# Auditory compensation for reverberation in normal hearing listeners

JENS BO NIELSEN AND TORSTEN DAU

*Centre for Applied Hearing Research, Ørsted•DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark*

The present study investigates the ability of the auditory system to compensate for the immediate effect of reverberation on speech intelligibility. In the experiments, the spectro-temporal properties of a speech signal are modified by reverberation that smears the signal. Reverberation can be viewed as a natural lowpass modulation filter and generally reduces speech intelligibility. However, recent research has shown that the auditory system appears to compensate for this effect by taking the reverberant environment into account (Watkins, 2005). This compensation mechanism is functional after a short adaptation period (less than 1 s) and enhances the ability of the auditory system to detect smeared speech modulations. In the present study, the compensation mechanism was investigated in a simulated room with a high degree of reverberation. Two experimental setups were used in order to verify the existence of compensation for reverberation, but the results of the experiments could not confirm the earlier findings of Watkins (2005).

## INTRODUCTION

It is well-known that the auditory system performs both a frequency analysis and a temporal analysis of the speech signal. The temporal analysis is crucial not only for parsing the speech stream into units (words, syllables etc.), but also for identifying the amplitude modulations that are the main cue to stop-consonants such as [t] or [p]. Anything that reduces a listener's ability to identify such modulations may lead to deterioration in speech intelligibility (Houtgast and Steeneken, 1985). Reverberation smears the speech signal and generally reduces speech intelligibility, but the auditory system can to some extend compensate for the modulation smearing by taking a reverberant environment into account (Watkins, 2005). Watkins showed that a listener's ability to identify the stop-consonant [t] in the word 'stir' decreases when the word is played in surroundings with increasing levels of reverberation. However, if the word is embedded in a carrier sentence with similar reverberation the detectability of the [t] is partly restored. Watkins refers to this effect as *extrinsic compensation*, because the intelligibility of the word is affected by the carrier, i.e., the intelligibility depends on information about the reverberation collected from a speech signal 'outside' the word itself. Watkins also describes the compensation as *perceptual*, because it is assumed to originate in an increased ability of the listener to *perceptually* differentiate between the direct sound and the reverberation.

In the first part of the present study, the intention was to create a test setup that supports the existence of extrinsic compensation in normal hearing listeners. The idea was

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.). ISBN: 87-990013-1-4. Print: Centertryk A/S.

then to use the setup to study this effect in hearing-impaired listeners. In particular, it would be interesting to investigate whether certain perceptual deficits of hearing-impaired listeners can be related to a lack of such reverberation compensation. These investigations were, however, not conducted since the compensation mechanism for normal hearing listeners could not be replicated in the present study.

## ASSESSING THE EFFECT OF EXTRINSIC COMPENSATION

In the study, two words that only differ by a stop-consonant were used to assess the impact of reverberation. Stop-consonants are sensitive to reverberation, because a main cue to their identification is a pronounced amplitude modulation dip that is highly receptive to smearing by reflections (Drullman et *et al.*, 1994). From the Danish word pair 'sok-stok', a continuum of 11 words, changing in steps from plain 'sok' (step 0) to plain 'stok' (step 10), was created, as shown in Fig. 1. The words were generated by determining the envelope of both ‘sok’ and ‘stok’ and imposing different ratios of these envelopes on the waveform of ‘sok’. Whether a word from the continuum is heard as ‘sok’ or ‘stok’ depends on listener bias, on the amount of reverberation added, and, as shown by Watkins (2005), on the surrounding carrier.

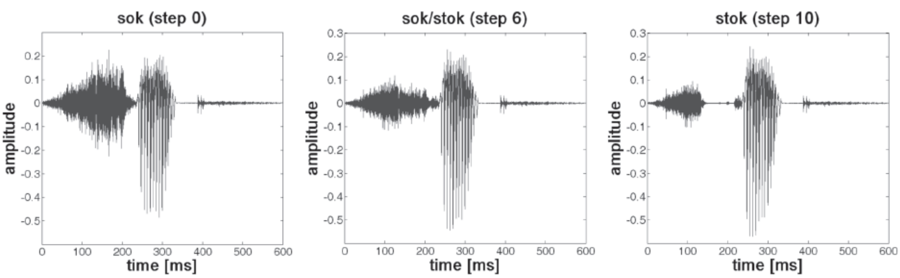


**Fig. 1**: Three instances of the ‘sok-stok’ continuum. Left: The original waveform of ‘sok’. Middle: The waveform of ‘sok’ imposed with a ratio of the ‘sok’ and ‘stok’ envelopes. Right: The waveform of ‘sok’ combined with the plain envelope of ‘stok’.

