Speech segregation in rooms: Monaural, binaural, and interacting effects of reverberation on target and interferer

Mathieu Lavandier[a] and John F. Culling[b]

School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3AT,
United Kingdom

Speech reception thresholds were measured in virtual rooms to investigate the influence of reverberation on speech intelligibility for spatially separated targets and interferers. The measurements were realized under headphones, using target sentences and noise or two-voice interferers. The room simulation allowed variation of the absorption coefficient of the room surfaces independently for target and interferer. The direct-to-reverberant ratio and interaural coherence of sources were also varied independently by considering binaural and diotic listening. The main effect of reverberation on the interferer was binaural and mediated by the coherence, in agreement with binaural unmasking theories. It appeared at lower reverberation levels than the effect of reverberation on the target, which was mainly monaural and associated with the direct-to-reverberant ratio, and could be explained by the loss of amplitude modulation in the reverberant speech signals. This effect was slightly smaller when listening binaurally. Reverberation might also be responsible for a disruption of the mechanism by which the auditory system exploits fundamental frequency differences to segregate competing voices, and a disruption of the "listening in the gaps" associated with speech interferers. These disruptions may explain an interaction observed between the effects of reverberation on the targets and two-voice interferers.
© 2008 Acoustical Society of America. [DOI: 10.1121/1.2871943]

I. INTRODUCTION

In order to understand speech in a noisy environment or when many people are talking at the same time, the auditory system has to segregate the target speech from the competing speech or noise. The task becomes even more complicated in rooms, where the auditory system also has to cope with the effects of reverberation. Speech intelligibility against noise or speech interferers decreases when reverberation increases. Lavandier and Culling (2007) showed that the loss of intelligibility could be associated with at least two distinct effects of reverberation.

The reverberation reduced intelligibility by affecting the interfering source. This effect was interpreted as resulting from the decorrelation of the interferer at the two ears. Licklider (1948) showed that speech intelligibility in noise diminishes when the interaural coherence of the noise is reduced, and Robinson and Jeffress (1963) observed a similar effect for tone detection in noise. This effect is predicted by all binaural unmasking theories. For example, the equalization-cancellation (E-C) mechanism of Durlach (1972) predicts that a less correlated masker will be more difficult to equalize at the two ears, and consequently more difficult to cancel, resulting in lower speech intelligibility or poorer tone detection. This effect of reverberation on the interferer depends on the interaural coherence of the interfering source. The interaural coherence of a source evaluates the similarity of the sound waveform it produces at the two ears of the listener. It

is calculated by taking the maximum value from the interaural cross-correlation function. The interaural coherence of the interfering source in a room is degraded by the multiple sound reflections reaching the listener, because these reflections are not the same at the two ears (unless the configuration is symmetrical within the room).

The reverberation also directly affected the target. This effect was interpreted as the intrinsic degradation of speech intelligibility in reverberation, occurring even without the interferer. Target intelligibility decreases because of the temporal and spectral distortions resulting from the multiple delayed sound reflections mixing with the direct sound in the room. This effect is related to the direct-to-reverberant (D/R) ratio of the source at the listening position. It can be described by objective measurements such as the useful-to-detrimental ratios (Bradley, 1986; Bradley *et al.*, 1999) or the speech transmission index (STI), which takes into account the reduction of amplitude modulation in the speech signals due to reverberation (Houtgast and Steeneken, 1985). In Lavandier and Culling (2007), the loss of intelligibility due to the effect of reverberation on the target was partially accounted for by STI calculations.

Lavandier and Culling (2007) adopted a parsimonious model to describe their data, hypothesizing that there was no interaction between the two effects described earlier. These two effects are based on different mechanisms (monaural for the target and binaural for the interferer), so there was no reason to suspect that they could interact. However, this remained to be tested. Moreover, discrepancies between STI

[a]Electronic mail: lavandiermn@cardiff.ac.uk.
[b]Electronic mail: cullingj@cardiff.ac.uk.

predictions and speech reception thresholds (SRTs) measured with the target at different levels of reverberation were not explained by this model.

Three new experiments were designed to verify that there was no interaction between the effects of reverberation on target and interferer. They used methods very similar to those of Lavandier and Culling (2007). SRTs (i.e., the level of the target compared to that of the interferer for 50% intelligibility of the target) were measured for spatially separated targets and interferers in a virtual room. The room simulation allowed variation of the absorption coefficient of the room boundaries, thus modifying the D/R ratio and the interaural coherence of sources. A different absorption coefficient was used for target and interferer. This experimental design is not realistic as it implies listening simultaneously to two sources in rooms having different reverberation characteristics, but it offers the advantage of controlling interferer and target reverberation conditions independently, decomposing the influence of reverberation on speech intelligibility. The effects of the room and the head on the sound level of each source at the listener's ears were eliminated by equalizing all stimuli in level, in order to interpret the effect of interaural coherence/direct-to-reverberant ratio unambiguously. Target and interferer were also clearly identified, so that the listener knew which voice to listen to[1].

In Lavandier and Culling (2007), four absorption coefficients were tested for the source with varying reverberation, while a unique absorption coefficient was used for the source with fixed reverberation. In experiments 1 and 3 presented here, four absorption coefficients were used for both sources. They were varied independently for target and interferer across conditions, so that SRTs were measured for the four types of target against the four types of interferer. A potential interaction between the effects of reverberation on target and interferer could then be investigated, by comparing the effect of varying reverberation on the first source at the different levels of reverberation of the second source. Experiment 1 used speech interferers, whereas experiment 3 used noise interferers. Experiment 2 used a subset of the conditions from experiment 1.

As an interaction was observed between the effects of reverberation on the targets and speech interferers, the parsimonious model hypothesized by Lavandier and Culling (2007) had to be refined. Two further experiments investigated more precisely and individually the effects of reverberation on target and interferer. It should be noted that D/R ratio and interaural coherence do not necessarily covary. The example of a symmetrical configuration of the listener and the sound source within a room illustrates this point. Because the configuration is symmetrical, the interaural coherence of the source is 1 even if its D/R ratio is reduced by making the room more reverberant. If the sound reflections are the same at the two ears, they do not affect the interaural coherence. In experiments 4 and 5, the D/R ratio and interaural coherence of sources were then varied independently by considering both binaural and diotic listening conditions. By definition, diotic stimuli had a fixed interaural coherence of 1, whereas the coherence of binaural stimuli was reduced by reverberation. The aim of experiments 4 and 5 was to tease apart the binaural and monaural effects of reverberation, in order to elucidate the mechanism responsible for the observed interaction, and to show that the main effect of reverberation on the target was monaural and associated with the D/R ratio, wheareas the main effect of reverberation on the interferer was binaural and mediated by the coherence.

II. GENERAL METHODS

A. Stimuli

A male voice was used as the target source in each experiment. The corpus of sentences was from the Harvard Sentence List (IEEE, 1969). The anechoic recordings of a male voice digitized at 20 kHz with 16-bit quantization were used as the basis of all stimuli. The sentences have low predictability, and each sentence contains five key words. For instance, one sentence was "the POINT of the STEEL PEN was BENT and TWISTED". Interfering sentences were generated by feminizing the male voice using the Praat PSOLA speech analysis and resynthesis package. The voice was increased in fundamental frequency by a factor of 1.8 (Rendall et al., 2005), and a vocal tract 9% shorter (Fitch and Giedd, 1999) was simulated by shifting the spectral envelope up in frequency (Darwin and Hukin, 2000). The resulting female voice was used to create two-voice interferers. The aim of using a female voice for the interferer was to eliminate some attentional effects encountered in previous experiments using the same voice for target and interferer (Lavandier and Culling, 2007). The difference of voice between target and interferer is a cue that helps the listener to focus his/her attention on the target voice, and this cue has been shown to be very resistant to reverberation (Darwin and Hukin 2000). Noise interferers were also used. They were speech-shaped noises, obtained by filtering Gaussian noises with a finite impulse response (FIR) filter designed to match the long-term excitation pattern (Moore and Glasberg, 1983) of the two-voice interferers.

