Real-time Speech Exchange (RSE)

Manual

Version 2.0

January 2021

Björn Breidegard
Certec, Lund University, Sweden

Andreas Lind
Lund University Cognitive Science, Sweden

Contact, Andreas Lind: jan.oscar.andreas@gmail.com

Acknowledgments

The development of RSE was made possible by the following grants awarded to Andreas Lind: two grants from the **Crafoord Foundation** (20101020 and 20170734), a grant from the **Swedish Research Council** (2015-00446), and a grant from **Uno Otterstedt's Foundation** (EKDO2010/54).

Contents

List of figures

List of tables

1. Overview

1.1. Main functions of RSE

Real-time Speech Exchange (RSE) is an open-source software (it runs on Windows 10 and requires ASIO4ALL to be installed) that allows the researcher to, covertly or overtly, selectively exchange or otherwise manipulate the auditory feedback that test participants receive of their own voice via headsets. The basic operation mode of RSE is a general low-latency feedback mechanism relaying what the speaker is saying. An adjustable low-pass filter is added to all feedback (non-manipulated as well as manipulated) in order to mimic the bone conducted contribution to hearing one's own voice and to make the feedback sound as natural as possible (Shuster & Durrant, 2003). Using a sound detection-based trimming- and triggering system, utterances can be recorded and, later in the test, inserted with precise timing at the exact moment that the participant says something else. By simultaneously blocking the feedback of what the participant is saying, situations are created where participants say one thing but hear themselves saying something else.

In extensions of this basic functionality, RSE also allows for the insertion of pre-prepared WAV-files using the same mechanism. This allows the experimenter to insert any other sound that is desirable for the experiment at hand. This can include, for example, the following:

- The speaker's own voice recorded previously and processed off-line
- Someone else's voice
- Masking sounds, such as Brownian noise
- Emotionally charged vocalizations, such as gasps, laughs, cries etc.
- Bodily sounds, such as burbs, coughs, sneezes etc.

A further function is the insertion of WAV-files timed not to the utterances of the participant, but instead timed in relation to the start of the trial. For example, an insertion can be made at a specified time after trial start, e.g., 700 ms, allowing the experimenter to insert e.g., a recording of the participant's own voice but slightly before they start speaking (we call this "prelayed auditory feedback" in distinction to "delayed auditory feedback").

Breaks can be inserted in the experiment, with simultaneous visual presentation of e.g. a question, for example in order to evaluate the participant's experience of a specific feedback manipulation.

RSE is controlled via an intuitive visual interface and includes a visual stimuli presentation function that can show any JPG- or PNG-file, or simple text stimuli (with the option to colour the text). New experiments can easily be created and modified by editing a number of TXT-, RTF-, and XML-files.

1.2. Who developed RSE?

RSE was developed and designed at Lund University, Sweden, by Björn Breidegard (Certec) and Andreas Lind (LUCS) in collaboration with Petter Johansson, Lars Hall and Christian Balkenius (all at LUCS).

A critical contribution to the sound processing part of the system was the large and versatile hierarchical class library developed by Björn Breidegard over the last 30 years. The library, which is implemented using Visual C++, includes classes for video/audio capture, playback, and digital image/audio processing. The library has been continually developed in different projects (Breidegard & Andersson, 1992; Breidegard & Balkenius, 2003; Breidegard, 2007; 2013; Caltenco, Breidegard, & Andreasen Struijk, 2013), and has enabled the exchange and recycling of software components between the different projects.

1.3. How to cite RSE

Please cite RSE as:

Breidegard, B., & Lind, A. (2020). Real-time Speech Exchange (RSE) (version 2.0) [Computer software]. http://www.realtimespeechexchange.wordpress.com.

or

Breidegard, B., Hall, L., Johansson, P., Balkenius, C., & Lind, A. (in prep.). Real-time Speech Exchange (RSE): A technique for manipulating the auditory feedback of self-produced speech.

Publications involving RSE:

Lind, A., Hall, L., Breidegard, B., Balkenius, C. & Johansson, P. (2015). Auditory feedback is used for self-comprehension: When we hear ourselves saying something other than what we said, we believe we said what we hear. *Psychological Science, 26*, 1978-1980. DOI: https://doi.org/10.1177/0956797615599341

Lind, A., Hall, L., Breidegard, B., Balkenius, C. & Johansson, P. (2014). Auditory feedback of one's own voice is used for high-level, semantic monitoring: the "self-comprehension" hypothesis. *Frontiers in Human Neuroscience, 8:166*. DOI: https://doi.org/10.3389/fnhum.2014.00166

Lind, A., Hall, L., Breidegard, B., Balkenius, C. & Johansson, P. (2014). Speakers' acceptance of real-time speech exchange indicates that we use auditory feedback to specify the meaning of what we say. *Psychological Science, 25*, 1198-1205. DOI: https://doi.org/10.1177/0956797614529797

2. Getting started

This manual is written as a step-by-step guide on how to download RSE, set up the necessary hardware and software, and design, run and analyse an experiment.

