RESEARCH TRAJECTORIES IN CLASSROOM ACOUSTICS: INVESTIGATING CHILDREN PERCEPTION BEYOND TASK PERFORMANCE

Nicola Prodi* and Chiara Visentin†
Dipartimento di Ingegneria, Università di Ferrara, Ferrara, Italy

Résumé

L'acoustique des salles de classe a été un sujet de recherche pendant des décennies et suscite toujours l’intérêt des scientifiques en raison des multiples aspects impliqués, allant des problématiques autour de l’auditeur, comme les compétences cognitives, à la conception acoustique de la pièce ou l’influence du bruit ambiant sur l’orateur. Le professeur Murray Hodgson a été une personne clé dans ce domaine, et ses recherches ont été une source d’inspiration pour la communauté scientifique, et en particulier pour les auteurs de ce papier. Dans cette étude, il est d’abord discuté que les indicateurs acoustiques objectifs permettant d’obtenir une intelligibilité de la parole adéquate dans une salle de classe ne permettent pas de discriminer des conditions d’écoute perçues différemment par les sujets. Pour cette raison, une évaluation supplémentaire a été développée, basée le concept d’«effort d’écoute» issu de l’audiologie. La méthode est implémentée par une des mesures indirectes de l’effort d’écoute, qui est le temps de réponse au stimulus auditif. Le présent travail est un état de l’art des principaux résultats obtenus quand l’approche a été appliquée à plusieurs problèmes en lien avec l’acoustique des salles de classe.

Mots clefs : acoustique des salles de classe, bruit, intelligibilité de la parole, temps de réponse

Abstract

Classroom acoustics has been a topic of research for decades and still attracts the interest of scientists because of the many aspects that are involved, which range from listener-centered issues such as cognitive proficiency to the acoustical design of the room and also to the speaker’s voice alterations in noisy conditions. Prof. Murray Hodgson has been a key person in the field and his findings have inspired many others that followed, including the present authors. In this work it is firstly discussed that the objective acoustical indicators whose provision warrants suitable speech intelligibility in the classroom are not able to disentangle listening conditions which are perceived and rated differently by the listeners. For this reason an additional assessment has been developed based on the audiologically-oriented concept of “listening effort”. The method is implemented by means of one of the proxy measures of listening effort that is the response time to the auditory stimulus. The present work reviews the main results obtained when the approach was applied to several problems related to classroom acoustics.

Keywords: classroom acoustics, noise, speech intelligibility, response time

Foreword

During the year 2003 a short course on classroom acoustics was held at the University of Ferrara, Italy, where prof. Murray Hodgson was invited to present his findings. It was a special occasion to discuss open problems and share views with him. He was a leading expert in the field and the topic was having a revival in Italy. In fact many complaints had been raised on the inadequacy of the acoustical environment inside classrooms. In addition measurement campaigns had already confirmed the prevalence of poor acoustics in the large majority of the school buildings. In the discussion it was a shared view that research should have guided a process of improvement of classroom acoustics and it should have provided insight into the many aspects that make up the impact of acoustics on the learning process.

These were the first moves of a research trajectory that was clarified in the years to follow and that is still under development. The present work is primarily an account of the work done so far by means of a review of the main published works. Although the work does not add new results to the already disseminated literature, it is primarily intended as a resume of the path whose start dates back to the meeting with Prof. Murray Hodgson.

1 Introduction

The literature on the topic of classroom acoustics in the mid years 2000 was already rich and comprised several types of studies covering a large set of chronic and acute problems that can be roughly listed as:

A) Epidemiological studies on the effects of noise emitted by transport infrastructures on the academic achievements of the exposed students, with few longitudinal assessments;
B) Studies that investigated the link between internal classroom acoustics and verbal communication;
C) Studies on specific issues (data collection, simulation and assessment, interventions and acoustical treatments);
D) Studies on cognitive mechanisms, hearing impairment and various types of interactions.

In addition, also very accurate reviews were available that resumed the state-of-the-art knowledge [1, 2] (later on the more cognitive-oriented review of Klatte et al. [3] came up). Moreover at that time in North-America the interest in the topic was high and for instance the Acoustical Society of America (ASA) had revised its guidelines [4]. As a backing to these activities several studies investigated the limits of objective indicators in order to achieve optimal conditions (such as [5] and later [6]).

