capstract:2611.00059v1 [eess.AS] 20 Nov 2026

Estimating Reader-Hours from the Addressees of Recorded Quiet Notices

Maelis Quessavrel

***Abstract*—A request for quiet identifies a seat and supplies a nominal addressee even when the addressee makes no further sound. We examine an attendance estimator that credits the seat with a reader for one minute after each recorded shush. Two desks with local microphones and notice playback were simulated in a vacant reading room. One scheduled playback initiated each trial and was excluded from attendance. At a threshold of 40 dB SPL, the installation registered 1.013 reader-hours over the next hour, averaged over 160 desk placements (95% interval 0.859 to 1.166). At 80 placements each desk supplied sufficient acoustic evidence to renew the other seat indefinitely. The long-run attendance over the sampled placements was exactly one reader per vacant room. Shortening the attendance interval from 60 to 10 s reduced the hourly estimate by 0.010 reader-hours. Subtracting the direct sound path reduced it by 0.998 reader-hours. Exact playback subtraction eliminated attendance in vacancy and retained the response to all independently audible page rustles in a paired overlap test.**

***Index Terms*—Acoustic occupancy sensing, sound event detection, acoustic echo cancellation**

I. Introduction

A seating sweep records a reader where an observer finds one [1]. It gives no arrival or departure time between sweeps. Desk microphones provide continuous timestamps but have little to record during sustained reading. We consider supplementing the seating record with the addressees of requests for quiet. A request made at desk B opens an attendance interval at B. A further request renews the interval on the assumption that the reader is still there. Records at different desks may overlap. Within one desk they count only once.

This convention assigns a reader to a place without requiring a verbal acknowledgment. It does require that a notice have a recipient. We study that requirement in an empty room equipped with two desks, each capable of detecting a disturbance and delivering its own notice. A scheduled sound check is identifiable in the log and receives no seat credit. The question is how much subsequent attendance can be recovered from records that carry no such exclusion.

We integrate occupied seats over time. Multiple notices to one seat can then describe a single visit. The true number of reader-hours in every vacancy trial is zero.

II. From microphone level to attendance

Each recording desk has one seat. When it detects a disturbance, the desk plays a 180 ms shush. It enters the time and its own seat identifier in the notice register. No sound is attributed to a named reader. Let the attendance interval be W seconds from issue. At each seat we take the union of its open intervals, truncate at the reporting boundary, and sum their durations across seats. Dividing this total by 3,600 gives reader-hours. We set $W = 60$ s to allow for quiet reading between observations.

The vacancy trial begins with a sound check at desk A at time zero and an empty attendance register. If B detects the sound check, B addresses its own seat at 0.4 s. This is an ordinary notice in the register. A may then address its seat at 0.8 s. Thus the first two inferred readers, when present, have different seat identifiers. Deduplication within seats cannot combine them.

A. Sources and propagation

Each desk holds 32 versions of a 180 ms shush, sampled at 16 kHz. A token consists of first-differenced uniform white noise under a squared-sine envelope with a single maximum at 90 ms. The RMS level over the whole token at 1 m is drawn uniformly from 46 to 54 dB SPL. Each playback selects a token independently with replacement. The two banks are generated once and kept across installations, as they would be on a pair of devices moved between rooms. These synthetic tokens represent only the unvoiced part of a request. Natural fricatives also vary in spectral shape [2].

We used 160 independent placements in a 10 m by 7 m room with a 3 m ceiling. Both transducers are at a height of 1.2 m. The first desk is uniform over the floor area more than 1 m from any side wall. The second is drawn from the same area until its distance from the first is between 1.5 and 7 m. Each loudspeaker is colocated with its own microphone. A transmission to the other desk includes the direct path and the six first-order wall, floor and ceiling images [3]. A path of length d meters has pressure gain $1/d$ for direct sound and $0.55/d$ for an image. Delays are rounded to the nearest sample at 343 m/s. The channel is reciprocal and constant within a placement. The model includes early reflections only.

B. Decision and playback

Starting from a forced notice at desk A, time is divided into 400 ms slots. The desk playing a notice is deaf for that slot. Its peer forms unweighted RMS pressure levels in successive, nonoverlapping 20 ms frames during the first 240 ms. All seven arrivals end within this window. Two adjacent frames at or above the threshold schedule one notice for the start of the next slot. A desk issues at most one notice per slot. The first missed response ends the exchange.

Within a placement, let p_A be the fraction of A’s bank that elicits a reply from B, and p_B the converse. The number N of replies after the initial shush satisfies

$$P(N \ge h) = p_A^{\lfloor h/2 \rfloor} p_B^{\lfloor h/2 \rfloor}. \tag{1}$$

In particular a pair is certain to exchange notices indefinitely under this model if every token clears the opposite desk’s threshold. If either bank has even one token that fails, the pair eventually falls silent with probability one.