The virtual rooms used in the experiments were simulated using a ray-tracing method (Allen and Berkley, 1979; Peterson, 1986), implemented in the |WAVE signal processing package (Culling 1996). The absorption coefficients of the room internal surfaces were all set to the same value, and varied together when the reverberation condition was changed. The listener was modeled as two ears with no head between them, using omnidirectional microphones suspended in space at 1.5 m from the floor. The head was not modeled in order to avoid head shadow effects which could complicate the interpretation of the effects associated with interaural coherence. The interferer and target were placed at different positions in the room, also at 1.5 m from the floor. Binaural stimuli were produced by calculating the impulse responses between the source positions and each ear, and convolving the speech samples with these impulse responses.

The convolution by the room impulse response modifies the sound levels of the original stimuli (Bradley et al., 1999). The levels change differently depending on the positions used within the room and the ear considered. To avoid mixing the effects of room-induced changes in sound level and interaural coherence/direct-to-reverberant ratio, all stimuli

M. Lavandier and J. F. Culling: Speech segregation in rooms II

were equalized in level after convolution by the room impulse responses. Left and right channels were equalized independently, removing potential interaural level differences.

The D/R energy ratio was calculated at each ear, for every tested source position and absorption coefficient, using the impulse response between the source position and the considered ear. The direct-path and reverberant portions of the impulse response were separated by a time-windowing procedure. This procedure was done by eye and was trivial as our simulated impulse responses were not contaminated by noise. The energy of each portion was computed, and the D/R energy ratio was expressed in decibels.

The interaural coherence of each source at the listening position was also calculated in each experiment. It was computed as the maximum of the absolute value of the crosscorrelation of the waveforms reaching the two ears from the source (Hartmann *et al.*, 2005). These waveforms were filtered between 20 and 1500 Hz prior to calculation, in order to consider only the frequency range for which binaural unmasking is most effective in broadband noise.

B. Procedure

SRTs were measured using a one-up/one-down adaptive threshold method (Plomp and Mimpen, 1979). For each SRT measurement, ten target sentences were presented one after another against the same interferer. For the two-voice interferers, the interfering sentences were displayed on a screen in front of the listeners as they were listening. They were instructed to disregard the female sentences appearing on the screen, and to listen for the male sentence. The target-to-interferer level ratio was initially very low (−32 dB). On the first trial, listeners could either enter a transcript on a computer keyboard, or replay the stimuli. If stimuli were replayed, the target level was increased by 4 dB. Stimuli had to be replayed until the target was loud enough to be judged partially intelligible. Listeners were instructed to attempt a transcript of this first target sentence when they believed that they could hear more than half the words of the sentence. Once the first transcript was entered, the correct transcript was displayed on the computer terminal, with the five key words in capitals. The listener self-marked the number of correct key words. Subsequent target sentences were presented only once, and self-marked in a similar manner. The target level was decreased by 2 dB if the listener correctly identified three or more of the five key words in the previous sentence, and otherwise increased by 2 dB. The SRT for a given condition was taken as the mean target-to-interferer level ratio on the last eight trials.

Each SRT measurement used a different interferer. The session began with two practice runs using unprocessed speech, in order to familiarize listeners with the task. The following runs measured SRTs in each of the N tested conditions in a random order (N=16 in experiments 1–3, N=8 in experiments 4 and 5). The order of the conditions was rotated for successive listeners, whereas sentence materials remained in the same order. Each target sentence was thus presented to every listener in the same order and, across a group of N listeners, a complete rotation of conditions was
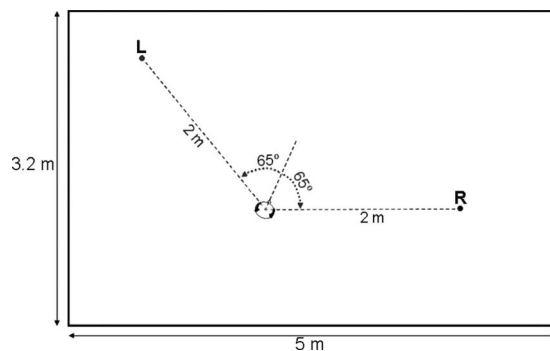

FIG. 1. Virtual room and spatial configuration used in experiments 1–3.

achieved. Each experiment therefore used a multiple of N listeners. This procedure also ensured that each condition was presented in each serial position within the experimental session, counterbalancing order effects.

Signals were digitally mixed, D/A converted, and amplified using a 24-bit Edirol UA-20 sound card and an MTR HPA-2 headphone amplifier. They were presented to listeners over Sennheiser HD650 headphones in a single-walled IAC sound-attenuating booth within a sound-treated room. A computer terminal screen was visible outside the booth window. A keyboard was inside the booth to gather the transcripts of listeners.

C. Listeners

Listeners all reported normal hearing and English as their first language. None of them were familiar with the sentences used during the test. Each listener participated in a single session of the first three experiments and/or of the last two experiments. As experiments 4 and 5 used different target sentences from experiments 1 to 3, a listener could participate in a single session from each group of experiments. Listeners were paid for their participation.

III. INTERACTION BETWEEN THE EFFECTS OF REVERBERATION ON TARGET AND INTERFERER

SRTs were measured for a spatially separated interferer and target at fixed positions within the room, but the room absorption coefficient was varied across conditions, independently for target and interferer. The absorption coefficients were chosen so that they led to different values of direct-to-reverberant ratio and interaural coherence at the listener position.

A. Two-voice interferers (experiment 1)

1. Design

Experiment 1 used the same room and listener position as Lavandier and Culling (2007). The room was 5 m long, 3.2 m wide and 2.5 m high (Fig. 1). The listener was modeled as two ears, separated by 18 cm, placed along an axis at 25° to the 5-m wall on either side of a center point located at 1.2 m from the 5-m wall and 2 m from the 3.2-m wall. The source positions R and L were considered. They were situated at 2 m from the center of the ears, at 65° of azimuth on each side of the virtual listener. The interferer was always

TABLE I. Direct-to-reverberant ratio (dB) at the left and right ears for a source at the target (R) and interferer (L) positions, as a function of the room absorption coefficient used in experiments 1–3.

Absorption coefficient	Target (R)		Interferer (L)	
	Left	Right	Left	Right
1	∞	∞	∞	∞
0.7	0.53	0.79	0.39	−0.22
0.5	−3.70	−3.46	−3.80	−4.48
0.2	−10.93	−10.61	−10.76	−11.67

placed at position L, whearas the target was at position R, keeping a constant azimuth separation of 130° between the two competing sources. When computing the impulse responses between the sources and each ear, four absorption coefficients were used for each source: 1, 0.7, 0.5 and 0.2. Table I shows that the direct-to-reverberant ratio at the listening position decreased with the absorption coefficient. In the anechoic room with the absorption coefficient of 1, the direct-to-reverberant ratio was infinite as there was no reverberant field.

