On the Effect of an Expected Auditory Task on the Performance of an Ongoing Task

Ewald Strasser[1], Christiane Thiel, Steven van de Par[1]

[1] *Acoustics, Department of Medical Physics and Acoustics, Carl von Ossietzky University,*
E-Mail: Ewald.Strasser@uni-oldenburg.de

[2] *Biological Psychology, Department for Psychology, Carl von Ossietzky University*

Introduction

Imagine the following situation: You are in a busy restaurant and you are talking to your companion while monitoring your surrounding for the waiter who brings your meal.
The question arises if your attention will be divided between listening/talking to your companion and monitoring your surroundings for the waiter who brings your meal. If this is the case, another important question is whether your attention will it be divided from the moment you ordered on or after a certain time when it becomes more likely that your order will arrive?
From studies of divided attention it is known that listeners are capable of receiving information from two different speakers simultaneously [1]. It is also known that there are costs associated with this ability. Studies on divided attention have shown that listening effort increases when two listening tasks are employed at the same time (see for example [2]). Despite the extensive literature available on divided attention and dual task performance, very little is known about how much separation time between two tasks is needed so that they no longer compete with one another for attention. It is likely that, to a certain degree, simply the anticipation of a second task has an impact on the listeners response to the first task.
We argue that hearing-impaired and/or elderly people, who deploy more listening effort (see for example [3] and [4]), suffer a disproportional greater listening effort compared to normal hearing and/or young people. Elderly and/or hearing-impaired people would have to strain their attention disproportionally when they are in challenging listening situations with an anticipated sparse task.

Method

The experiment was done in a listening boot. For the presentation of the auditory stimuli we used a single loudspeaker (Genelec 1029A) placed in front of the participant. For the presentation of visual stimuli we use a Dell P2314T computer screen.
Because we wanted to get a better understanding of how an anticipated word recognition task influences an ongoing task 1-Back task, we looked at different dual task procedures as a research paradigm for us.
The majority of dual task studies use a concurrent experimental design (compare [5]), where both stimuli are presented at the same time. Such a paradigm obscures the potential time differences we are interested in.
To tackle this problem, we choose a sequential dual task paradigm. Obviously, the paradigm consists of two tasks, which are described in the following section.

1-Back Task

The first task is a 1-Back task. In a sequence of presented numbers, participants had to decide whether a number that was presented was either equal or different to the number that was presented directly before. For example, when the number that was presented was 7, then participants had to remember if the number that was presented before the 7 was also a 7, or another number. If the numbers were the same, participants should press the left arrow key, if they are different, they should press the right arrow key (see Figure 1). Each 1-Back block consisted of 8 to 12 trials and the answer needed to be given within 1,2 seconds after stimulus onset.

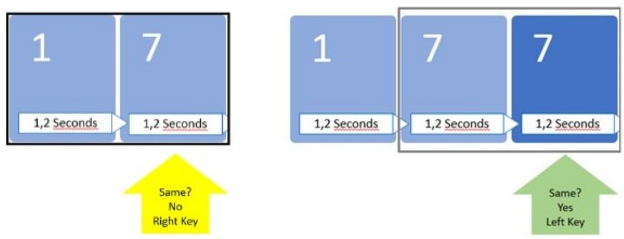

Figure 1: Sequential presentation of 1-Back trials. If the actual 1-Back trial is the same as presented before the left arrow key needs to be pressed otherwise the right arrow key.

To ensure a certain commitment in the task, participants got feedback on their performance. Below 75% correct responses at a single 1-Back block participants were informed that they should try to do better. Therefore, correct responses are not a very useful measure of the performance of participants. Instead we will use reaction time to quantify effort of participants.

In the analysis we used trials where only the 1-Back task was done as a baseline.

Auditory and Visual Presentation of 1-Back

The 1-Back task was presented auditorily and visually. This was done because crossing modalities within a dual task is likely to have an impact on the performance of the tasks (see [6] for an overview). Auditory stimuli (digits) were between 570-810ms long and visual stimuli were presented for 500ms.
The sound level of the digit words was adjusted by the participants before they started the experiment. They were instructed to adjust the level, to hear the digits pleasantly loud. This was done to exclude the factor intelligibility of the 1-Back words as an influence on performance.

Word recognition task

The second task was a word recognition task. We used 10 words from the WaKo monosyllable rhyme test [7]. Criterion for choosing words was that they could clearly be put in the categories ‘animate’ and ‘inanimate’. After pilot testing we decided to use the words ‘chicken’, ‘louse’, ‘cattle’, ‘animal’, ‘women’ for the category animate and ‘courtyard’, ‘hole’, ‘pipe’, ‘gate’, ‘wind’ for the category inanimate. **Table *1*** shows all words categorized and with their German translation.

