School of Psychology and Clinical Language Sciences

University of Reading

Monaural and dichotic effects of real-room reverberation on loudness perception

Andrew P Raimond | Anthony J Watkins

Background

Stecker & Hafter (2000) found that short sinusoids shaped with fast-onset and slow-offset temporal envelopes, "FS sounds", are judged to be less loud than their time-reversed counterparts , "SF sounds". The excitation-pattern model of Glasberg & Moore (2002) accounts for this loudness difference. However, the loudness difference is more apparent when the sounds to be judged are each preceded by FS sounds than when they are preceded by SF sounds. This 'context effect' might be due to the resemblance of these sounds' offsets to the decaying 'tails' caused by the reflected sound in everyday listening environments (Stecker & Hafter, 2000). The idea is that when successive sounds have similarly long tails, a 'perceptual constancy' results in sounds with slow offsets being parsed into two parts, which separately give information about the sound source and the listening environment. Energy in a decaying tail might thereby be discounted from loudness assessments if listeners judge only the source.

Previously, we have found a similar context effect with *dichotic* sounds that have tails from real-room reflections. These tails are relatively long in duration, they decay non-monotonically, and they are de-correlated at the two ears. This result confirms the idea that a perceptual constancy for room listening-conditions is likely to be responsible for the loudness context effect.

Certain room-reflection 'tail effects' that have been found in speech perception experiments are increased in monaural conditions (Watkins, 2005). Here we ask whether the loudness context effect is also increased when sounds are presented in *monaural* real-room reverberation.

FS and SF pairings

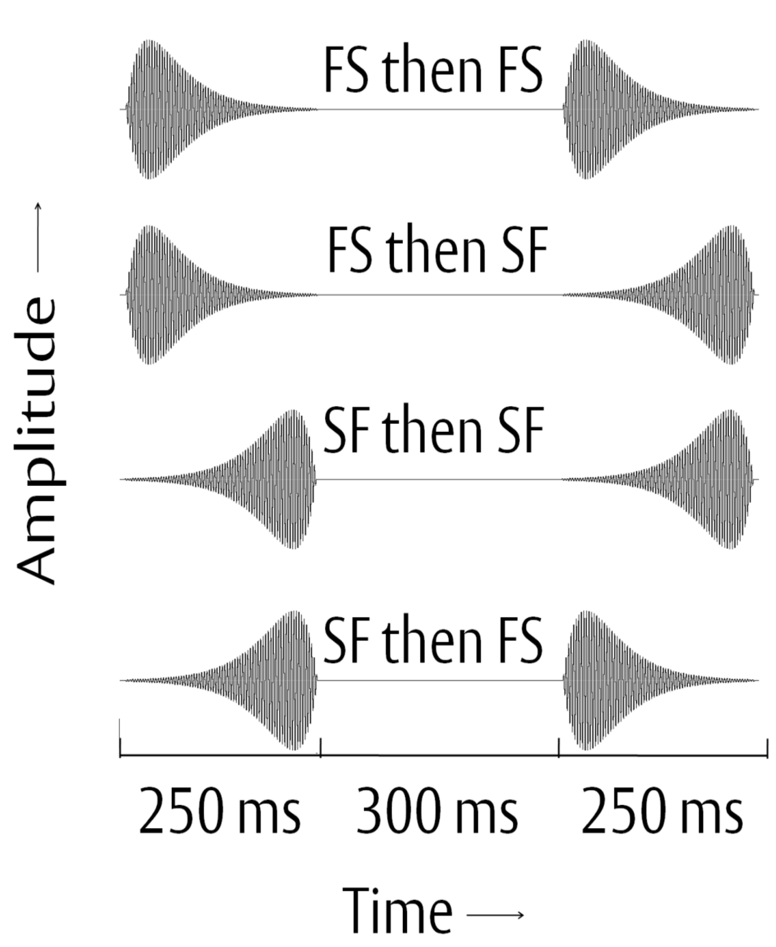


Figure 1. Function-gated 300Hz sinusoids, presented in a whole trial with 300 -ms ISI, as used by Stecker & Hafter (2000).