Assessing the acoustical conditions in the real classrooms of primary schools by measures with the latest technical means was deemed a good starting point also in the local Italian studies [7], together with the attempt to set acoustics in the context of the overall comfort in schools [8]. A first outcome was apparent: several noises could provide the same intelligibility experimentally, but they were perceived differently. Thus they had a peculiar impact of the communication process, and hence probably on learning. More generally, already in the literature [9] it was found that the same intelligibility could give raise to a broad subjective evaluation of listening difficulty. Objective qualification was not able to describe these facts even with speech indicators such as the Speech Transmission Index (STI) [10].

![Figure 1](from [7]): Measured speech intelligibility scores of pupils under several types of background noise tracked by the STI. A range of ambiguity starts from STI=0.5 on.

At that point the examination of the literature told that the limits of the acoustical qualification of classrooms were still serious. First, based only on the current acoustical indicators, the design of classrooms was not sufficiently refined to warrant at once high performance and comfortable listening conditions for pupils. In fact their intelligibility depended on the type of noise and, moreover, the threshold values defined for the acoustical parameters were fit for adults but not necessarily for children. Second, beside speech recognition also listening comprehension should be addressed in classrooms; however, this aspect was not under control of the conventional metrics. Third, other tasks besides linguistic ones may be impaired by the presence of adverse acoustic conditions, for instance the tasks related to reasoning and mathematics. No indication was available to control for that.

Despite several other unclear aspects had been raised in the literature, in this context the most urgent issue to address was deemed the first listed above. This involved improving the means of qualification and design of classrooms by going beyond the speech reception performance.

### 2 Speech reception beyond task performance: which direction?

Given the limits of the current objective indicators enforced in the standards it appeared that some important mechanisms underpinning speech reception could not be captured properly causing an unpredictable outcome, especially for students at disadvantage. Several studies [11, 12] replaced speech intelligibility with accuracy in task-specific tests and the effects on attainments were confirmed [13], but the methods were not easily adaptable to be used in the design and evaluation of remedies. So, an alternative solution should be searched for.

Cognitive psychologists discussed that listening «easily» at school is an essential pre-requisite to learning [14]. This happens because only when listening is not a difficult task there are cognitive resources available for further processing of the input information; those steps of learning requiring more cognitive load are thus facilitated. So, the problem could be theoretically turned back into a listening-related one, with a focus on the “ease” of listening and not just on the task performance. To develop this idea a basic approach could be borrowed from the field of audiology. In fact in that area the conditions of effortful listening were an increasingly urgent topic of research. A consensus definition of “listening effort (LE)” was outlined only much later [15] by resorting to psychophysics theoretical models of capacity and attention. LE depends on input demands, capacity and on their interplay, and it is influenced by unconscious and intentional attention. Moreover, LE is a complex and multi-faceted construct which is not possible to uniquely quantify with a single measure.

Several methods to grasp aspects of LE were already available. They could be divided into three broad typologies: subjective (e.g., self-ratings of effort), physiological (e.g., pupillometry, skin conductance, saliva cortisol etc...), behavioral (e.g., dual-tasks) methods [16, 17]. In particular one of the oldest behavioral quantities proposed in the literature is the response time (RT) [18]. This measure can be implemented in dual-task or in single task experimental paradigms. In the latter case, the RT to the auditory stimulus is used: it is defined as the time elapsed from the end of the presentation of the stimulus to the response given by subject, either verbally or manually. RT is not an estimate of LE, but it is a measure of speed of processing. By construction, it is assumed that an increase
of RT is associated to an increase of LE. A fairly large number of studies used it with profit in auditory experiments since it is relatively easy to collect, but data need careful handling to deal with inter and intra-subject variabilities.

In the above process of development the first step was thus to implement a mixed measurement process where the task was one (speech reception) and the measures retrieved were two, performance and response time.

3 First application in a virtualized classroom

Binaural impulse response measurements were taken inside two identical classrooms (V=250m³) which were close one another. One of them was treated with sound absorbing ceiling tiles while the other was not [19]. The rooms had both 24 desks with chairs for pupils and a bigger desk with chair for the teacher. They had a large window on one lateral wall. Their furniture consisted in two closets and few posters. Background noise was also sampled during the acoustical measurements and consisted in activity noise from occupants and babble noise from the adjacent corridor. Various combinations of reverberated target signal and noise were rendered under controlled conditions (unoccupied untreated, A; occupied untreated, B; occupied treated, C). The rendered sound fields had signal-to-noise ratios equal to 0, 6 and 12 dB. Diagnostic Rhyme Tests were proposed to 80 normal hearing pupils from III, IV e V grades (8-10 years) and to 42 normal hearing adults.

