

## 1 Übung

### 1.1

At first, a sinusoidal sound signal with  $L_p = 60$  dB SPL, a duration of 1 s and a frequency  $f = 1$  kHz was generated. This signal was sampled at a frequency  $f_s$  of 20 kHz. To introduce the sound pressure level (SPL) into the signal the following derivation was developed. As reference sound pressure, the threshold of human hearing  $p_0 = 20$   $\mu$ Pa was chosen.  $\hat{p}$  thereby indicates the peak value of the pressure wave, while  $\tilde{p}$  is the effective value of the sinusoidal pressure amplitude. (Eq. 1) Simplistically, the static pressure component was neglected for further computations.

Solving equation 1 and 2 after  $\hat{p}$  results in the solution (3). Following this, the final pressure wave could be computed in equation 4 with the given parameters. A snippet from the wave is also shown in figure 1.

$$\tilde{p} = \frac{\hat{p}}{\sqrt{2}} \quad (1)$$

$$L_p = 20 \log \left( \frac{\tilde{p}}{p_0} \right) \quad (2)$$

$$\hat{p} = \sqrt{2} \cdot 10^{\frac{L_p}{20}} \cdot p_0 \quad (3)$$

$$p(t) = \hat{p} \sin(2\pi ft) \quad (4)$$

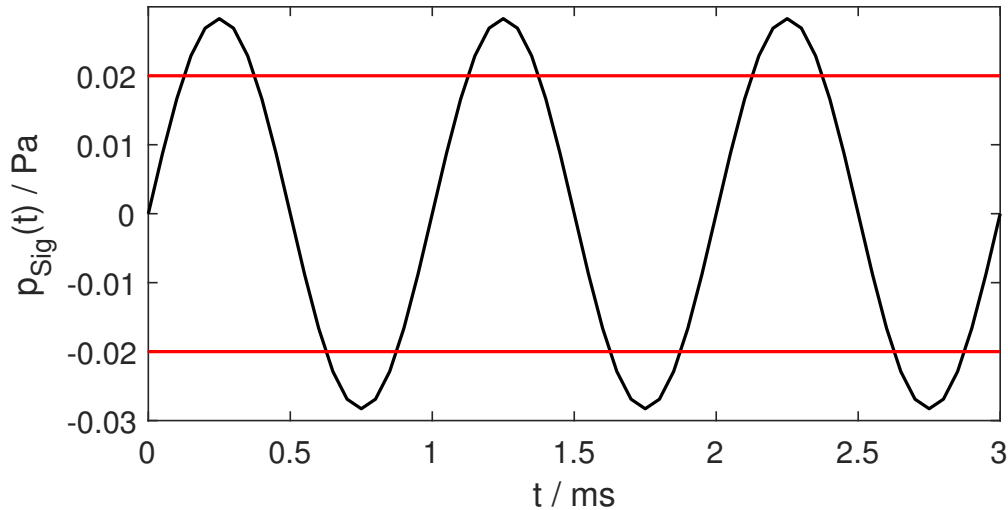


FIGURE 1 – Sinusoidal pressure signal for further analysis. The red lines indicate the RMS-SPL at 60 dB and lie at  $\approx 70\%$  ( $\frac{1}{\sqrt{2}}$ ) of the peak amplitude.

Subsequently, the acoustic pressure level was computed from the spectrum of the pressure wave from figure 1. We would assume a single peak slightly above 60 dB at 1 kHz with decaying flanks in the frequency domain. This is proven by the plot in figure 2.