Binaural room impulse responses, (BRIRs)

Sine-swept measurement signal (20-22kHz)

Played through KEMAR dummy with directional transducer mouthpiece

In real room

Recorded at Head-and-Torso-Simulator with mikes in ears.

To produce BRIRs

Real-room listening

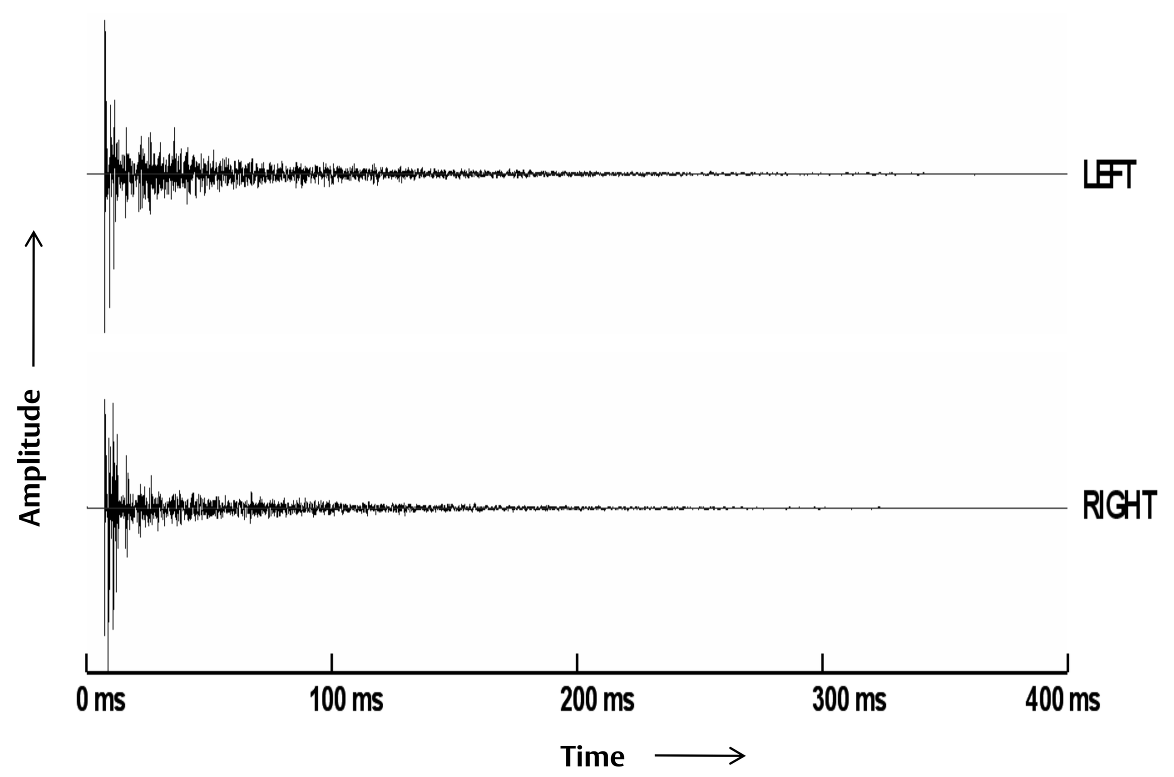


Figure 2. BRIR recorded in a real seminar–room with a distance of 2.50 m between source and listener, giving a ratio of early (50 ms) to late energy of 9.71 dB (C_{50}, left ear, A-weighted).