Experiment 1 used only two-voice interferers. The reverberation condition was varied across conditions, independently for target and interferer. A session of the experiment consisted of measuring the SRTs for the four types of target against the four types of interferer, resulting in 16 different conditions. One hundred and sixty target sentences and sixteen two-voice interferers were used to test these conditions. The interaural coherence was calculated for the 160 target stimuli at position R and the 16 interferers at position L. Table II presents the mean results for each type of stimulus. As expected, the coherence was 1 in the anechoic room and decreased with increasing reverberation.

For each listener, the SRTs were measured in the 16 tested conditions during a single 70-min session. Thirty-two listeners took part in the experiment.

2. Results

Figure 2 presents the mean SRTs measured in experiment 1. The two panels contain the same data plotted as a function of the absorption coefficients used for the interferer (top panel) or the target (bottom panel). The intelligibility decreased when the reverberation on either source increased. An analysis of variance (ANOVA) confirmed that the main

TABLE II. Mean interaural coherence with standard deviation for targets (in R) and interferers (in L), as a function of the room absorption coefficient used in experiments 1–3.

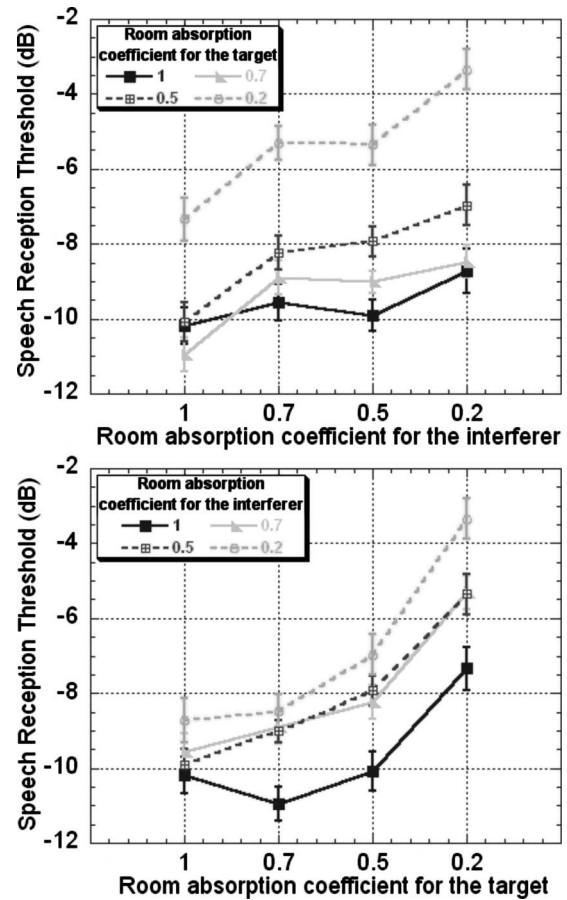| Absorption coefficient	Target (R)	Interferer (L)	
		Two-voice	Noise
1	1 (0.00)	1 (0.00)	1 (0.00)
0.7	0.86 (0.04)	0.86 (0.03)	0.87 (0.00)
0.5	0.75 (0.06)	0.76 (0.04)	0.77 (0.00)
0.2	0.62 (0.07)	0.63 (0.06)	0.61 (0.01)

FIG. 2. Mean SRTs with standard errors measured with two-voice interferers in experiment 1, plotted as a function of the absorption coefficient used for the interferer (top) or as a function of the absorption coefficient used for the target (bottom).

effect of the absorption coefficient was significant for both target $[F(3,93)=99.7, p<0.0001]$ and interferer $[F(3,93)=31.1, p<0.0001]$. Tukey pairwise comparisons were performed in both cases. For the target, only the absorption coefficients 1 and 0.7 were not significantly different $[q>5, p<0.01$ in every other case]. For the interferer, only the coefficients 0.7 and 0.5 were not significantly different $[q>5, p<0.01$ in every other case]. The interaction between the effects of the absorption coefficients used for target and interferer was not significant.

3. Discussion

As the absorption coefficient used for the interferer was decreased, reverberation reduced the interferer interaural coherence (Table II). As a result, this interferer produced more masking, in agreement with binaural unmasking theories, leading to higher thresholds. A less coherent interferer is more difficult to cancel using a binaural E-C mechanism, producing less masking release for spatially separated sources. SRTs also increased with decreasing target direct-to-reverberant ratio (Table I), as the absorption coefficient used for the target was decreased. This loss of target intelligibility is in agreement with the intrinsic degradation of speech in-

telligibility in reverberation, occurring even when no interferer is involved, and associated with the loss of amplitude modulation in the speech signal.

As reverberation increased, speech intelligibility suffered first from the effect of reverberation on the interferer, its effect on the target appearing only at higher reverberation levels. When decreasing the absorption coefficient used for the target, SRTs started increasing only when the coefficient reached 0.5, with the main deterioration of intelligibility appearing when the absorption further decreased to 0.2. For the interferer, decreasing the absorption coefficient from 1 to 0.7 had the strongest effect on SRTs. In the most common situation of conversations taking place in moderately reverberant rooms, the loss of intelligibility due to the room might then be due to the detrimental effect of reverberation on the interferer rather than to its effect on the target.

The interaction between the effects of reverberation on target and interferer did not reach significance in experiment 1. However, Fig. 2 shows evidence of interaction, at least when comparing the extreme absorption coefficients 1 and 0.2. Decreasing the absorption coefficient used for the interferer from 1 to 0.2 led to a 1.5-dB SRT increase with the anechoic targets, whereas a 4-dB increase was measured with the most reverberant targets. The interpretation of two independent effects of reverberation described above would not predict such a difference. Experiment 2 was run to test whether this difference was only due to the inherent measurement variability, or if it had not reached significance due to a lack of statistical power.

B. Two-voice interferers (experiment 2)

1. Design

Experiment 2 used a subset of the stimuli from experiment 1, and tested only the conditions involving the extreme absorption coefficients 1 and 0.2. The 16 conditions of experiment 1 were thus reduced to four conditions in experiment 2. These four conditions were tested four times for each listener. A single session of experiment 2 involved 16 SRT measurements and two practice trials, like experiment 1. The same 160 target sentences and 16 interferers were used. Eight listeners took part in a 70-min session of the experiment.

2. Results

Figure 3 presents the mean SRTs measured in experiment 2 as a function of the two absorption coefficients used for interferer and target. SRTs were about $1-1.5$ dB lower in experiment 2 compared to experiment 1 (Fig. 2). The observed trends were similar in both experiments, the intelligibility decreasing with increasing reverberation on either source. An ANOVA confirmed that the main effect of the absorption coefficient was significant for both target $[F(1,31)=228.0, p<0.0001]$ and interferer $[F(1,31)=63.2, p<0.0001]$. Decreasing the absorption coefficient used for the interferer from 1 to 0.2 led to a 2-dB SRT increase with the anechoic targets, whereas a 4.5-dB increase was measured with the reverberant targets. The interaction between the effects of the absorption coefficients used for
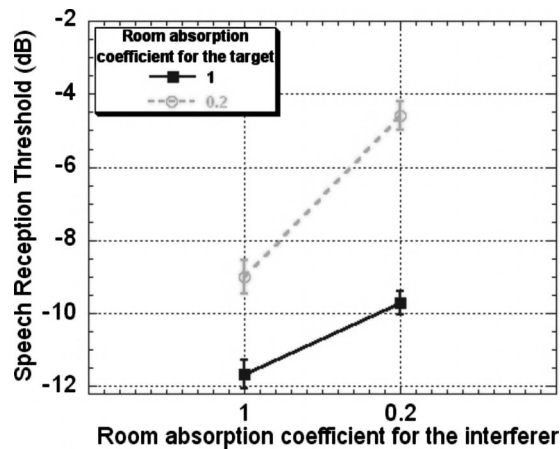

FIG. 3. Mean SRTs with standard errors measured with two-voice interferers in experiment 2, plotted as a function of the absorption coefficient used for the interferer.

target and interferer was significant $[F(1,31)=17.3, p<0.001]$. A simple-main-effects analysis of this interaction showed that the effect of the absorption coefficient used for each source was significant at both levels of reverberation of the other source $[F(1,31)>12, p<0.0015$ in each case].