Table 1: Words that were used for the word recognition task. The translation is in parentheses.

Belebt (Animate)	**Unbelebt (Inanimate)**
Huhn (chicken)	Hof (courtyard)
Laus (chicken)	Loch (hole)
Rind (cattle)	Rohr (pipe)
Tier (animal)	Tor (gate)
Weib (women)	Wind (wind)

To address the differences in intelligibility of these words and for differences in the hearing ability of our participants, we measured an SRT for each word for each subject before the actual test started. At the actual test we measured for each of these words and each participant separately a speech perception threshold, before the actual dual task started. In the actual test words were presented at -2, 0, 2, 4, 6, 8 dB compared to the individual SRT. These level differences were chosen to make the task challenging for our participants. This is important because the anticipatory effect will probably be greater when participants spend a lot of cognitive energy on it.

Dual Task

For the dual task, the 1-Back and the word recognition task was combined (Figure *2*). Each dual task block consisted of 8-12 1-Back trials which were followed by a word recognition task. The variation in the number of 1-Back trials was done to make it less predictable when the word recognition task would occur. The jitter should also make the anticipation effect more pronounced.

Figure 2: Dual Task setting. After 8-12 1-Back trials the word recognition task was presented.

We were interested whether dual task blocks had different response times for the 1-Back trials than the 1-Back trials alone.

Speech Shaped Noise

Each condition was accompanied by speech shaped noise which was generated from all the speech material from the WaKo [7]. This was done to make the listening situation more challenging.

Hearing Status

To see whether hearing-impaired people have a disproportional disadvantage when they are in a dual task scenario, we had to recruit hearing impaired and compared them with a normal hearing control group of approximately the same age.
At this stage of the study we gathered data from ten normal hearing participants and five hearing impaired participants.
Hearing impaired people were recruited with the help of the “Hörzentrum Oldenburg”. Participants had to have a PTA-4 of below 40dB. Additionally, we defined borders for band specific pure tone hearing which can be found in

Table 2. Normal hearing participants had to have no greater hearing loss than 30 dB of each frequencies between 125 Hz and 3000 Hz. Additionally, the mean of the frequencies between 2000 Hz and 8000 Hz should not show a hearing loss greater than 30 dB.

Table 2: Borders for frequency specific hearing deficits for hearing impaired and normal hearing. Frequency indicates the pure tone of interest. The columns ‘Hearing Impaired’ and ‘Normal Hearing’ indicate which values were acceptable to be part of the regarding group. ‘None’ indicates that there was no indication for this frequency in the regarding group.

Frequency	Hearing Impaired	Normal Hearing
125 Hz	0 – 40 dB	>30 dB
250 Hz	0 – 40 dB	>30 dB
500 Hz	0 – 40 dB	>30 dB
1000 Hz	0 – 60 dB	>30 dB
1500 HZ	none	>30 dB
2000 Hz	20 – 60 dB	>30 dB
3000 Hz	none	>30 dB
4000 Hz	40 – 70 dB	none
8000 Hz	40 – 130 dB	none

All participants had to be within an age range of 50 to 75 years old.

Results

The average age of the normal hearing participants was ~63 years and they had a PTA-4 of 22,5 dB. The hearing-impaired participants had an average age of ~68 years had a PTA-4 of 47,5 dB.

We used a generalized linear mixed-effects model with the following variables as predictors:

- Dual Task: 1-Back task v.s dual task
- Position: of 1-Back trial

- Status: hearing status
- Modality: auditory v.s visual presentation of 1-Back task

Dual Task (1-Back task only v.s 1-Back task followed by word recognition task) codes whether a 1-Back trial was done in a run where the word recognition task was presented or not.
Position (position of 1-Back trial) codes at which position each 1-Back trial was presented within the 1-Back block. Only the positions 2 to 8 were considered. Position 1 was not considered because it had no predecessor and therefore no obvious valid answer could be given. Positions greater than 8 were not considered because at position 9 was earliest possible onset for the word recognition task. This means that till position 8 every trial was always a 1-Back trial.
Status (hearing status) codes whether participants were in the group hearing impaired or normal hearing.
Modality (auditory v.s visual presentation of 1-Back trials) codes whether in a whole run the 1-Back block was presented auditory or visually.
We used a generalized linear mixed-effects model approach to run a random intercept model in Matlab [8]. We modelled random intercepts for Position nested within participants and nested within Status.

Dual Task, Position, Status and Modality were used as fixed effects.
The fixed-effects were modelled as a full factorial design. We found significant main effects for Dual Task ($p < .001$, $b = 9.4$, $SE = 2.4$), for Position ($p < .001$, $b = -15.8$, $SE = 4$) and Modality ($p < .001$, $b = 135$, $SE = 1.3$). Also, the intercept was significant ($p < .001$, $b = 678$, $SE = 15.2$).

