformative visual cue (‘N’). The informative cue ‘E’ meant that the subsequent sound would be easy, while the informative cue ‘D’ meant that subsequent sound would be difficult. **B)** Behavioral performance for the auditory AM-rate change discrimination task. Psychometric functions are displayed separately for frequency decreases and frequency increases, and separately for informative and uninformative cues. The difference in PSE between frequency decrease and frequency increase (i.e., indexing the magnitude of the AM-rate change illusion) is shown in the bar graph, separately for informative and uninformative cues. Asterisks within the bars indicate a significant difference from zero. Error bars reflect the SEM. *$p < .05$.

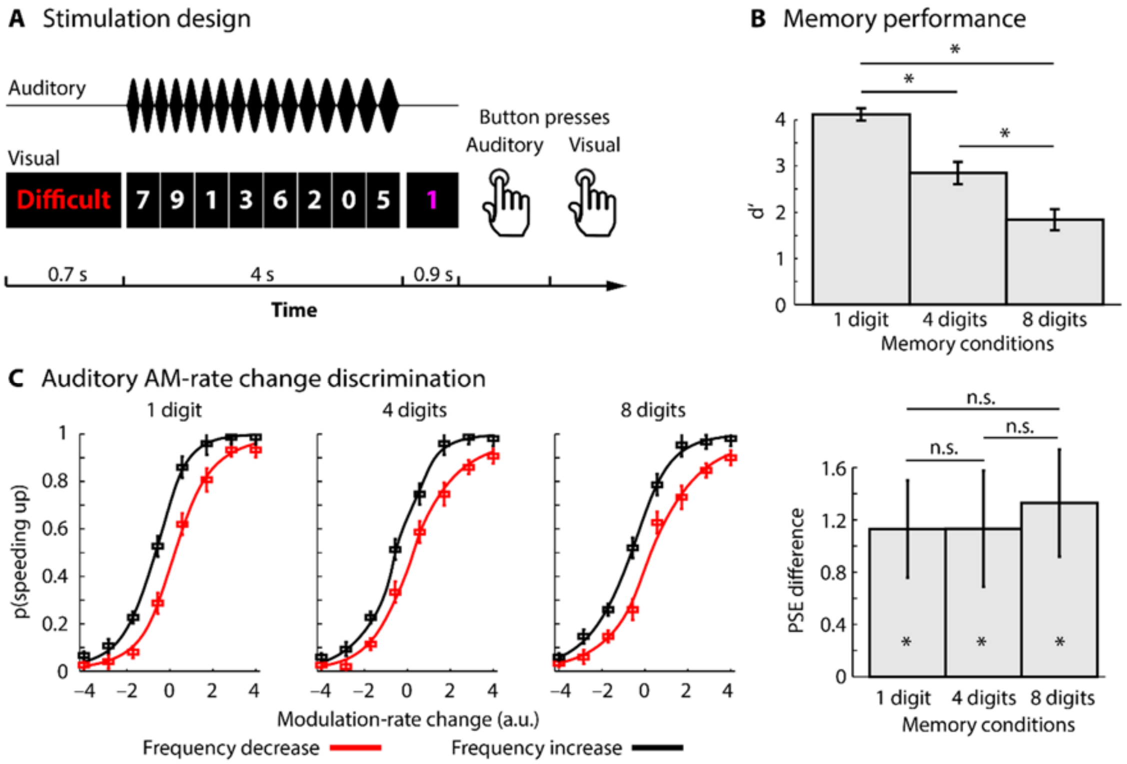

Figure 4: Experimental design and results for Experiment 4. A) Experimental design. A visually presented cue indicated whether the visual distractior memory task would be ‘Easy’, ‘Moderate’, or ‘Difficult’ on that trial (an example for a ‘Difficult’ trial is shown). Subsequently, an amplitude-modulated sound was presented simultaneously with 8 sequentially presented digits. Participants had to perform a dual task, judging whether the amplitude-modulated sound was “slowing down” or “speeding up” and holding the digits in memory in order to subsequently indicate whether a single probe digit had been presented in the stream of digits. **B)** Behavioral performance (d′) in the visual digit memory task at three levels of difficulty: one, four, or eight different digits. **C)** Behavioral performance for the auditory AM-rate change discrimination task. On the left, for each memory condition, psychometric functions are displayed separately for frequency decreases and frequency increases. The difference in PSE between frequency decrease and frequency increase is shown on the right and indexes the magnitude of the AM-rate change illusion. Asterisks within the bars indicate a significant difference from zero. Error bars reflect the SEM. **p* < .05, n.s. – not significant.