3. Discussion

Experiment 2 clarified the results of experiment 1. By remeasuring the SRTs in only a few conditions, the statistical power of our analysis increased. It revealed that the effects of reverberation on the target and two-voice interferers were indeed interacting, and that the parsimonious model adopted by Lavandier and Culling (2007) was not sufficient to fully describe speech segregation in rooms. Mechanisms other than the decorrelation of the interferer and the decrease in D/R ratio of the target need to be taken into account to explain the observed interaction. Figure 2 indicates that these mechanisms became important when extreme reverberation conditions were considered.

Experiment 3 investigated whether these mechanisms were also involved when speech interferers were replaced by noise. If so, the effects of reverberation on target and interferer should interact using noise interferers, as they did using two-voice interferers.

C. Speech-shaped noise interferers (experiment 3)

1. Design

Experiment 3 was similar to experiment 1, but with the two-voice interferers being replaced by noise interferers. They were speech-shaped noises based on the long-term excitation pattern of the 16 two-voice interferers of experiment 1 concatenated. The mean interaural coherences of the 16 noise interferers in the four reverberation conditions are presented in Table II. Thirty-two listeners took part in a 70-min session of the experiment.

2. Results

Figure 4 presents the mean SRTs measured in experiment 3. The two panels contain the same data plotted as a
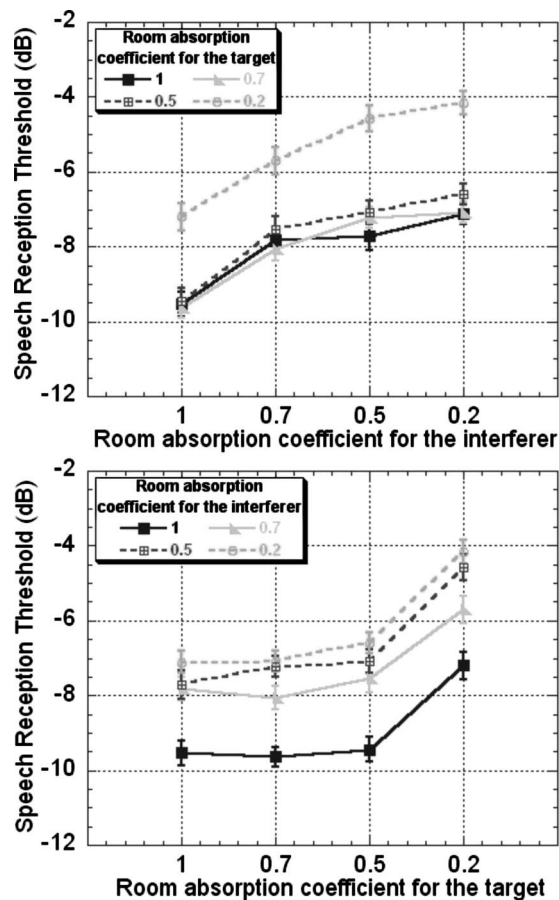
FIG. 4. Mean SRTs with standard errors measured with noise interferers in experiment 3, plotted as a function of the absorption coefficient used for the interferer (top) or as a function of the absorption coefficient used for the target (bottom).

function of the absorption coefficients used for the interferer (top panel) or the target (bottom panel). As with the two-voice interferers of experiment 1 (Fig. 2), the intelligibility decreased when the reverberation of either source increased. An ANOVA confirmed that the main effect of the absorption coefficient was significant for both target [$F(3,93)=85.7$, $p < 0.0001$] and interferer [$F(3,93)=76.5$, $p < 0.0001$]. Tukey pairwise comparisons were performed in both cases. For the interferer, only the coefficients 0.5 and 0.2 were not significantly different [$q > 4.5$, $p < 0.01$ in every other case]. For the target, the absorption coefficient 0.2 led to SRTs significantly higher than those measured with the three other coefficients [$q > 16$, $p < 0.001$ in each case], but the coefficients 1, 0.7 and 0.5 were not significantly different. The effects of the absorption coefficients used for target and interferer did not interact significantly.

3. Discussion

As with the two-voice interferers, speech intelligibility suffered from the effect of reverberation on the noise at lower levels of reverberation than those which affected intelligibility when applied to the target. Decreasing the absorption coefficient used for the target from 1 to 0.7 or 0.5 had no discernible effect on SRTs. The deterioration of intelligibility appeared only when this absorption was further decreased to

0.2. For the interferer, decreasing the absorption coefficient from 1 to 0.7 already resulted in a large increase in SRTs. A further reduction of absorption increased the SRTs only slightly. The coefficients 0.5 and 0.2 did not differ significantly. This outcome is in agreement with the results of Lick-lider (1948) who showed that, for speech in noise, most of the variation in intelligibility observed as a function of the noise interaural coherence occurred for a noise coherence varying between 1 and 0.75. Table II shows that in experiments 1–3, this coherence range corresponded to an absorption coefficient varying between 1 and 0.5. Decreasing the absorption coefficient below 0.5 still decreased the coherence, but below 0.75, and the influence on intelligibility was then limited.

Whereas the effects of reverberation on target and interferer interacted using two-voice interferers, they did not with noise interferers. Therefore, a parsimonious model involving only the decorrelation of the interferer and the decrease in D/R ratio of the target was sufficient to describe the influence of reverberation in experiment 3. The other mechanisms potentially taking place when speech interferers were involved, responsible for the interaction observed in experiments 1 and 2, would then have to be associated with segregation cues which were available with speech interferers but not noise. Fundamental frequency (F0) differences might have constituted such an additional cue to segregate the male target from the female interfering voices (Brokx and Nooteboom, 1982; Culling and Darwin, 1993). Silent periods or "gaps" in the speech masker could also have allowed the listener to hear "glimpses" of the target (Cooke, 2006; Dusquesnoy, 1983; Festen and Plomp, 1990), whereas this cue was not available with the continuous noise interferers. It should be noted that "listening in the gaps" was probably already limited with two-voice interferers compared to what would have been possible with one-voice interferers (Bronkhorst and Plomp, 1992). These two additional cues available with speech interferers and not noise are known to be disrupted by reverberation, and so could be responsible for the interaction observed in experiments 1 and 2. Culling et al. (2003, 1994) showed that reverberation was detrimental to the segregation of nonmonotonous voices by their F0 difference. Bronkhorst and Plomp (1990) showed that listening in the gaps was impaired by reverberation.