A listener switches from identifying the target words as 'sok' to 'stok' at what is designated the *decision boundary*. When reverberation is added to the word, the stop-consonant [t] is smeared, and more words of the continuum will be heard as 'sok‘. As a result, the decision boundary will move towards ‘stok’ (step 10). If now the perceptual ability of the listener to separate the direct speech signal from the smearing reverberation is improved, more words will be heard as ‘stok’ and the decision boundary will move back towards ‘sok’ (step 0). Hence, the decision boundary can be used as a measure of a listener’s ability to detect speech modulations that have been smeared by reverberation. Specifically, it can be used to measure whether the listener’s ability to identify the [t] in ‘stok’ improves when the test word is embedded in a reverberant carrier sentence.

The carrier sentence was chosen to be: "Læs nu den korte tekst" *target word* ”og tryk på den." ("Now read the short text ... and press it.") A target word from the 'sok-stok' continuum was embedded in the sentence. The target word and the carrier sentence were then manipulated by reverberation in three different ways and the decision boundary was determined:

1. Reverberation was not added to the target word nor to the carrier sentence.
2. Reverberation was added to the target word. This should make more of the continuum words sound like 'sok' because the [t] is smeared. As a result, the decision boundary moves towards 'stok'.
3. The same reverberation was added to both the target word and the carrier sentence. Reverberation on the carrier sentence should activate the extrinsic compensation, and facilitate the identification of the [t] in 'stok'. More target words should now sound like 'stok' again and the decision boundary move back towards 'sok'.

The expected shift of the decision boundary towards ‘sok’ when reverberation is added to the carrier also (i.e., the boundary position difference between situation 2 and 3) would be an indication of extrinsic compensation for reverberation.

## EXPERIMENTAL PROCEDURE

Two experimental setups were designed to investigate extrinsic compensation: 1) A method of constant stimuli, where each step of the 'sok-stok' continuum was presented five times to the listener in random order. This experiment followed Watkins’ original experimental design and determined the decision boundary. 2) A two interval, alternative forced-choice (AFC) method in combination with an adaptive procedure (1-up 3-down). This experimental design did not determine the decision boundary, but measured the listener’s ability to identify a step distance between words from the continuum.

In the constant stimuli experiment (1), the listener’s task was to decide whether he/she perceived the continuum word as 'sok' or 'stok'. Based on the responses, a psychometric function could be derived, as shown in Fig. 2. At the decision boundary, there is a transition from ‘sok’ to ‘stok’ responses; the boundary is defined as the crossing with the 50% line (not shown). This experimental method requires that the listener selects between two options (‘sok’ or ‘stok’) of which none is the ‘correct’ choice. The listener must establish a decision criterion, and the responses are therefore affected by bias.

The extrinsic compensation effect predicts more ‘stok’ responses with reverberation added to the whole sentence, compared to reverberation on the target word only. Does this reflect a *subjective tendency* of the subject to shift towards ‘stok’, or does it reflect an improved *ability* to detect the [t] in ‘stok’? In other words, does the carrier reverberation facilitate the detection of the [t] or does it change the bias of the listener? To answer this question, experiment 2 was set up with the AFC method. The carrier sentence was the same as in the constant stimuli experiment, but the target word was replaced by *two* words that were presented right after one another: A fixed base word, which was chosen to be step 2 from the ‘sok/stok’-continuum, and a variable target word starting at step 10 of the continuum, and changing according to the listener's responses. The two words were presented in random order, and the listener was asked to select the interval that sounded *most* as 'stok'. A 1-up 3-down algorithm was applied. This adaptive procedure balances around the step number in the continuum, where the listener can distinguish the base word (step 2) and the variable target word with a likelihood of 59%.