FS and SF sounds convolved with BRIRs

FS and SF sounds convolved with BRIRs → Played through headphones → "Real-room" listening

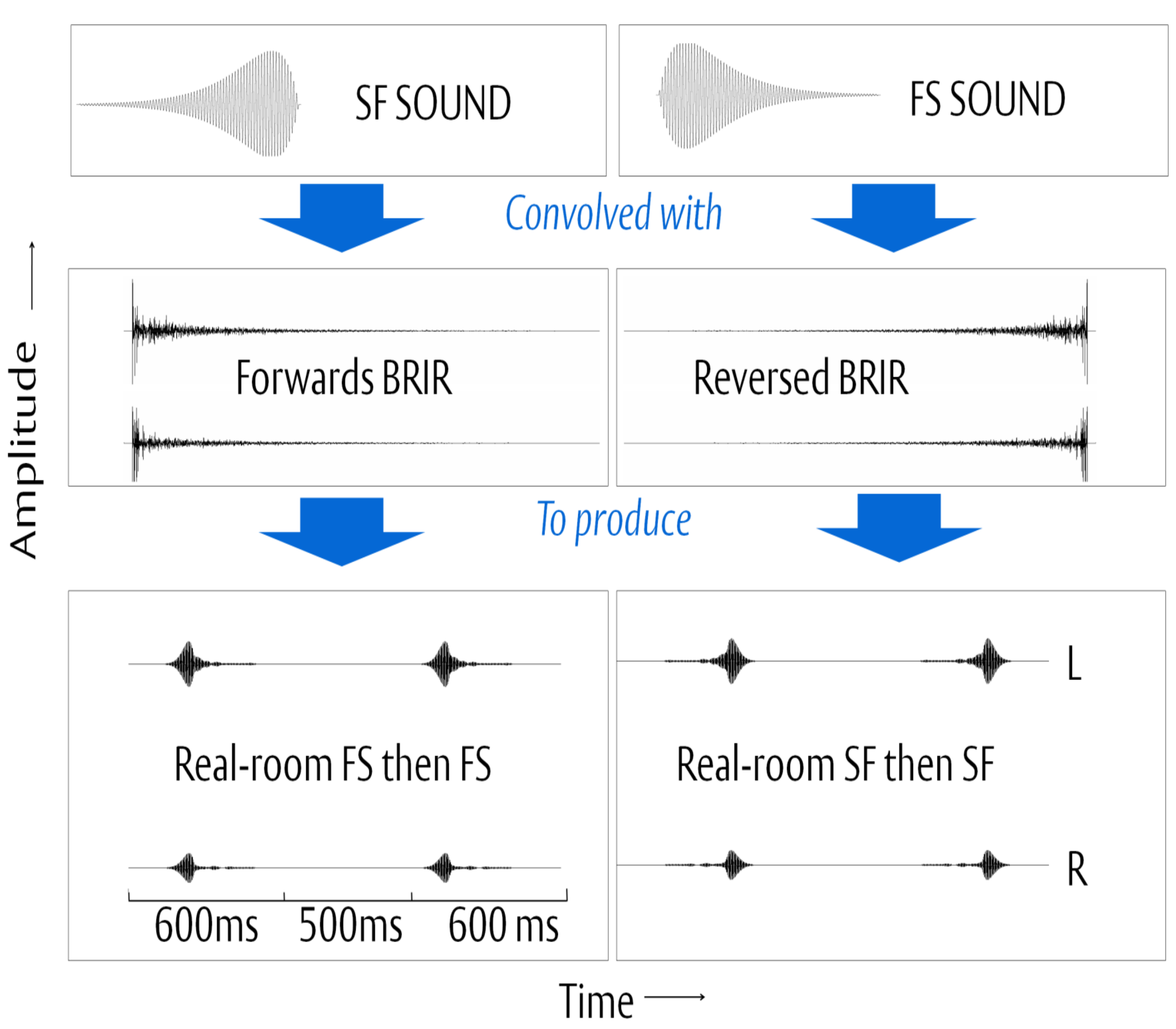


Figure 3. BRIRs were convolved with function-gated stimuli in such a way that that the time-direction of the function-gated 'tail' was the opposite of the BRIR's tail. Hence, SF stimuli were convolved with forwards BRIRs, and FS stimuli were convolved with reversed BRIRs. This gave a 'real room FS' sound that has a (forwards) BRIR-tail at the end, and a 'real room SF' sound that has a slow onset from a (reversed) BRIR at its onset. Real-room FS with FS pairings were formed with these sounds.