III. Calculation and controls

For the one-hour account we use 9,000 slots of length $\Delta = 0.4$ s. A response born at the left boundary occupies that slot. The initial test at slot zero contributes no credit. Let $L = W/\Delta$ and consider slot n. The earliest notice that can still support attendance there has index $j = \max(1, n - L + 1)$. Because the exchange alternates without gaps until extinction, attendance depends only on whether the exchange reached the first odd and the first even index in the eligible window. Writing $S(h) = P(N \ge h)$ from Eq. (1) gives the expected occupied seats

$$E[C_n] = S(j) + \mathbf{1}_{j<n} S(j+1), \quad n \ge 1. \tag{2}$$

Summing Eq. (2) over the hour and multiplying by $\Delta/3600$ gives the expected reader-hours for a fixed placement. This calculation includes the last credited minute after a pair stops exchanging notices.

We enumerated both 32-token banks for every placement and receiver. All installations have equal weight. Intervals for means are 1.96 standard errors across 160 placements, conditional on the installed banks. We report 10 s and 300 s attendance intervals as sensitivity checks. The 32-bit linear congruential generator used initial state 48171, multiplier 1664525 and increment 1013904223. Draws for the two playback banks and then the rustle bank preceded placement.

A. Receiver controls

The baseline receiver measures the unweighted level. A peer veto ignores the listening window if the other desk has declared a playback. Three reference receivers subtract the known remote waveform before measuring level. One uses the direct arrival only. One uses all seven true taps. The timing control uses those same seven taps delayed by one sample.

A vacant-room result alone cannot distinguish a receiver that cancels a notice from one that has stopped accepting local sounds. We used a separate paired overlap test, following the distinction between remote playback and local activity in echo cancellation [4]. Page turns occur in synthetic office event evaluations [5]. Our probe is a 40 ms burst of uniform noise under a squared-sine envelope, scaled to a received RMS level drawn uniformly from 40 to 48 dB SPL. A bank of 32 probes was rotated by placement, direction and token index. The probe began 80 ms into a shush, near its level maximum. Each of the 10,240 mixtures had a shush-only and a probe-only counterpart.

When a probe alone cleared the threshold, the receiver received credit for hearing it only if adding it to the shush changed the decision from no notice to notice. Onset agreement, as used in event scoring [6], would credit a receiver already responding to the playback.

IV. Results

A. Attendance during vacancy

With the default one-minute interval the level gate returned 1.013 reader-hours per hour of vacancy (95% interval 0.859 to 1.166). The mean exceeded one because short exchanges also furnish attendance near the beginning of the hour. In the first minute alone the register averaged 1.695 occupied seats. By the end of 15 minutes the cumulative mean had fallen to 1.050. True occupancy throughout all three periods was zero.

Table I shows the effect of the interval used to retain each addressee. The first-minute means coincide for 60 and 300 s because no record can expire within either window during that minute. Shortening the interval sixfold removed less than one hundredth of a reader-hour from the hourly estimate. In an installation where both seats keep receiving notices every 0.8 s, even the 10 s rule renews each reader well before expiry.

TABLE I
Mean occupied seats credited by the level gate at 40 dB SPL over reporting periods measured from the sound check. A one-hour mean also gives reader-hours accrued.

Retention (s)	First minute	First 15 min	First hour
10	1.157	1.011	1.003
60	1.695	1.050	1.013
300	1.695	1.240	1.060

In 80 placements, every token from each desk cleared the peer’s threshold. In the other 80 there was at least one opportunity for the exchange to end on each round trip. For any fixed retention of at least two slots, the first group converges to two occupied seats and the second to zero. The limiting mean for these installations is consequently one registered reader. Desk separation predicted which pair could maintain a complete register (Fig. 1). All 54 pairs separated by less than 3 m sustained both seats indefinitely, against 26 of 75 at 3 to 5 m and zero of 31 beyond 5 m.

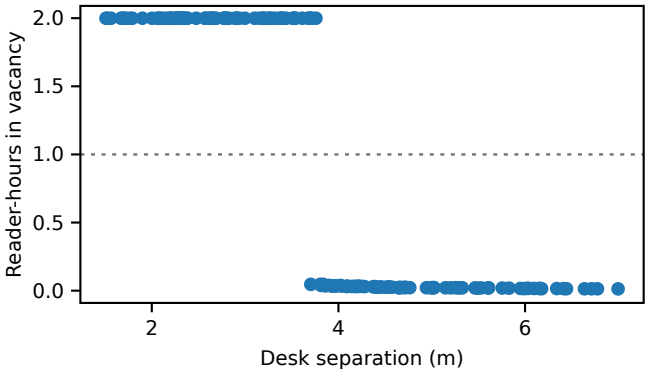


Fig. 1. Seat credit over one hour against desk separation, at 40 dB SPL with 60 s retention. Each point is an expectation over token sequences for one fixed placement. The reference line marks one occupied seat throughout the hour.