An illustration of the effect of the Dual Task condition can be found in **Figure *3***. It shows that 1-Back only tasks have a slight, but significant, advantage in response time.

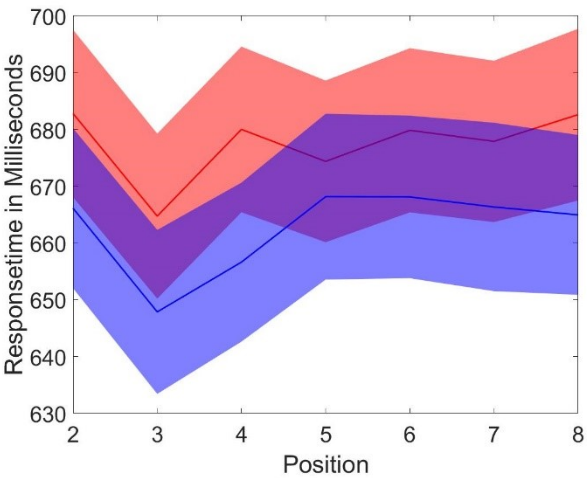

Figure 3: Response time for 1-Back trials in milliseconds. The solid lines indicate the mean at each position of the 1-Back trial and the tubes indicate the standard error. Red is the condition with the word recognition task and blue shows the 1-Back only trials.

Figure 4 shows the main effect for the position at which the 1-Back trial was presented. The model found that position 3 differs significant from the rest of the position. This can clearly be seen but more interesting is the tendency that form position 3 on we see a constant rise in response time. This would very well fit with the theory that the anticipated task draws cognitive energy form the ongoing task. One might rightfully say that also 1-Back only blocks contribute to the main effect of Position but it has to be kept in mind that also at the end of a 1-Back only block a task switch happened, namely to the next 1-Back only block. However, the effect should be much weaker in the 1-Back only block because this switch should not be as cognitive demanding. As we will see this prediction is very much in line with our data.

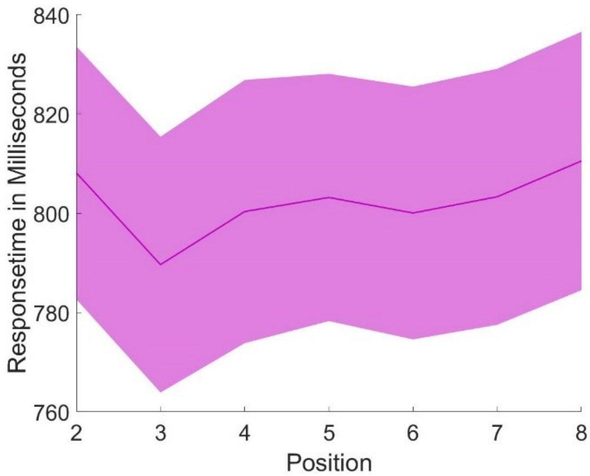

Figure 4: Response time for 1-Back trials in milliseconds. The solid line indicates the mean at each position of the 1-Back trial and the tube indicates the standard error. The tube includes all conditions.

In Figure 5, we see all data split for the visual and auditory condition. It is obvious that the visual condition has a huge advantage over the audio condition in the means of response time. This effect is simply due to the general differences in the perception process of audio and visual stimuli. Whereas a single digit word that is presented visually can be processed immediately after onset, one must listen at least for some time to understand a single digit word that is presented auditorily.

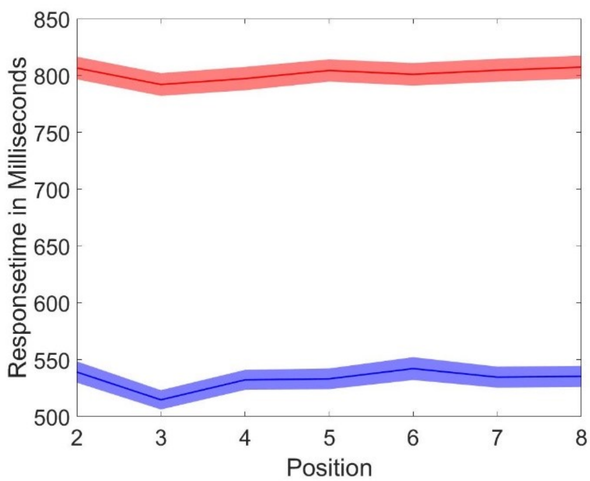

Figure 5: Response time for 1-Back trials in milliseconds. The solid lines indicate the mean at each position of the 1-Back trial and the tubes indicate the standard error. Red is the auditory condition and blue is the visual condition.