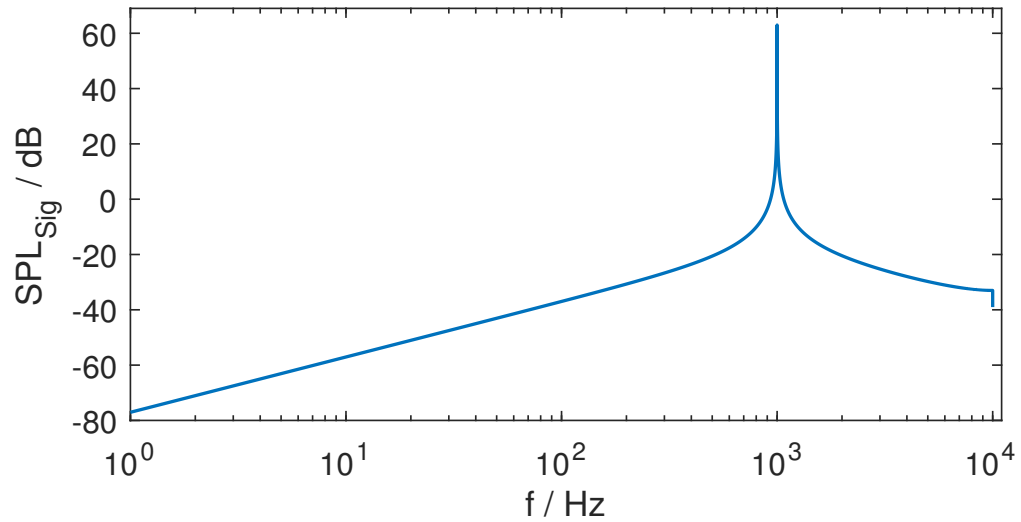


FIGURE 2 – Acoustic pressure level of the Fourier-transformed pressure wave from figure 1

## 1.2

By playing the sound on the headphones, a cracking noise is noticeable in the beginning and the end of the signal. This happens due to the non differentiable step in the signal at both sides. To counteract this, start and end of the tone were filtered with the corresponding half a Blackman window to generate a fading-effect.

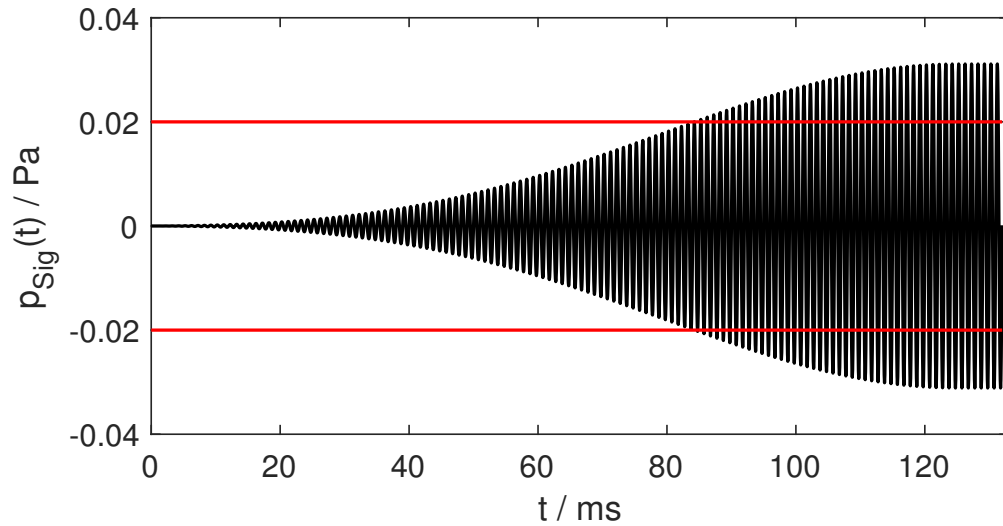


FIGURE 3 – Blackman - windowed sound signal for fading start and end to reduce the cracking noise. The red lines indicate the 60 dB SPL.

### 1.3

In this exercise, two recorded audio signals with the word "Nuss" and the letter "w" (spoken "double-u") were loaded into the program and scaled to a RMS-SPL of 60 dB SPL. Then, the corresponding spectrum of the SPL was computed. (Figs. 4 and 5).

Afterwards, both signals were A-weighted. (Figs. 6 and 7) This filter is applied to convert the machine-measured loudness of a signal to the non-linear relative loudness perceived by humans. The particular "A" filter corresponds to the 40-phon Fletcher–Munson curve, which is considered to be the most accurate equal-loudness contour for human hearing.

The pressure level spectra were also evaluated for both signals. (Figs. 8 and 9). In comparison to the levels of the unfiltered SPL, a clear dampening of frequencies below ( $\approx 10$  Hz) and above ( $\approx 20$  kHz) the absolute auditory threshold could be observed.