A single transmission test can conceal this effect. At 40 dB, 89.2% of A’s notices elicited a response from B and 89.7% passed in the reverse direction. Applying these spatial averages successively gives a probability of 0.00024 that an exchange is still issuing notices at 30 s. Keeping the desks at their installed positions gives 0.509 (Fig. 2). Successive evidence for the continued presence of a particular reader comes through the same pair of acoustic paths.

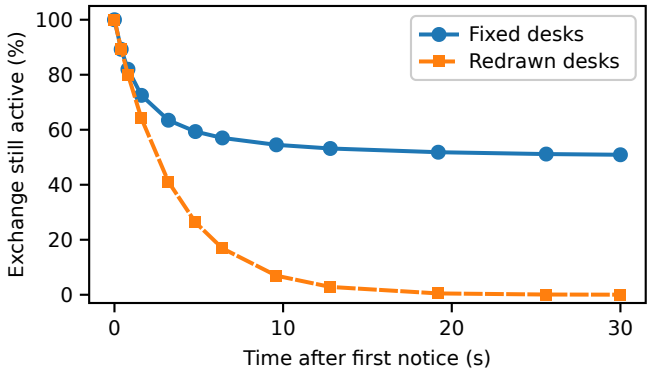


Fig. 2. Probability of a new notice at elapsed time following the exempt sound check. Fixed and independently redrawn placements have the same response probability for the first transmission.

B. Source attribution

Table II compares receivers at the same threshold. Direct-path subtraction reduced the mean attendance to 0.014 reader-hours, or 51 s of seat use in an hour. Every placement then had a token that could end the exchange. The residual credit includes the time allowed for a reader to finish a visit after the last notice. Both peer veto and exact subtraction cleared the attendance register. Of 32 local probes, 31 were audible at 40 dB in isolation. Exact subtraction preserved each of these responses during playback. The playback veto preserved none.

TABLE II
ONE HOUR IN A VACANT TWO-SEAT ROOM, AT 40 DB SPL AND 60 S RETENTION. THE FINAL COLUMN GIVES PAIRED DETECTION OF INDEPENDENTLY AUDIBLE LOCAL RUSTLES.

Receiver	Reader-hours	Page detection (%)
Level gate	1.013	10.5
Peer veto	0	0
Direct path	0.014	47.5
All paths	0	100
All paths, 1 sample late	1.9997	0

With the full subtraction filter one sample late, every placement recorded two readers after the first 0.8 s. The missing 0.0003 reader-hours are the initial delays before B and A issue their first notices. A one-sample error multiplies the magnitude of the playback residual by $2|\sin(\pi f/f_s)|$ relative to the incident component at frequency f. This exceeds unity in the upper part of the sampled band, where the differenced noise has most of its energy. The occupancy error therefore depends on the spectrum of the notice even though the attendance rule never uses frequency.

V. Implications for Seating Records

An empty room and the same room containing two acoustically silent readers have identical microphone inputs under the stipulated fixed paths. They must return the same attendance estimate, although their true occupancies differ by two. Cancellation can establish whether a local sound occurred during a notice. It cannot establish how long a reader remained after producing that sound. Conversely, a finite attendance interval cannot resolve the origin of a sound that keeps renewing it.

The exact one-reader limit belongs to the sampled set of placements and the two finite playback banks. A different spacing distribution would change the proportion of pairs able to sustain both records. Later reflections, ambient transients and moving bodies would also change the results. The simulated rustles test attribution at one overlap time and do not measure recognition of natural page turns.

A vacancy check for this estimator requires a reporting interval as well as a sound-level reading. Exempting the initiating test removes one entry. The attendance attributable to it includes the open interval at each seat and all subsequent renewals.

References

[1] L. M. Given and G. J. Leckie, “"Sweeping" the library: Mapping the social activity space of the public library,” *Library and Information Science Research*, vol. 25, no. 4, pp. 365–385, 2003.

[2] A. Jongman, R. Wayland, and S. Wong, “Acoustic characteristics of English fricatives,” *Journal of the Acoustical Society of America*, vol. 108, no. 3, pp. 1252–1263, 2000.

[3] J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” *Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[4] J. Benesty, D. R. Morgan, and J. H. Cho, “A new class of doubletalk detectors based on cross-correlation,” *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 2, pp. 168–172, 2000.

[5] G. Lafay, E. Benetos, and M. Lagrange, “Sound Event Detection in Synthetic Audio: Analysis of the DCASE 2016 Task Results,” in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2017, pp. 11–15.

[6] A. Mesaros, T. Heittola, and T. Virtanen, “Metrics for polyphonic sound event detection,” *Applied Sciences*, vol. 6, no. 6, pp. 162, 2016.