The following experiments were designed to investigate more precisely and individually the effects of reverberation on the target (experiment 4) and on the interferer (experiment 5). To separate the effects of reverberation on monaural and binaural processing, the direct-to-reverberant ratio and the interaural coherence were varied independently, by considering both binaural and diotic listening conditions. As in experiments 1 to 3, the coherence of the binaural stimuli was reduced by reverberation, whereas the coherence of the diotic stimuli was by definition fixed to 1 in every condition. In experiments 1 to 3, the D/R ratio and coherence of the sources decreased simultaneously when the room became more reverberant (Tables I and II). In experiments 4 and 5, with the diotic stimuli, the D/R ratio of the sources could be varied without modifying their coherence. In this way, the binaural and monaural effects of reverberation were teased
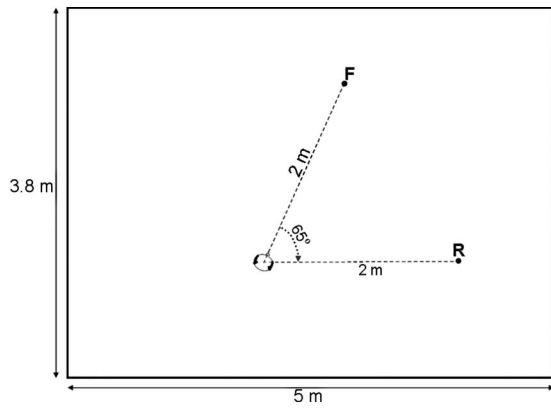
FIG. 5. Virtual room and spatial configuration used in experiments 4 and 5.

TABLE III. Direct-to-reverberant ratio (dB) at the left and right ears for a source at positions F and R, as a function of the room absorption coefficient used in experiments 4 and 5.

Absorption coefficient	Position F		Position R	
	Left	Right	Left	Right
1	∞	∞	⋯	⋯
0.5	⋯	⋯	−3.92	−3.28
0.2	−10.52	−10.50	⋯	⋯

apart, in order to reveal the potential mechanisms responsible for the observed interaction in experiments 1 and 2, and to show that the main effect of reverberation on the target was monaural and associated with the D/R ratio, wheareas the main effect of reverberation on the interferer was binaural and mediated by the coherence.

IV. MONAURAL AND BINAURAL EFFECTS OF REVERBERATION ON THE TARGET (EXPERIMENT 4)

In experiment 4, the effect of reverberation on the target was investigated. The listening mode and the reverberation condition of the target were varied, whereas the interferer remained unchanged through the tested conditions.

A. Design

Figure 5 presents the spatial configuration considered in experiment 4. Both binaural and diotic listening modes were tested for the target in this experiment. To avoid any change in azimuth separation between target and interferer across conditions, the azimuth of the target had to remain the same in the binaural and diotic conditions, and thus the target had to be placed in front of the listener. In this spatial configuration, the interaural coherence of the target decreased less with reverberation than when it was on the listener's side. Therefore, experiment 4 used a slightly larger room than experiments 1–3, in order to avoid having the position F too close to the wall and to get a sufficient effect of reverberation on the coherence of the frontal source. The new room had the same length and height as the previous room, but was 3.8 m wide. The listener was still modeled in the same way, and remained with its center point located at 1.2 m from the 5-m

wall and 2 m from the 3.8-m wall. The interferer was placed at position R, 65° to the right of the listener, as in experiments 1 to 3, such that the competing sources had a constant azimuth separation of 65°. Both F and R were situated 2 m from the center of the ears.

When computing the impulse responses between the sources and each ear, different absorption coefficients were used for the two sources. For the interferer at position R, it was fixed to 0.5. This intermediate value was chosen to avoid the interacting effect found for the extreme levels of reverberation in experiment 2. Noise and two-voice interferers were used. For the target at position F, the two absorption coefficients 1 and 0.2 were tested. Table III presents the D/R ratios computed at the listening position as a function of the room absorption coefficient and the source position. The stimuli corresponding to the interferer were always binaural, whereas two listening modes were considered for the target. The target stimuli were either binaural using the impulse responses between the position F and each ear, or diotic using the signal of the right ear (chosen arbitrarily) for both ears. The results of the interaural coherence calculations for each type of stimulus are presented in bold in Table IV. By definition, the interaural coherence of the diotic targets was 1, even in the reverberant condition. Comparing these coherence values with those of Table II, it can be noted that the coherence of a frontal source in the larger room with the absorption coefficient 0.2 was close to the coherence of a source on the side in the smaller room with the higher absorption coefficient 0.5.

By considering the binaural and diotic listening modes and the two reverberation conditions for the target, we were able to decrease the D/R ratio of the target, and simultaneously either decreasing its coherence in the binaural case or keeping this coherence constant at 1 in the diotic case. If the effect of reverberation on the target is mediated by the

TABLE IV. Mean interaural coherence with standard deviation for targets and interferers at positions F or R, as a function of the room absorption coefficient and the listening mode used in experiments 4 (bold) and 5 (italic).

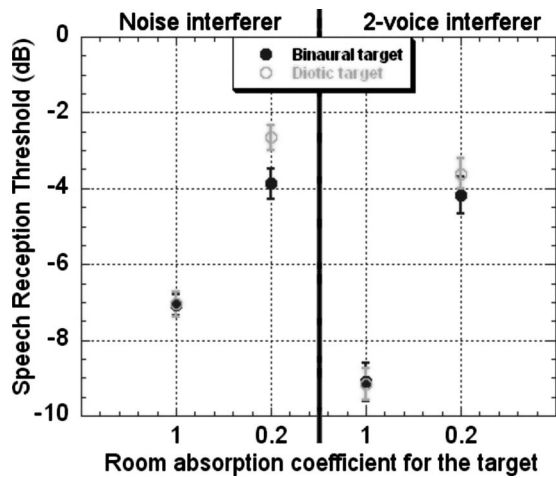| Absorption coefficient, position and listening mode	Target	Interferer	
		Two-voice	Noise
0.5, R, binaural	*0.77 (0.04)*	***0.72 (0.03)***	***0.74 (0.01)***
1, F, binaural and diotic	***1 (0.00)***	*1 (0.00)*	*1 (0.00)*
0.2, F, binaural	***0.77 (0.08)***	*0.74 (0.05)*	*0.70 (0.01)*
0.2, F, diotic	***1 (0.00)***	*1 (0.00)*	*1 (0.00)*

FIG. 6. Mean SRTs with standard errors measured with noise or two-voice interferers in experiment 4, for targets in two reverberation conditions and two listening modes.

D/R ratio and not by the coherence, SRTs in the diotic case should increase with reverberation as in the binaural case.

SRTs were measured in eight conditions corresponding to the two types of interferer (noise and two voice), and the two reverberation conditions (anechoic and reverberant) and two listening modes (binaural and diotic) for the target. Eighty target sentences were used to test these conditions. The sentences were all different from those used in experiments 1–3. Eight interferers of each type were paired with the eight lists of ten target sentences. The noise interferers were based on the long-term excitation pattern of the eight two-voice interferers concatenated. Listeners only heard four interferers of each type, different ones for different listeners, but the order of the conditions was rotated for successive listeners, whereas sentence materials remained in the same order. Across a group of eight listeners, each interferer thus contributed equally to each condition, and any order effect was counterbalanced. Thirty-two listeners took part in a 40-min session of the experiment.

B. Results

Figure 6 presents the mean SRTs measured with noise and two-voice interferers in experiment 4, for the target in the two reverberation conditions and listening modes. For the two types of interferer and the two listening modes, SRTs increased by 3–5 dB when the target was placed in a reverberant room. An ANOVA confirmed that reverberation on the target impaired intelligibility $[F(1,31)=513.7, p<0.0001]$, and also that SRTs were lower with two-voice interferers than with noise $[F(1,31)=31.5, p<0.0001]$. On average, SRTs were higher for the diotic targets than for the binaural ones $[F(1,31)=4.8, p<0.05]$. The interaction between the effects of the listening mode and the reverberation condition of the target was significant $[F(1,31)=4.5, p<0.05]$. A simple-main-effects analysis of this interaction showed that the effect of the listening mode was significant only in the reverberant room $[F(1,31)=9.8, p<0.01]$. The effect of reverberation was significant in both binaural and diotic listening $[F(1,31)>208, p<0.0001$ in each case]. The interaction

between the effects of the interferer type and the reverberation condition of the target was also significant $[F(1,31)=5.6, p<0.05]$. A simple-main-effects analysis of this interaction showed that the two-voice interferers caused less masking than the noise in the anechoic room $[F(1,31)=9.8, p<0.01]$, but not in the reverberant one. The effect of reverberation was significant for both types of interferer $[F(1,31)>181, p<0.0001$ in each case].