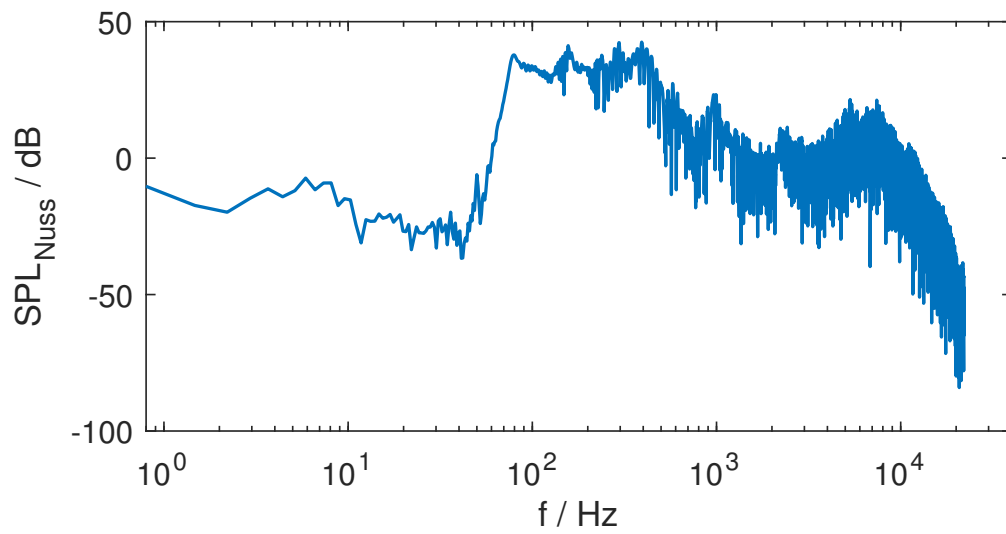


FIGURE 4 – SPL spectrum of an audio record of the german word "Nuss".

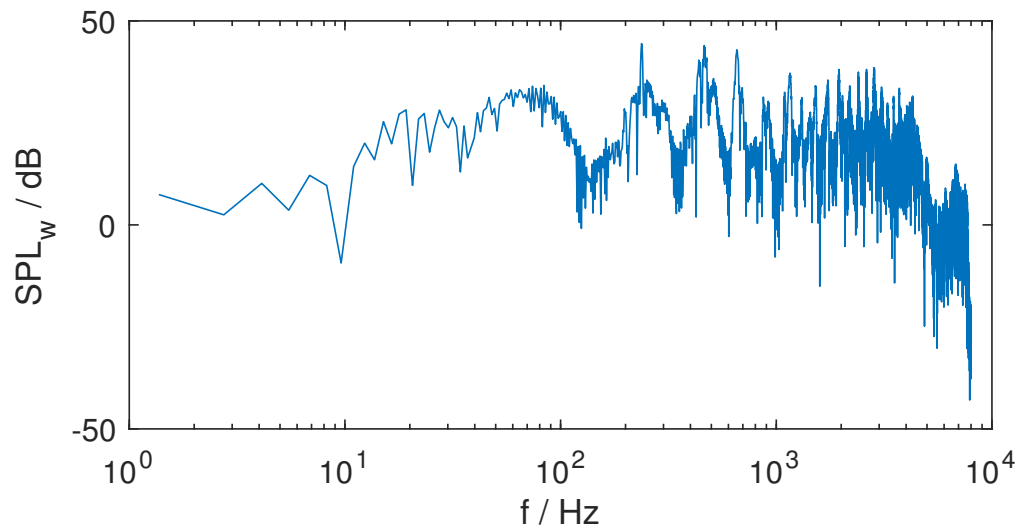


FIGURE 5 – SPL spectrum of an audio record of the english letter "w".