Analysis of variance (ANOVA) was used for the statistical analysis; the p value at the level of 0.05 was the statistical tool to test the significance of the effects. When ANOVA was applied to the speech intelligibility scores (IS) there was a non-significant difference between the grades (p>0.05) and a slight significant difference (p=0.046) between all pupils and adults. Moreover only condition C was disentangled statistically, but IS could not differentiate the other comparisons. When RT was analyzed two facts become evident: the indicator was very correlated with IS but it was able to resolve some of the previous limits. In fact at equal IS there were significant differences between grades III and V (p<0.001), III and IV (p=0.02), IV and V (p=0.034) and between all pupils and adults (p=0.039) (Fig. 2).

Therefore, it was decided to combine the two metrics into a single indicator being their ratio. This responded to two instances: first, making the method of assessment more compact and second, framing the method into the psychophysics concept of cost. In fact, in the area of psychophysics it is not unusual to define the ratio of a performance measure to the time needed to achieve it as the “cost” or “efficiency” of the related process [20]. By using this analogy the ratio of IS versus RT, termed “listening efficiency (DE)”, was by definition the number of items correctly recognized per second (units [s⁻¹]).

The quantity was employed to analyze the same data set and provided further insight into the bias of performance between pupils and adults. In particular it was found that the room-acoustical criteria developed for adults may cause a severe underperformance of pupils, which could be quantified as equal to nearly 0.1 STI units, or 0.2 s⁻¹ in listening efficiency terms. Said it more practically, in the above experiment the gap would be equivalent to a loss of an additional entire item (a disyllabic word in this case) every five seconds employed in the speech recognition process.

4 Studies in real classrooms

Later on studies considered real classrooms. A first study had the twofold aim of ranking harmful noises in the classroom and of better detailing the performance with age. The work involved as much as 741 pupils distributed over 47 classes at six primary schools. A test bench was developed so that an entire class could be tested at once inside their classroom. The target signal and the noises were played back through separate loudspeakers. Three types of noises were included, that is “babble and activity” – A, “tapping” – Tp and “traffic” – Tr. As in the previous studies a speech recognition test was used. The signal-to-noise ratios were fixed at 0, 6 and 12 dB and, depending on the noise type, the level of the target signal was set at 60, 66 or 72 dBA at 1 m in front of the loudspeaker. Each pupil was equipped with a touchscreen smartphone to collect responses. The test lasted approximately 45 min. A dedicated statistical procedure was set up to compare the noises. It was based on the aggregation of the data into separate strata corresponding to the rating intervals of the STI.

While taken separately, both IS and RT could not achieve a fully unambiguous ranking of the noises. On the contrary, when using DE the full ordering of the noises was statistically significant. The result was driven by the better listening conditions; it was found that the A noise was the less efficient, followed by Tp and Tr in succession (Fig. 3 and Tab. 1).

Within the same study, it became evident a developmental effect on RT, since older pupils were always faster to respond than the youngest pupils, even in quiet conditions.
Such age proficiency was controlled for and finally it could entirely account statistically for the age effect found in the tests.

Data so far always considered the mean values of the quantities over a set of trials that span along the test duration. Such analysis could not depict the eventual changes of performance during a real lesson of approximately 45 minutes. On the other hand, hints were given that performance could have not been stable during the lesson. In fact, during private communications, often teachers outlined a loss of performance during lessons due to a continuous noise.

The subset of data for 8-10 years old pupils was used in [22] to address this issue. It was found that during a lesson period only older pupils’ performance changed. In particular, when the acoustical conditions were more favorable, they suffered a decrement of performance.

This loss occurring in more favorable acoustics could not be firmly ascribed to eventual “fatigue” occurring to children during the lesson. In fact the construct of “fatigue”, which differs from LE, needs dedicated evaluation methods as pointed out in specific studies [23, 24], and RT measures alone are not entirely appropriate for the scope.

Anyhow, the changes of performance during a lesson period highlighted by the study told that LE could be modulated by noise and age and thus needed confirmation with a dedicated data set, and for younger pupils too.

For this reason later on a specific investigation [25] addressed the impact of noise over the lesson period by splitting it into two equal parts and by running the same tests twice (Fig. 4). A quiet condition and one with a speech shaped stationary noise were used. Pupils aged 5 to 7 years took part in the experiments which consisted in With Picture Identification Tests (5 years) and Diagnostic Rhyme Tests (6 and 7 years). The tests were conducted inside their classrooms with the same apparatus as in prior studies.

Although the results for the 5 years old pupils were mixed, those for the 6 and 7 years old were clear and consistent between the two groups.