Experiment

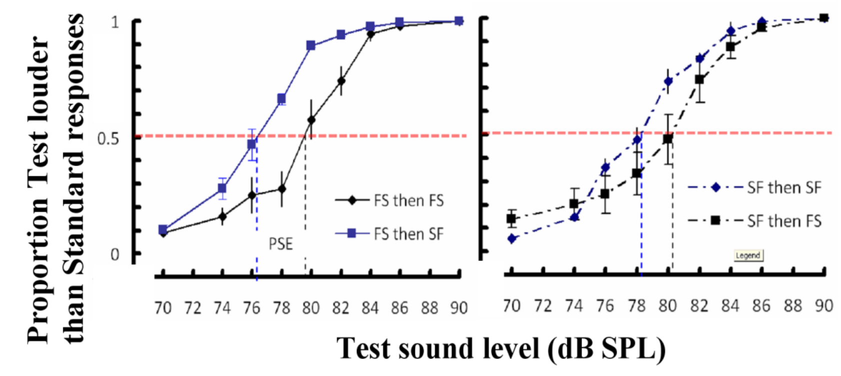
On a trial, listeners seated in an IAC 1201 booth heard one of the 4 'real-room' FS with SF pairings over headphones, to which they responded with a two-interval relative-loudness judgment (2IC) . The 80-dB SPL standard was always the first of the pair, and it was followed by test sounds whose 9 possible levels were selected from an 18-dB range around 80 dB SPL. Listeners responded by clicking one of two buttons (first or second sound 'is louder') on a computer screen viewed through the booth's window. SF or FS standards were judged in different trial-blocks, so that each block had two of the 4 types of pairing. There were three repeats of each trial-type in a block's randomised 54-trial sequence. There were two of these blocks for each of the three types of presentation configuration (two monaural and one dichotic configuration). Four participants aged between 20 and 27 (2 males 2 female) completed all 6 of these blocks ten times in separate 12-block experimental sessions.

Conditions

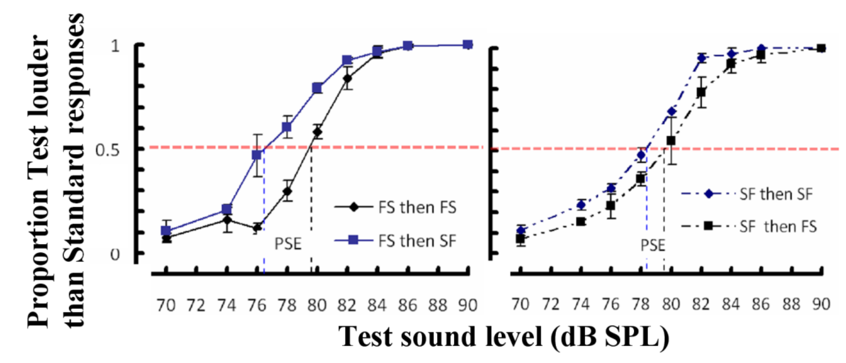
The three factors were: **pairing** (real room FS with SF, FS with FS, SF with SF and SF with FS), **presentation configuration** (Monaural Left , Monaural Right or Dichotic) and **test sound level** (70, 72, 74, 78, 80, 82, 82, 86, or 90 dB SPL) .