C. Discussion

As in experiments 1–3, the reverberation on the target increased SRTs for binaural listening. However, it was also the case for diotic listening, even if the coherence of the reverberant targets was then 1 as in the anechoic room. The loss of intelligibility in reverberation was therefore primarily a monaural effect related to the target D/R ratio, in agreement with the intrinsic degradation of speech intelligibility in rooms, associated with the loss of amplitude modulation in the reverberant speech signals.

A small but significant binaural advantage was measured in reverberation, indicating that having two ears slightly reduced the deleterious effect of reverberation on the target. This binaural improvement in intelligibility may be associated with the "squelching" effect of binaural hearing on perceived reverberation described by Koenig (1950). It was previously measured by Nábělek and Robinson (1982), at different levels of reverberation, using recordings of a target speech in a room with variable absorption characteristics. The percentage of correct words identified by each listener was evaluated using earphones, binaurally and monaurally (with one earphone disconnected). The binaural scores were in average 5% better than the monaural scores. These measurements involved target speech in quiet, with no interferer in the room. Experiment 4 showed that squelching could also take place in the presence of interferers.

SRTs were lower with two-voice interferers than with noise for the anechoic targets. Therefore, at least one additional segregation cue was available with speech interferers which was not available with noise. Moreover, this cue must have been monaural, as the difference of SRTs between the two-voice and noise interferers was found even in diotic listening. As SRTs were the same for the two-voice interferers and the noise when the target was in the reverberant room, this cue must have been ineffective when reverberation was added to the target.

Fundamental frequency differences between the competing voices may have constituted an additional segregation cue available with the two-voice interferers. Culling et al.. (2003, 1994) previously observed a disruption of this cue in reverberation. In those studies, the reverberation was varied simultaneously for target and interferer, but it was hypothesized that the disruption was due to the blurring of the F0 contour of the interferer rather than that of the target, because the segregation mechanism is probably based on the harmonic cancellation of the interferer rather than the harmonic enhancement of the target (de Cheveigné et al., 1995). As the interferer remained unchanged across the tested conditions in experiment 4, if the F0 difference cue was used

and disrupted by reverberation, then it might be relevant to consider also the fundamental frequency of the target.

Listening in the gaps of the speech interferers may also have constituted a segregation strategy which was not available with the continuous noise interferers. These gaps may not have been so useful when the target was reverberant, resulting in an interaction between listening in the gaps and target degradation in reverberation. The deterioration of the target intelligibility would then have been worse when the target was heard through the gaps of the interferer rather than in quiet. It should be noted here that the gaps in the interferer remained the same across the tested conditions, because reverberation was varied for the target but not for the interferer.

Experiment 4 showed that the degradation of intelligibility associated with the loss of amplitude modulation in speech constituted the main effect of reverberation on the target, but that it was not the only mechanism affecting speech segregation in rooms.

V. BINAURAL AND MONAURAL EFFECTS OF REVERBERATION ON THE INTERFERER (EXPERIMENT 5)

In experiment 5, the listening mode and the reverberation condition of the interferer were varied, whereas the target remained unchanged through the tested conditions.

A. Design

Experiment 5 used the same protocol as experiment 4, but replacing the target by the interferer and vice-versa. The same room, listener and source positions were used (Fig. 5), with the same source material for target and interferer, but this time the target was always at position R, whereas the interferer was at position F. The target at position R had a fixed reverberation, with a room absorption coefficient of 0.5. For the interferer at position F, the two absorption coefficients 1 and 0.2 were tested. The target stimuli were always binaural, whereas binaural and diotic listening modes were tested for the interferer. The diotic stimuli used the right channel of the binaural stimuli for both ears. Table III presents the corresponding D/R ratios computed at the listening position and the results of the interaural coherence calculations for each type of stimulus are presented in italic in Table IV.

The coherence of the diotic interferers was always 1, even in the reverberant condition. If the effect of reverberation on the interferer is mediated by its coherence, SRTs should increase with increasing reverberation in the binaural case, but remain constant even when the reverberation is increased in the diotic case.

SRTs were measured for the interferer in eight conditions: two types of stimuli (noise and two-voice), two reverberation conditions (anechoic and reverberant) and two listening modes (binaural and diotic). Thirty-two listeners took part in a 40-min session of the experiment.
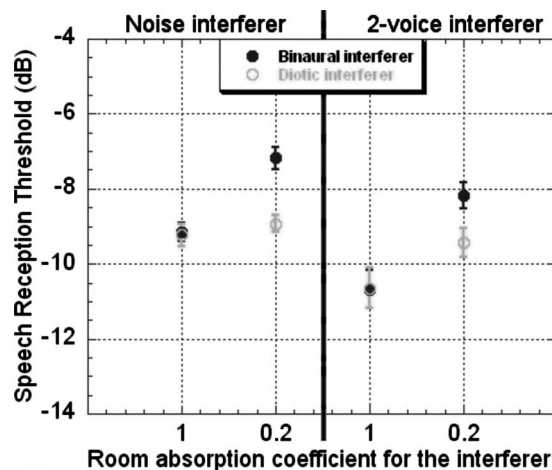

FIG. 7. Mean SRTs with standard errors measured in experiment 5, for noise or two-voice interferers in two reverberation conditions and two listening modes.

B. Results

Figure 7 presents the mean SRTs measured in experiment 5, for noise and two voice interferers in two reverberation conditions and two listening modes. SRTs increased by 2–2.5 dB when the binaural interferers were placed in a reverberant room. The effect of reverberation was very limited for the diotic interferers. An ANOVA confirmed that reverberation on the interferer impaired intelligibility $[F(1,31)=69.0, p<0.0001]$, and also that SRTs were lower with two-voice interferers than with noise $[F(1,31)=21.5, p<0.0001]$. On average, SRTs were lower for the diotic interferers than for the binaural ones $[F(1,31)=10.6, p<0.01]$. The interaction between the effects of the listening mode and the reverberation condition of the interferer was significant $[F(1,31)=10.6, p<0.01]$. A simple-main-effects analysis of this interaction showed that the effect of the listening mode was significant only in the reverberant room $[F(1,31)=20.5, p=0.0001]$. The effect of reverberation was significant in both binaural and diotic listening $[F(1,31)>8.8, p<0.01$ in each case].

C. Discussion

As in experiments 1–3, the reverberation on the interferer increased the SRTs for binaural listening. However, experiment 5 showed that the loss of intelligibility was considerably smaller for diotic interferers. This suggests that adding reverberation to the interferer had a detrimental effect on intelligibility when the interferer coherence was reduced. The effect of reverberation on the interferer was therefore mainly binaural and mediated by interaural coherence. This effect is in agreement with previous binaural unmasking experiments in which reducing the correlation of the interferer at the two ears increased masking (Licklider, 1948; Robinson and Jeffress, 1963).