The IS showed a main effect of noise ($p<0.001$) but nothing more. On the contrary, RT showed a significant effect of noise ($p<0.001$), an effect of test repetition ($p<0.001$) and an interaction of the two ($p=0.002$). In Fig. 5 the results for the 7 years old pupils are shown. The meaning is that the effect of noise is relevant in the repetition only, and that it happens while the IS are still unaffected.

### 5 Using response time in the acoustical design

The studies so far proved that RT is sensitive to several relevant issues of speech perception in rooms for younger pupils. Then one of the original questions could be raised: is it possible to employ the concepts above during acoustical design? Said it differently, is such indicator also sensitive to changes of the room acoustics, such as shape and properties of materials? Should this be the case, RT would have potentials in directing the acoustical design towards less effortful listening and hence to a more learning-oriented
one. This was the main question of a successive study [26] that employed a mix of field and auralized experiments. In this case university students of Italian and German mother tongue were recruited.

They performed tests first in a real classroom which was a box-shaped room with a volume of 197 m³. It had flat surfaces apart the lateral partition with the adjacent corridor which was acoustically treated with a sound absorbing paneling. The classroom was furnished with wooden desks and chairs and hosted a maximum of 25 students. Then an auralized version of the same space was developed where either the sound treatment was removed or the volume was doubled. As always done in similar experimental designs, in order to minimize learning effects and to familiarize the participants with the test procedure a set of trials was proposed before the test in the real classroom took place.

RT was capable of detecting the changes, while IS were largely unsuccessful (Fig. 6). Moreover, RT proved more reliable than a subjective assessment, mainly because the anchors that the two groups employed were likely to differ in a relevant manner.

In the same study the metric RT was validated between field and auralized tests. This step allowed taking RT in the virtual field as an indicator of effort during listening in the corresponding real space.

6 Addressing second language learners

A successive investigation [27] was accomplished at UBC Vancouver with the direct participation of Prof. Hodgson. The study regarded the so-called English second language students (L2) that is those who went to Canada from non-English speaking Countries. In particular impulse responses in a real classroom were used to create reverberated conditions and recorded noises were added at various signal-to-noise ratios. The tests took place in an anechoic chamber and the panel included both mother tongue L1 (13) and L2 (24) listeners. Given the L2 disadvantage already reported in the literature for worse acoustical conditions, the sound fields were set to provide near ceiling IS (i.e. close to 100%).

The first condition was anechoic; the other two shared the same value of STI=0.57 but the former had reverberation only, while the latter had some added noise.

The results of IS confirmed no significant effect neither of condition nor of groups and no interaction between the two variables. On the other hand, when RT was analyzed (Fig. 7) there were two main effects of listeners’ group ($p=0.02$) and listening condition ($p<0.001$) but there was no interaction between the two variables.

The data showed that L2 were always at a disadvantage compared to their L1 peers. The absence of interactions witnessed that such gap between the groups did not depend on room acoustics but on more fundamental internal processing.

7 Concluding remarks and future directions

The research trajectory that was exposed in this work took its first moves in the early years 2000 and thanks to the suggestions and discussions that on many occasions the authors had with Prof. Hodgson.

The data accumulated so far in diverse scenarios and with different panels of users do confirm that RT is a sensitive quantity that can output information not accessible with accuracy measures. In particular environmental variables such as the type of noise and room acoustics have been reflected in the results. Several additional variables can be included for instance by altering the characteristics of the speaker, those of the listener and finally the transmission path between the two. It is believed that the method could display a benefit when applied to numerous practical applications encompassing both the perception of speech in room acoustics and the audiological fields where it was firstly introduced. In agreement with the initial intentions, one of the most relevant fields of application is the design of rooms for speech in order to provide conditions where listeners can be highly efficient in the listening task.

![Figure 6](image1.png)  
*Figure 6 (from [26]): Values of RT obtained for native Italian and native German students. Front and back positions in a virtual classroom with and without sound absorbing treatment.*

![Figure 7](image2.png)  
*Figure 7 (from [27]): Values of RT obtained for L1 and L2 across the three acoustical conditions. A: anechoic; B: reverberated only; C: reverberated with noise.*
References


It's that quiet.

Cut down on noise and acoustic interference with Soundproof windows.

www.bquiet.ca
1.877.475.9111
NOISE MONITORING SOLUTIONS | Models 831C & NMS044

SoundAdvisor from Larson Davis sets a new standard for connectivity, access, and control of your noise monitoring using a network connection:

- Connect over cellular, WiFi, or wired networks
- Control meter and view data via web browser
- Receive real-time alerts on your mobile device
- Monitor continuously with a solar powered outdoor system