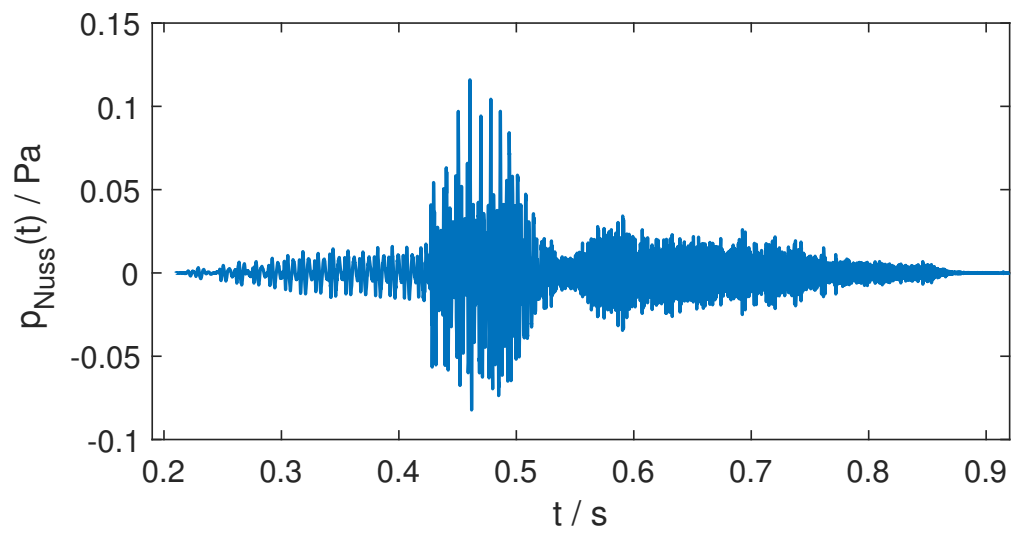


FIGURE 6 – A-weighted signal of an audio record of the german word "Nuss".

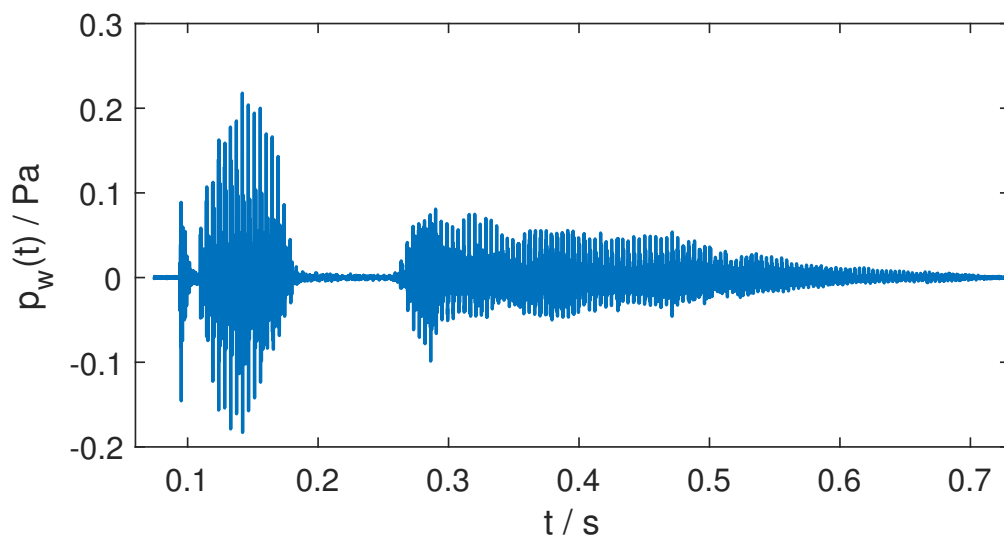


FIGURE 7 – A-weighted signal of an audio record of the english letter "w".

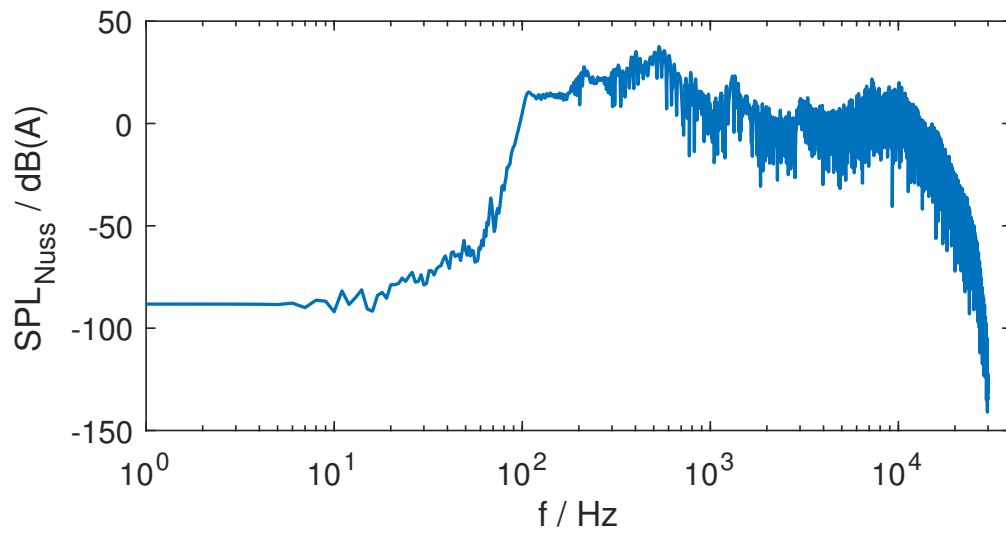


FIGURE 8 – A-weighted SPL spectrum of an audio record of the german word "Nuss".