The decorrelation of the masker constituted the main influence of reverberation on the interferer, but it was not the only one. The effect of reverberation for diotic listening was small but significant. This monaural effect could result from

the reverberation disrupting the segregation cue based on F0 differences between the competing voices. If this cue allows canceling one of the interfering voices by following its F0 contour (de Cheveigné *et al.*, 1995), the disruption could result from the blurring of this contour as reverberation was added to the interferer. Reverberation could also have disrupted the segregation cue based on listening in the gaps of the speech interferers. These interferers could have become more efficient maskers when their potential gaps were filled by the added reverberation, as shown by Bronkhorst and Plomp (1990) when they compared the amount of masking produced by continuous and modulated noises in different reverberation conditions.

The effect of reverberation on the interferer in experiment 5 led to a smaller loss of intelligibility than its effect on the target in experiment 4 (Figs. 6 and 7). However, it should be noted that in these two experiments the source with variable reverberation was placed in front of the listener, and that, in this room, reverberation affects the coherence of a frontal source less than the one of a source on the side, whereas its effect on the D/R ratio mainly depends on the source-listener distance and is rather independent of the azimuth of the source. As a consequence, the decrease of D/R ratio tested for the target in experiment 4 was comparatively larger than the decrease in coherence tested for the interferer in experiment 5.

VI. GENERAL DISCUSSION

The effects measured in experiments 1–5 can be summarized by describing how increasing reverberation sequentially affected speech segregation in the virtual rooms. The increasing reverberation first reduced intelligibility by decorrelating the interferer at the listener's ears, making it a more effective masker as shown by previous binaural unmasking experiments (Licklider, 1948; Robinson and Jeffress, 1963). For example, the equalization-cancellation mechanism (Durlach, 1972) predicts that a less correlated masker should be more difficult to equalize at the two ears, and consequently more difficult to cancel, resulting in lower speech intelligibility. This binaural effect was mediated by the interaural coherence of the interfering source, and affected both the noise and the two-voice interferers. It reduced intelligibility for low levels of reverberation. Further decreases in coherence had less influence, as most of the variation in intelligibility for speech in noise occurs for a noise coherence between 1 and 0.75 (Licklider, 1948). It should be noted that taking into account the influence of the head between the two ears in our simulations might have led to impairments of intelligibility at even lower levels of reverberation. The head would lead to extra interaural decorrelation in asymmetric configurations, particularly for sources to the side (Lindevald and Benade, 1986). This additional decorrelation would make the interferer's coherence drop at even lower levels of reverberation. Because of the floor effect associated with coherence, the maximum intelligibility loss with increasing reverberation should be similar with or without a simulated head.

After decorrelating the interferer, the increasing reverberation distorted the target speech signals (experiments 1–4), as a result of the multiple sound reflections mixing with the direct sound. This monaural effect directly affected the intelligibility of the target for relatively high levels of reverberation. The intelligibility loss was mediated by the direct-to-reverberant ratio of the target at the listening position, and might be described by existing objective measurements (Bradley, 1986; Bradley *et al.*, 1999; Houtgast and Steeneken, 1985). The monaural speech distortions affected the target, but also the two-voice interferers (experiments 1, 2, and 5). Because of these distortions, the fundamental frequency patterns of both target and interferer might have been more difficult to track in reverberation by the auditory system, and thus segregation based on F0 differences could have been disrupted, as noted by Culling *et al.* (2003, 1994). Reverberation might also have filled some silent periods in the two-voice interferers (Bronkhorst and Plomp, 1990). These monaural effects may explain the small but significant additional intelligibility losses measured at the higher levels of reverberation. They might also be responsible for the interaction of the effects of reverberation on the target and two-voice interferers observed in experiments 1 and 2. Because of this interaction, the loss of intelligibility due to the reverberation on one source (target or speech interferer) was larger when the other source was also reverberant rather than anechoic. The D/R ratio of the sources could be relevant to the prediction of these monaural effects. Further investigations are required. For example, it is not clear if the influence of reverberation on the segregation by fundamental frequency difference related to the reverberation of the target, of the interfering voice, or both.

In experiment 4, a small but significant binaural advantage was measured for reverberant targets, with both noise and speech interferers. This advantage might constitute further evidence of a squelching effect (Koenig, 1950; Nábĕlek and Robinson, 1982). In this effect, having two ears may ameliorate the effect of reverberation on target intelligibility, regardless of the interferer. The monaural D/R ratio of the target would not be sufficient to predict this binaural effect, and the target interaural statistics might have to be considered as well.

In experiment 3, the effects of reverberation on the target and noise interferers did not interact. This is consistent with the interaction observed with speech interferers being associated with the disruption of the fundamental frequency difference cue and/or listening in the gaps. As these two cues did not exist when the interferer was noise, there were no grounds for an interaction. Lavandier and Culling (2007) found discrepancies between SRTs measured with noise and STI calculations. An interaction between the effects of reverberation on target and noise was hypothesized as one of the potential explanations for these discrepancies. Experiment 3 of this study did not support this hypothesis. A second account for the discrepancies between measured SRTs and STI predictions could come from the STI predictions being less reliable below the intelligibility level for which the STI is best calibrated (Steeneken and Houtgast, 1980).

M. Lavandier and J. F. Culling: Speech segregation in rooms II

Even though the experiments presented in this paper were realized with a fixed azimuth separation of target and interferer, their results can be generalized to other spatial configurations, as long as target and interferer are not co-located. Spatial unmasking in anechoic situations results partly from head shadow improving the signal-to-noise ratio at the best ear, and partly from binaural interaction associated with the difference in interaural time delays of the competing sources (Bronkhorst and Plomp, 1988). Spatial unmasking depends on the azimuth separation of sources because the head shadow contribution is very dependent on the source azimuths. The binaural interaction contribution proved to be relatively independent of these azimuths as long as the sources are not co-located (Bronkhorst and Plomp, 1988). With increasing reverberation, the head shadow component progressively disappears (Plomp, 1976), and spatial unmasking is reduced to its binaural interaction component. Consequently it becomes independent of the size of the azimuth separation of sources (Beutelmann and Brand, 2006; Plomp, 1976). As our investigations only considered the spatial unmasking associated with binaural interaction, their results should not depend on the magnitude of the tested azimuth separations.

In order to predict speech intelligibility in rooms in the presence of directional noise interferers, the two main effects of interferer coherence and target intrinsic intelligibility should be taken into account, and the parsimonious model hypothesized by Lavandier and Culling (2007) is sufficient. These two effects have been modeled independently. The influence of noise coherence in rooms is implemented in two models based on the equalization-cancellation mechanism (Durlach, 1972). Zurek et al. (2004) proposed a model predicting the segregation of a narrow band noise target from a broadband noise interferer, and Beutelmann and Brand (2006) recently developed a model predicting the intelligibility of a near field speech target against a noise interferer. The STI (Houtgast and Steeneken, 1985) or the useful-to-detrimental ratios (Bradley, 1986; Bradley et al., 1999) evaluate the intrinsic degradation of speech by reverberation. Our goal is to implement these two main effects of reverberation in a single model, which should also account for the squelching effect, a binaural mechanism that ameliorates the deleterious influence of reverberation on target speech. Our model also needs to predict the sound levels of target and interferer at the listener's ears. This will allow consideration of the influence of the room on the sound level of each source at the head (Bradley et al., 1999), and the influence of head-shadow (Bronkhorst and Plomp, 1988). This model will require refinements to be able to predict speech segregation against speech interferers in rooms. The two monaural cues based on the F0 differences between competing voices and the gaps in the speech interferers will have to be taken into account, as they are also influenced by reverberation. Even if these cues were limited in the conditions we tested, they could become much more important in other conditions. For example, F0 differences might become predominant for co-located sources, where coherence becomes irrelevant because spatial unmasking is abolished. Listening in the gaps is known to be much larger with one-voice interferers than with the two-voice interferers we used (Bronkhorst and Plomp, 1992). The influence of reverberation on these two cues might become quite large in these situations. If it cannot be assumed that the listener knows who/where to listen to, then additional attentional effects also have to be incorporated in the model (Kidd et al., 2005; Shinn-Cunningham et al., 2005).