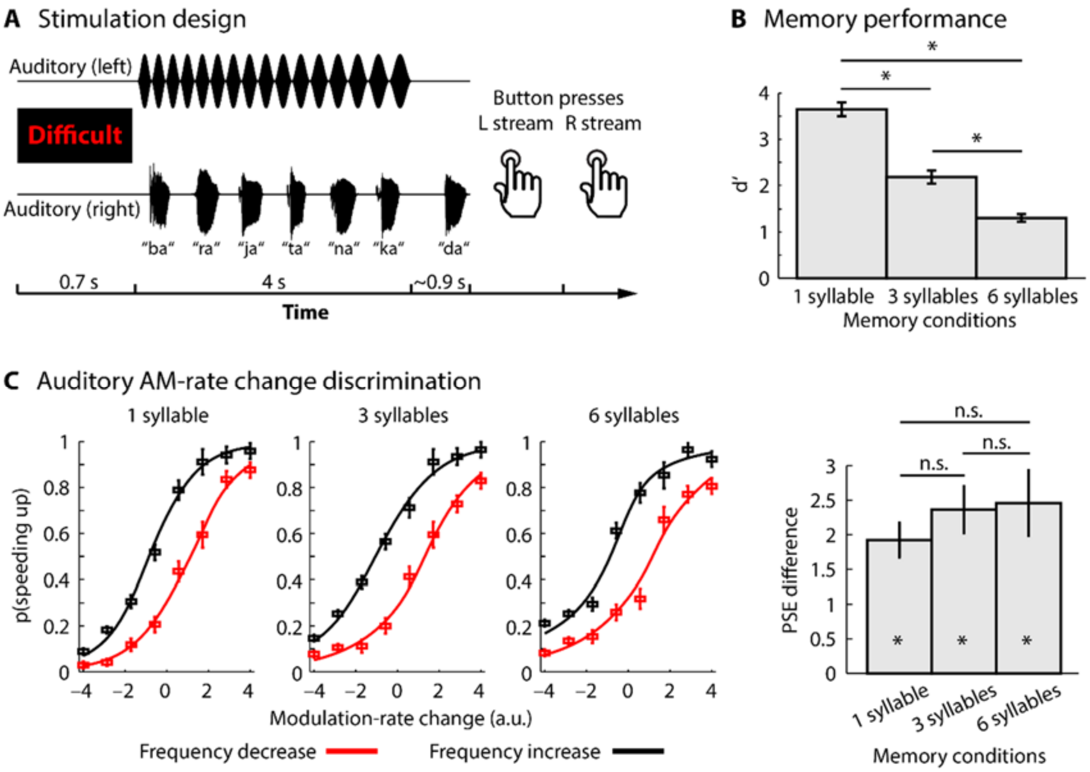

Figure 5: Experimental design and results for Experiment 5. A) Experimental design. A visually presented cue indicated whether the auditory distractor memory task would be ‘Easy’, ‘Moderate’, or ‘Difficult’ on that trial (an example for a ‘Difficult’ trial is shown). Subsequently, an amplitude-modulated sound was presented to one ear and 6 syllables to the other ear. Participants had to perform a dual task, judging whether the amplitude-modulated sound was “slowing down” or “speeding up” and holding the syllables in memory in order to indicate whether a single probe syllable had been presented in the stream of syllables. **B)** Behavioral performance (d′) in the auditory syllable memory task at three levels of difficulty: one, three, or six different syllables. **C)** Behavioral performance for the auditory AM-rate change discrimination task. On the left, for each memory condition, psychometric functions are displayed separately for frequency decreases and increases. The difference in PSE between frequency decrease and increase is shown on the right and indexes the magnitude of the AM-rate change illusion. Asterisks within the bars indicate a significant difference from zero. Error bars reflect the SEM. *$p < .05$, n.s. – not significant.