The interaction between Modality and Dual Task ($p < .01$, $b = -4.2$, $SE = 1.3$) as well as the interaction of Modality with Dual Task with Status ($p < .001$, $b = -5.4$, $SE = 1.3$) became also significant. Since the three-way interaction (of Modality with Dual Task with Status) is more solid in terms of p and the effect it has on reaction time we will focus at this effect.
As we can see in Figure 6, in the audio condition (upper tubes at the left figure), hearing impaired subjects show a higher response time in the 1-Back condition with the word recognition task condition than the 1-Back only condition. Compared to all other conditions there are large areas that do not overlap. This means that hearing impaired people spend way more response time in situations with a sparse task, especially when the ongoing task is also an auditory task.
For all the other conditions we see a high overlap between the 1-Back only and the 1-Back condition with the word recognition task. This does not mean that there is no difference at all but that the effect we see is most pronounced in the hearing-impaired group when they did the auditory version of the 1-Back task.

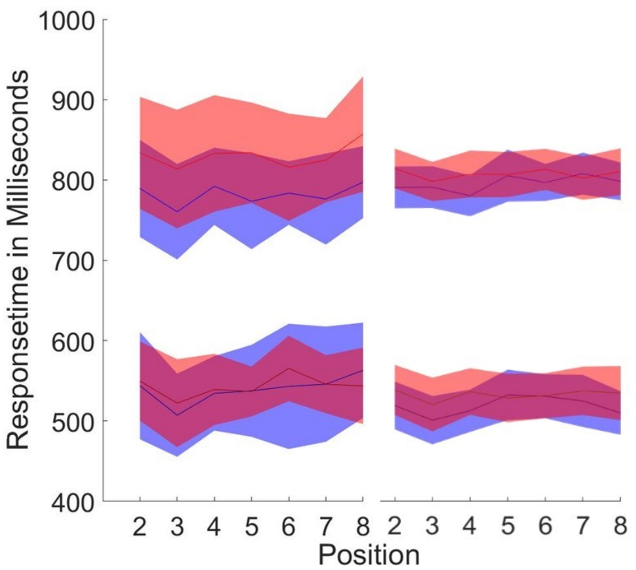

Figure 6: Response time for 1-Back trials in milliseconds. The solid line indicates the mean at each position of the 1-Back trial and the tube indicates the standard error. On the left side the data for hearing impaired are presented and the right site shows the normal hearing group. The upper tubes show the data for all 1-Back blocks with auditory stimuli and the lower tubes are indicate the visually presented stimuli. Red codes the 1-Back with the word recognition task and blue without.

Conclusion and Outlook

We have shown that the word recognition task draws cognitive energy from the ongoing task. This is especially or even exclusively true for hearing impaired people in situations with an ongoing auditory task.

This result helps to explain why such listening situations are so tiring for hearing impaired people.

For future research we are very interested in the effects the ongoing task has on the performance of the word recognition task. This effect will not only help to understand the challenges for hearing impaired but also tabs into areas like alerting systems and driverless cars.

We also want to quantify the tiring effect of our task and therefore employed pupillometry to measure the so-called Pupil Unrest Index which is a profound measure of sleepiness.

Literature

[1] Best, V., F.J. Gallun, A. Ihlefeld, and B.G. Shinn-Cunningham: The influence of spatial separation on divided listening. The Journal of the Acoustical Society of America 120 3, (2006), 1506-1516

[2] Koelewijn, T., B.G. Shinn-Cunningham, A.A. Zekveld, and S.E. Kramer: The pupil response is sensitive to divided attention during speech processing. Hearing research 312 (2014), 114-120

[3] Neher, T., G. Grimm, and V. Hohmann: Perceptual consequences of different signal changes due to binaural noise reduction: do hearing loss and working memory capacity play a role? Ear and hearing 35 5, (2014), 213-227

[4] Gosselin, P.A. and J.-P. Gagné: Older adults expend more listening effort than young adults recognizing speech in noise. Journal of speech, language, and hearing research 54 3, (2011), 944-958

[5] Gagne, J.-P., J. Besser, and U. Lemke: Behavioral assessment of listening effort using a dual-task paradigm: A review. Trends in hearing 21 (2017), 2331216516687287

[6] Wickens, C.D.: Multiple resources and performance prediction, in *Varieties of attention*, R. Parasuraman and D.R. Davies, Editors. 1984, Academic Press New York:.

[7] Wallenberg, E. and B. Kollmeier: Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten. Audiol Akustik 28 (1989), 50-65

[8] Inc., T.M.: *MATLAB*, in 2018a. (2018).