VII. CONCLUSION

SRT measurements in virtual rooms showed that reverberation impaired speech intelligibility first by a decorrelation of the interferer at the two ears, and then by a degradation of the intrinsic target intelligibility. The first effect of reverberation is binaural and mediated by the coherence for the interferer. The second is monaural and associated with the direct-to-reverberant ratio for the target. Having two ears reduced to a small extent the detrimental effect of reverberation on the target. Additional effects of reverberation were observed with speech interferers. The sound reflections in rooms might be responsible for a disruption of the mechanism by which the auditory system exploits fundamental frequency differences to segregate competing voices, and a disruption of the listening in the gaps associated with speech interferers. These disruptions might explain an interaction observed between the effects of reverberation on the targets and two-voice interferers, and could be important to model in very reverberant rooms.

[1]There are some listening situations in which target and interferer dynamically change over time, and other situations in which the listener knows who he/she is listening to, as is the case in many conversations. In our experiments, target and interferer were clearly identified and redirection of attention was not the key issue.

Allen, J. B., and Berkley, D. A. (**1979**). "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am. **65**, 943–950.

Beutelmann, R., and Brand, T. (**2006**). "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **120**, 331–342.

Bradley, J. S. (**1986**). "Predictors of speech intelligibility in rooms," J. Acoust. Soc. Am. **80**, 837–845.

Bradley, J. S., Reich, R. D., and Norcross, S. G. (**1999**). "On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility," J. Acoust. Soc. Am. **106**, 1820–1828.

Brokx, J. P. L., and Nooteboom, S. G. (**1982**). "Intonation and the perceptual separation of simultaneous voices," J. Phonetics **10**, 23–36.

Bronkhorst, A. W., and Plomp, R. (**1988**). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am. **83**, 1508–1516.

Bronkhorst, A. W., and Plomp, R. (**1990**). "A clinical test for the assessment of binaural speech perception in noise," Audiology **29**, 275–285.

Bronkhorst, A. W., and Plomp, R. (**1992**). "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing," J. Acoust. Soc. Am. **92**, 3132–3139.

Cooke, M. (**2006**). "A glimpsing model of speech perception in noise," J. Acoust. Soc. Am. **119**, 1562–1573.

Culling, J. F. (**1996**). "Signal processing software for teaching and research in psychoacoustics under UNIX and X-windows," Behav. Res. Methods Instrum. Comput. **28**, 376–382.

Culling, J. F., and Darwin, C. J. (**1993**). "Perceptual separation of simultaneous vowels: Within and across-formant grouping by f0," J. Acoust. Soc. Am. **93**, 3454–3467.

Culling, J. F., Hodder, K. I., and Toh, C. Y. (**2003**). "Effects of reverberation on perceptual segregation of competing voices," J. Acoust. Soc. Am. **114**, 2871–2876.

Culling, J. F., Summerfield, Q., and Marshall, D. H. (**1994**). "Effects of simulated reverberation on the use of binaural cues and fundamental-frequency differences for separating concurrent vowels," Speech Commun. **14**, 71–96.

Darwin, C. J., and Hukin, R. W. (**2000**). "Effects of reverberation on spatial, prosodic, and vocal-tract size cues to selective attention," J. Acoust. Soc. Am. **108**, 335–342.

de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M. (**1995**). "Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement," J. Acoust. Soc. Am. **97**, 3736–3748.

Durlach, N. I. (**1972**). "Binaural signal detection: Equalization and cancellation theory," in *Foundations of Modern Auditory Theory*, edited by J. Tobias (Academic, New York), Vol. **II**, pp. 371–462.

Dusquesnoy, A. J. (**1983**). "Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons," J. Acoust. Soc. Am. **74**, 739–743.

Festen, J. M., and Plomp, R. (**1990**). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," J. Acoust. Soc. Am. **88**, 1725–1736.

Fitch, W. T., and Giedd, J. (**1999**). "Morphology and development of the human vocal tract: a study using magnetic resonance imaging," J. Acoust. Soc. Am. **106**, 1511–1522.

Hartmann, W. M., Rakerd, B., and Koller, A. (**2005**). "Binaural coherence in rooms," Acta. Acust. Acust. **91**, 451–462.

Houtgast, T., and Steeneken, H. J. M. (**1985**). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," J. Acoust. Soc. Am. **77**, 1069–1077.

IEEE. (**1969**). "IEEE recommended practice for speech quality measurements," IEEE Trans. Audio Electroacoust. **17**, 227–246.

Kidd, G., Mason, C., Brughera, A., and Hartmann, W. (**2005**). "The role of reverberation in release from masking due to spatial separation of sources for speech identification," Acta. Acust. Acust. **91**, 526–535.

Koenig, W. (**1950**). "Subjective effects in binaural hearing," J. Acoust. Soc. Am. **22**, 61–62.

Lavandier, M., and Culling, J. F. (**2007**). "Speech segregation in rooms: Effects of reverberation on both target and interferer," J. Acoust. Soc. Am. **122**, 1713–1723.

Licklider, J. C. R. (**1948**). "The influence of interaural phase relations upon masking of speech by white noise," J. Acoust. Soc. Am. **20**, 150–159.

Lindevald, I. M., and Benade, A. H. (**1986**). "Two-ear correlation in the statistical sound fields of rooms," J. Acoust. Soc. Am. **80**, 661–664.

Moore, B. C. J., and Glasberg, B. R. (**1983**). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," J. Acoust. Soc. Am. **74**, 750–753.

Nábělek, A. K., and Robinson, P. K. (**1982**). "Monaural and binaural speech perception in reverberation for listeners of various ages," J. Acoust. Soc. Am. **71**, 1242–1248.

Peterson, P. M. (**1986**). "Simulating the response of multiple microphones to a single acoustic source in a reverberant room," J. Acoust. Soc. Am. **80**, 1527–1529.

Plomp, R. (**1976**). "Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise)," Acustica **34**, 200–211.

Plomp, R., and Mimpen, A. M. (**1979**). "Improving the reliability of testing the speech-reception threshold for sentences," Audiology **18**, 43–52.

Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (**2005**). "Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry," J. Acoust. Soc. Am. **117**, 944–955.

Robinson, D. E., and Jeffress, L. A. (**1963**). "Effect of varying the interaural noise correlation on the detectability of tonal signals," J. Acoust. Soc. Am. **35**, 1947–1952.

Shinn-Cunningham, B., Ihlefeld, A., Satyavarta, and Larson, E. (**2005**). "Bottom-up and top-down influences on spatial unmasking," Acta. Acust. Acust. **91**, 967–979.

Steeneken, H. J. M., and Houtgast, T. (**1980**). "A physical method for measuring speech-transmission quality," J. Acoust. Soc. Am. **67**, 318–326.

Zurek, P. M., Freyman, R. L., and Balakrishnan, U. (**2004**). "Auditory target detection in reverberation," J. Acoust. Soc. Am. **115**, 1609–1620.