Data

Monaural Left



Monaural Right



Dichotic

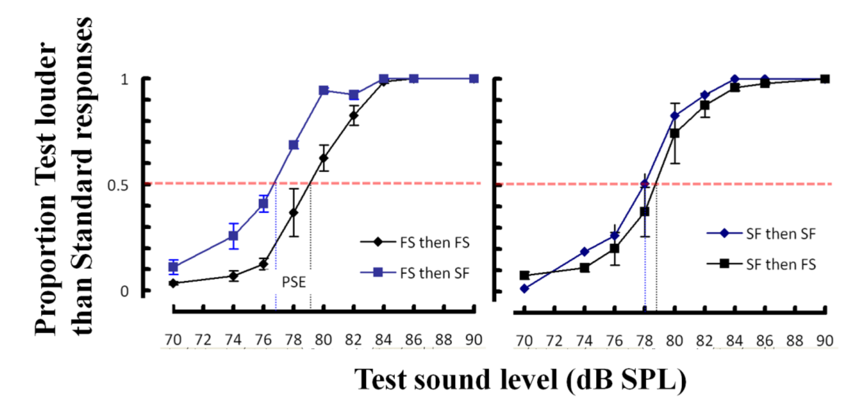


Figure 4. Proportion of trials where listeners responded that the test sound was louder than the standard sound (4 listeners, 120 observations per point). On the left, loudness functions for the real-room FS with FS pairings and real-room FS with SF pairings, under three presentation configurations. On the right, the corresponding functions for the real-room SF with SF pairing and real-room SF with FS pairings. The Points of Subjective Equality (PSE) for each pairing are indicated on the plots.

The effects of BRIR tails have overridden the effects of function-gated tails to produce a loudness asymmetry between real–room FS and real-room SF sounds. This is shown in both monaural conditions and in the dichotic condition.

Loudness context effect; pooled across tests' dB levels

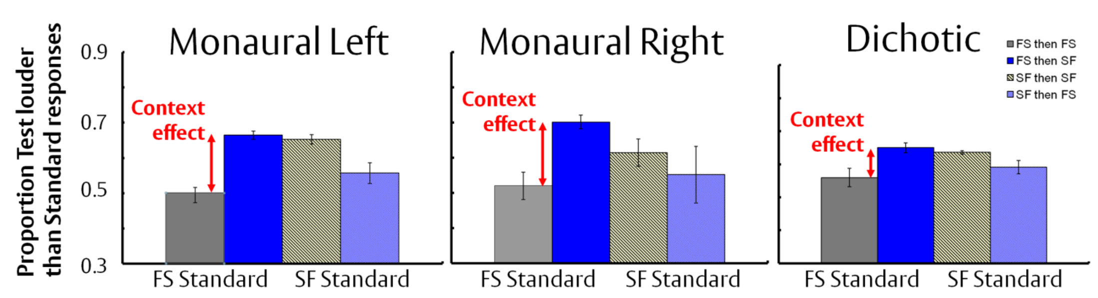


Figure 5. Proportion of trials where listeners responded that the test sound was louder than the standard sound, pooled across all 9 test sound levels and averaged across the 4 listeners. The loudness context effect is indicated in red.

The loudness context effect is obtained with the real-room FS with SF pairings for all presentation configurations. This context effect is substantially increased in the monaural conditions.

Conclusions

Effects of the real-room tails successfully oppose effects of the function-gated tails, as there was a substantial context effect that depended on the direction of the real-room tail.

As with 'tail effects' in speech, this context effect is found in both monaural and dichotic conditions, but is less prominent in dichotic conditions . There appears to be a 'dereverberation' in dichotic conditions that may be due to the de-correlation between the two ears' signals with the real-room BRIRs.

References

1. Moore, B.C.J, and Glasberg, B.R. (1997), A model for the prediction of thresholds, loudness and partial loudness. *J. Audio Eng. Soc*. **45**, 224-240.
2. Glasberg, B.R. and Moore, B.C.J. (2002), A model of loudness applicable to time-varying sounds. *J. Audio Eng. Soc*. **50 (5)**, 331-342.
3. Stecker, G. C. and Hafter, E. R. (2000). An effect of temporal asymmetry on loudness. *J. Acoust. Soc. Am*. 107 3358-3368.
4. Watkins, A.J. (2005) Perceptual compensation for effects of reverberation in speech identification. *J. Acoust. Soc*. Am. **118** 249-262

Acknowledgements

Supported by EPSRC