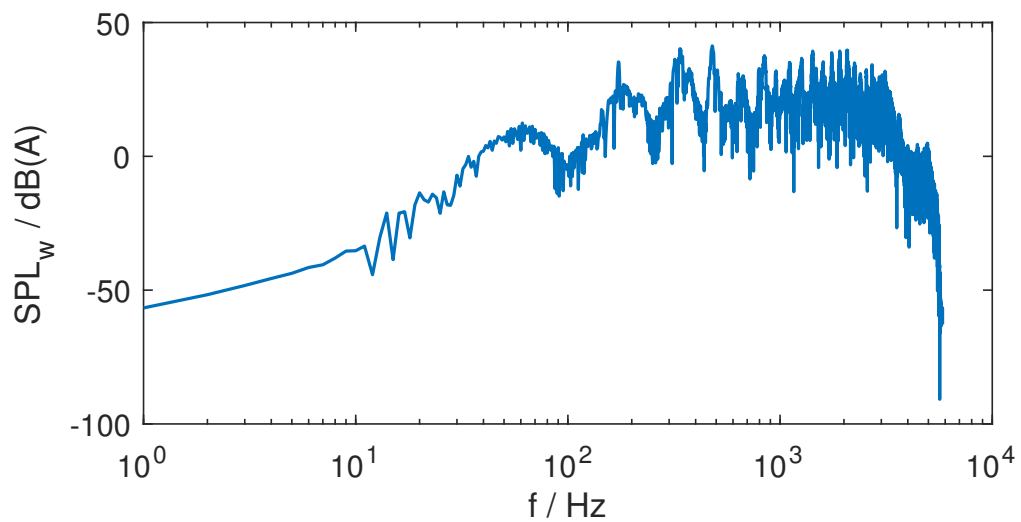


FIGURE 9 – A-weighted SPL spectrum of an audio record of the english letter "w".

#### 1.4

In the following, the audiosignals were plotted in a spectrogram (time x frequency x power). The dampening of the very high and very low frequencies by the A-filter can be observed as black stripes at the top and bottom of the corresponding plots. (Figs. 10 vs. 11 and 12 vs. 13).

For the word "Nuss", we can observe many frequency variations from letter to letter. The letter "N" can only be seen in very low frequencies with its harmonics above. "U" is then spoken in a slightly higher frequency band with more intense harmonics. In contrast, the hissing sound of the double "S" is a superposition of many medium-high-frequency components.

The letter "W" ("double-u") is constructed from more letters, which can be distinguished from the spectrogram. First, we can observe two mountain ranges from the parts "DOU" and "BL" with a small pause in between. "BL" then fades into the low frequent "U" component. Interestingly, "D" seems to range from low to high frequencies, while "OU" fades for above 6 kHz. "BL" is the first mountain after the short pause in the word and migrates to a unwritten "J"-like sound. The mountain raising from (2 kHz, 300 ms) to (5 kHz, 450 ms) and falling back to low frequencies from 500 ms to 600 ms account for the fade "BL" -> "J" and "J" -> "U". Similarly to "Nuss", "U" is then built only from low frequencies.

The parameters, with which the *spectrogram()* function is called, mainly affect the sharpness resp. smoothness between time- and frequency-steps. For the following, a Blackman-window with a length of 256 and a maximal overlap of 220 samples for adjoining windows was chosen. Fewer overlap hereby accounts for sharper transitions between time-frequency bins. Another parameter to refine the plot result is to maximize the number of DFT-bins "nfft". This was set to the MATLAB internal maximum of 256. By not evaluating desired frequencies specifically, MATLAB chose appropriate spaced frequencies from 0 to  $f_s/2$  itself.