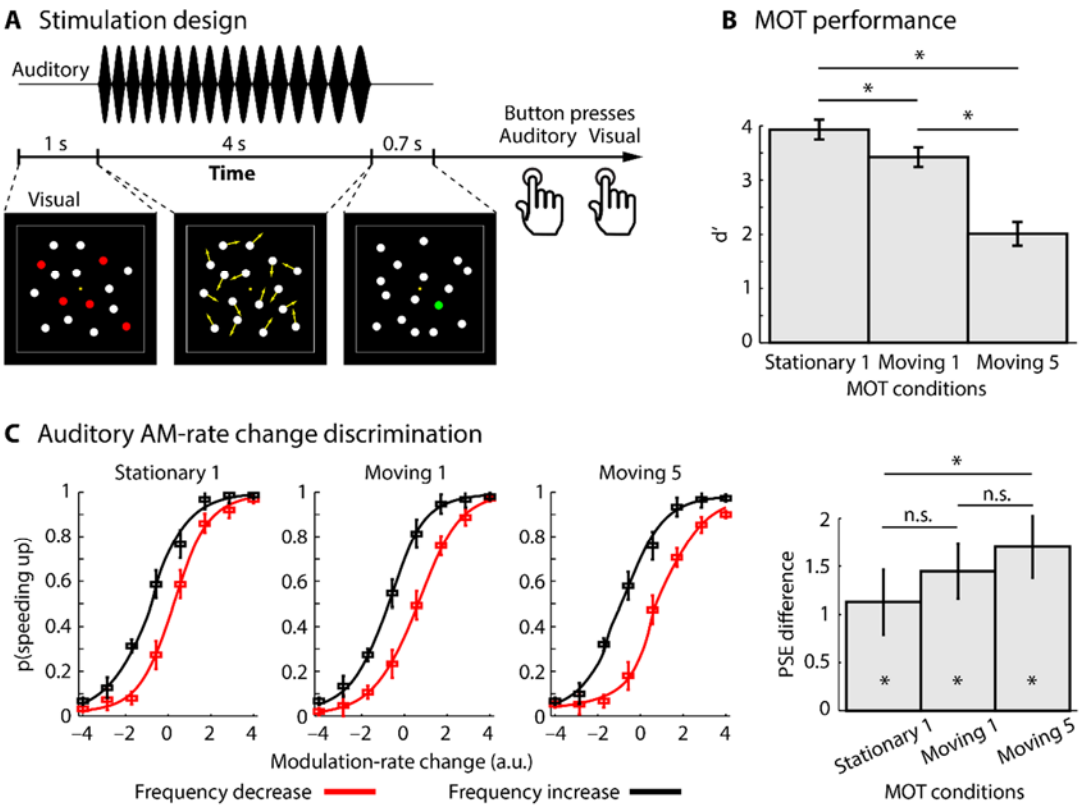

Figure 6: Experimental design and results for Experiment 6. A) Experimental design: Participants saw 16 dots on the screen of which either 1 or 5 were marked in red. Participants were asked to track the marked dots over time while the amplitude-modulated sound was presented. Three MOT conditions were presented: Track one stationary dot among 15 additional stationary dots (Stationary1); Track one moving dot among 15 additional moving dots (Moving1); Track five moving dots among 11 additional moving dots (Moving5). Following sound presentation, one dot was marked in green. Subsequently, participants indicated whether the AM sound had been "slowing down" or "speeding up", and then whether the green dot was among the ones the participant was asked to track. **B)** Behavioral performance (d′) in the MOT task for three different conditions. **C)** Behavioral performance for the auditory AM-rate change discrimination task. On the left, for each MOT condition, psychometric functions are displayed separately for frequency decreases and frequency increases. The difference in PSE between frequency decrease and frequency increase is shown on the right and indexes the magnitude of the AM-rate change illusion. Asterisks within the bars indicate a significant difference from zero. Error bars reflect the SEM. *$p < .05$, n.s. – not significant.