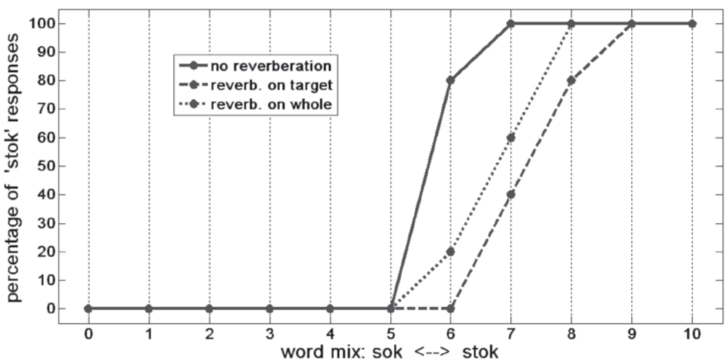


**Fig. 2**: Example of a result of one subject from the constant stimuli experiment. The listener responds 'sok' or 'stok'. Each stimulus was repeated 5 times, so response-percentages are in steps of 20%. The subject tends to respond 'sok' when reverberation is only added to the target word. When reverberation is added to the whole sentence (carrier and target word), the subject tends to respond 'stok' more frequently, as predicted by the extrinsic compensation effect.

## ADDITION OF REVERBERATION

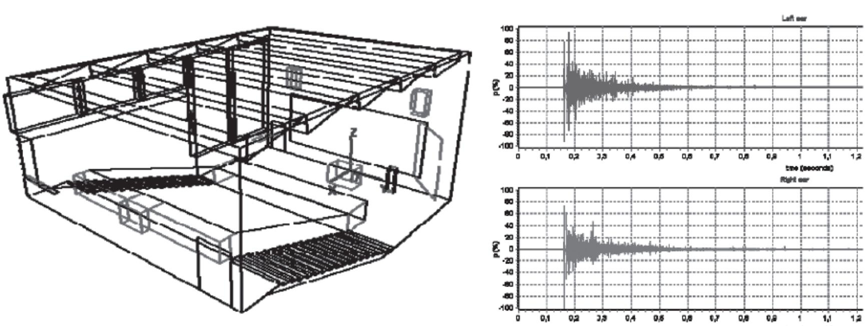

**Fig. 3**: The ODEON model of the auditorium and the corresponding BRIR. The room was modeled with no audience, which resulted in a reverberation time of 1.9 s at 2 kHz.

Reverberation was added to the target word and carrier sentence by convolving speech recordings from an anechoic chamber with impulse responses from computer simulations of real rooms. The room acoustics software ODEON was used to create binaural room impulse responses (BRIRs) of some specific locations at the Technical University of Denmark, e.g., a classroom, a stairway, a hallway and an unoccupied auditorium (with 150 seats). The reverberation in the auditorium was relatively high with an early decay time (EDT) of 1.9 s at 2 kHz, resulting in an effective smearing of the modulation from the [t] in ‘stok’ (Fig. 4). This BRIR was chosen for all experiments.

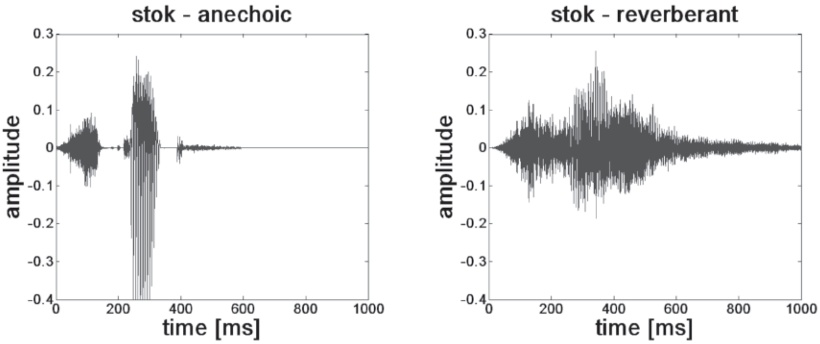


**Fig. 4**: The anechoic version of ‘stok’ (continuum step 10, see Fig. 1) and the same word convoluted with the BRIR of the auditorium. The [t] (at approx. 150 to 230 ms from word onset) is smeared, but still audible, because the [t] is very pronounced in the plain ‘stok’ word.

## RESULTS

The constant stimuli experiment was conducted with 6 normal hearing test listeners. The listeners were presented with 1) stimuli without reverberation, 2) stimuli with reverberation added to the target word only, and 3) stimuli with reverberation added to the whole sentence. Each step of the continuum was presented 5 times, resulting in 5 rep. x 11 steps x 3 conditions = 165 presentations. Normal distribution curves were fitted to the results; two examples are shown in Fig. 5. For each curve, the crossing with the 50% line was considered and a 2-way ANOVA was conducted, with the subjects and the three reverberation-conditions as factors. There was a significant difference between subjects [$F(5,2) = 5.4$, $p = 0.01$], but no significant difference between the reverberation conditions [$F(5,2) = 0.15$, $p = 0.86$]. Hence, the existence of extrinsic compensation was not confirmed in this experiment.