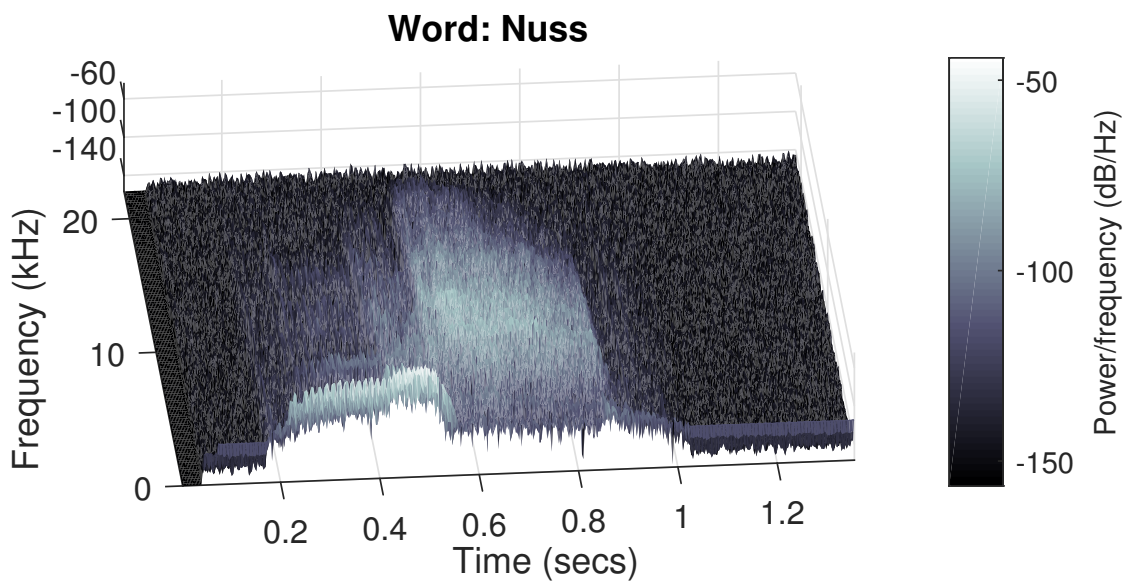


FIGURE 10 – Spectrogram of the word "Nuss".

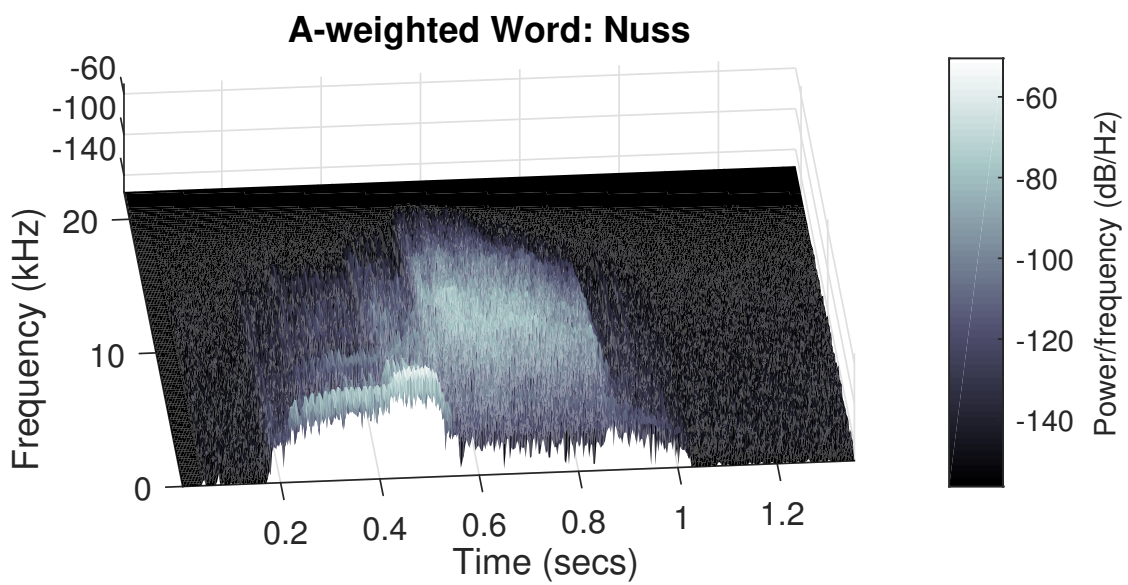


FIGURE 11 – A-weighted spectrogram of the word "Nuss".