The AFC experiment was conducted with five normal hearing listeners; two typical test runs are shown in Fig. 6. A statistical analysis was not performed. Figure 6 shows the number of continuum steps that the subjects needed to hear a difference between the base word (horizontal line at step 2) and the target word. An extrinsic compensation effect should result in a function for 'whole sentence reverberation' that lies closer to the base line than the function for 'target word reverberation'. Such an effect was not found in the present experiment. Thus, the existence of extrinsic compensation could not be confirmed in this experiment either.

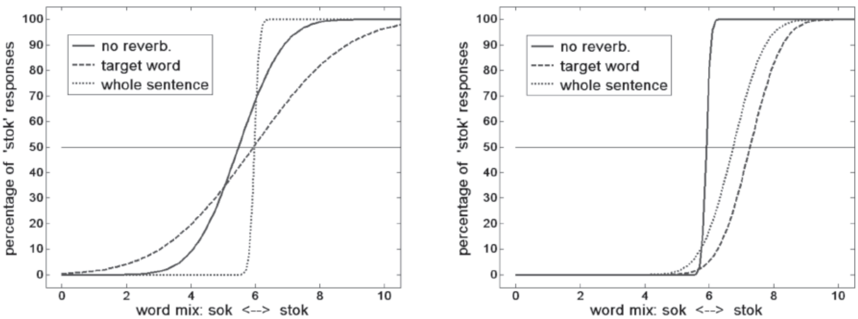


**Fig. 5**: The results of the constant stimuli experiment for two subjects. The fitted curves show at which step the listener's response shifts from 'sok' to 'stok'. The further to the left a curve lies, the easier it has been to detect the stop-consonant [t]. The crossing with the 50% line is defined as the decision boundary and used for a 2-way ANOVA analysis.

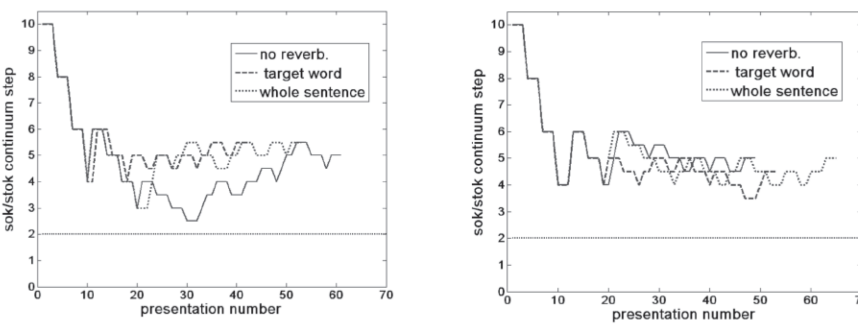


**Fig. 6**: The result of the AFC experiment shown for two normal hearing listeners. The curves converge at the step where the listener can just distinguish the variable word from the base word. A curve close to the base line indicates that the listener is able to detect small changes in the modulation stemming from the [t] in the continuum words.

## DISCUSSION

Unfortunately the results of the present study do not support the findings of Watkins (2005); none of the experiments have confirmed the existence of an extrinsic compensation mechanism. A reverberant carrier sentence did not improve the ability of the listeners to identify the [t] in the target word (compared to the situation with reverberation affecting the target word only). More surprisingly, the addition of reverberation to the target word only did not have a significant negative effect on the listeners' abilities to detect the [t] (compared to the situation with no reverberation at all). A too short reverberation time cannot explain these results, as the BRIR of an unoccupied auditorium was used (early decay time was 1.9 s). It is possible that differences in the recordings (of target words and carriers) are responsible for the deviations from the results in the study by Watkins (2005). More investigations are needed to explain these deviations.

## REFERENCES

Drullman, R., Festen, J. M., and Plomp, R. (**1994**). “Effect of temporal envelope smearing on speech reception," J. Acoust. Soc. Am. **95**, 1053-64.

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.

Watkins, A. J. (**2005**). “Perceptual compensation for effects of reverberation in speech identification,” J. Acoust. Soc. Am. **118**, 249-262.