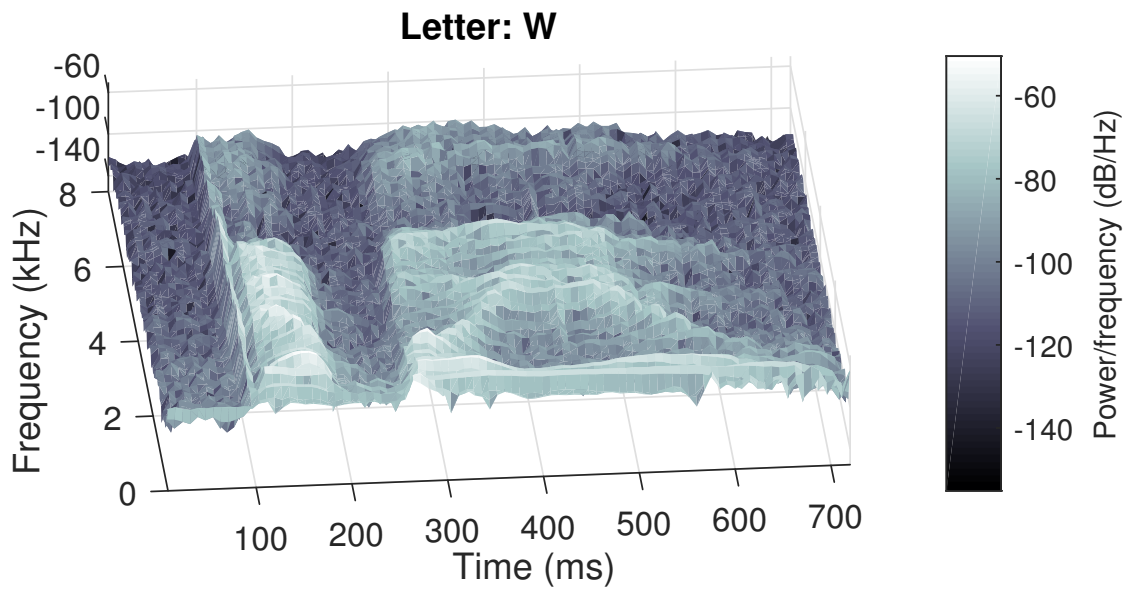


FIGURE 12 – Spectrogram of the letter "double-u".

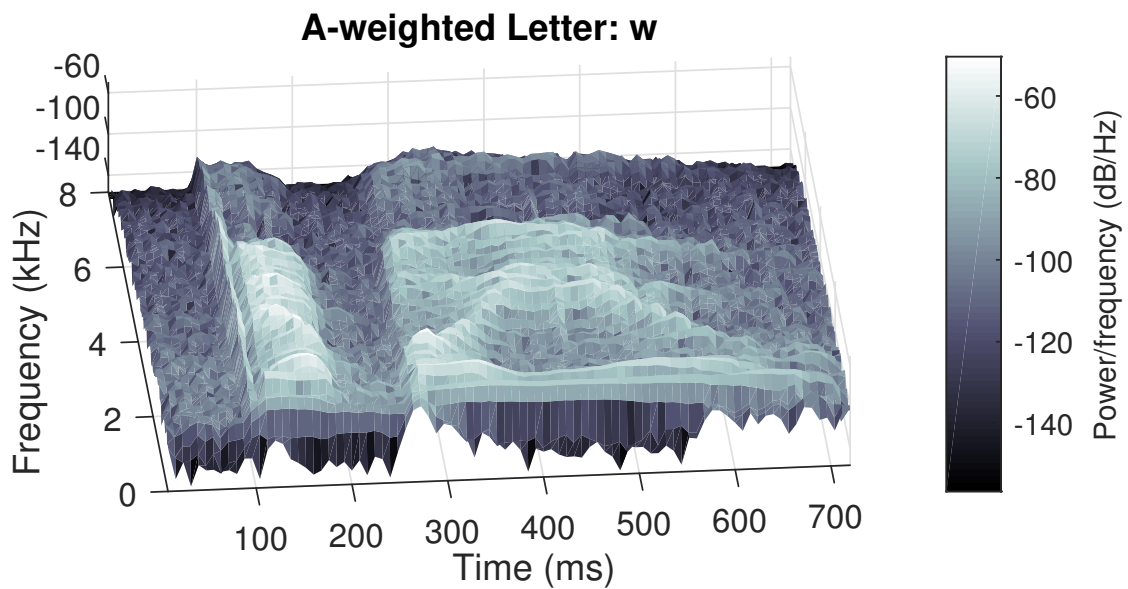


FIGURE 13 – A-weighted spectrogram of the letter "double-u".