A folder, "RSE", containing the executable version of **RSE** can be downloaded here. The "RSE" folder also contains the following:

- The folder "Demos", containing three demos (see section 2.2)
- The two files
    - "RSE_manual.pdf" (this manual)
    - "RSE_slide_example_small_size.pptx"
- The following six files
    - "BBDirect2DDll_d.dll"
    - "BBFormantWizard_d.dll"
    - "BBInteraction_d.dll"
    - "BBLoggerDll_d.dll"
    - "BBPlotter_d.dll"
    - "BBSoundProcessors_d.dll"

The code for RSE can be received from the authors upon request at the same webpage.

To run RSE you first need to install **ASIO4ALL**, which can be downloaded from the ASIO4ALL webpage. Follow download instructions on the webpage.

2.1. The Experiment folder

The files necessary to set up and run RSE are all contained in each of the three demos included in the download. The "Demo1-3" folders each contains the folder "Experiment", which, in turn, contains the following:

- The folder "Multimedia", which contains audio- and visual stimuli-files.
- The folder "LogFiles", to where log files are written.
- The folder "DebugFiles", to where debug files are written.
- The following files:
    - "BreakText.rtf" [*to be updated to a JPG-file*]
    - "CompletionText.rtf"
    - "IntroductionText.rtf"
    - "Experiment.txt"
    - "StimulusDefinitions.xml"

In addition to the "Experiment" folder, a folder, "SubstitutionWavFiles_YYYYMMDD_HHMMSS", will automatically be placed in the "Demo1-3" folders each time new substitution files are written.

2.2. Demos

There are three demos included which spell out experiments which allow you to familiarize yourself with RSE. See section 4 (specifically sections 4.1 and 4.2) for more information on how an experiment is structured, and section 8 for more instructions on how to use Demo 1.

2.2.1. Demo 1

This demo consists of 25 trials where the stimuli are Stroop words. There are two trials where recordings are to be made and two corresponding trials where these recordings are to be inserted. After each voice exchange the Break function shows the question "what did you say?"

2.2.2. Demo 2

This demo consists of 25 trials where the visual stimuli are JPG-images of line drawings taken from the Severens, Van Lommel, Ratinckx, and Hartsuiker (2005) set. There are six trials where recordings are to be made and six corresponding trials where these recordings are to be inserted. After each voice exchange the Break function shows the question “what did you say?”

2.2.3. Demo 3 [coming soon]

This demo consists of 25 trials where the stimuli are photos in JPG-format. There are three trials where recordings are to be made and three corresponding trials where these recordings are to be inserted using the pre-lay function [*coming soon*]. Additionally, there are five pre-recorded sounds, two words spoken by a man, two words spoken by a woman, and one 500 ms recording of Brownian noise (computer generated using Audacity: high-pass filtered, cutoff frequency 25 Hz and 48 dB rolloff per octave). These words/sounds are inserted on six different trials. After each voice exchange the Break function shows the question “what did you say?”

3. Hardware equipment and connections

Here we describe the required hardware (with suggestions for specific brands) and how to connect it. It is possible to use RSE with e.g., a combined mic- and headphone (4-pin) headset such as the Apple EarPods. However, while this can be used while familiarizing yourself with RSE, for experimental purposes you need a better microphone and more sound isolating headphones.

3.1. Mic/headphone splitter

Most laptops now have a combined mic- and headphone jack. As mentioned above, while not recommended for experimental purposes, RSE can be used with a combined mic- and headphone (4-pin) headset. To use RSE with separate microphone and headphone, the combined mic- and headphone jacks need to be separated with a mic/headphone splitter. A suggestion can be found here.

3.2. Microphone

We recommend that you use a high-quality microphone, such as the DPA 4088 Directional Headset Microphone, which can be found here. However, this microphone requires a phantom powered mixer (see below). If you want to avoid using a mixer, a 3.5 mm microphone can be connected directly to the microphone jack of the mic/headphone splitter.

3.3. Mixer (optional)

To use a microphone such as the DPA 4088, you need a mixer with phantom power. A suggestion, the Xenyx 802, can be found here. If you plug in your microphone directly in the microphone jack, you do not need a mixer.

3.4. Headphones

We recommend that headphones with high passive sound isolation are used, in order to prevent that the sound of the participant's actual utterances leak through the headphones during manipulated trials. Depending on the specific requirements of the experimental set up, we recommend for example:

- Etymotic insert earphones. These are especially good if you want to use RSE together with e.g., EEG or fMRI.
- Sennheiser HD 280 PRO. Note, however, that these might not be sound isolating enough for normal amplitude speech.
- You can use a combination of e.g., Apple EarPods together with hearing protection earmuffs with high passive sound isolation, such as 3M Peltor X5A, on top.
- For instructions on how to build headsets with extra passive sound isolation using Sennheiser HD 280 PRO loudspeaker drivers and 3M Peltor X5A hearing protection earmuffs (see figure 1), see the online supplementary material for Franken et al. (2019).

Figure 1. Specially built headphones with extra passive sound isolation.

3.5. External experimental screen

Connect the external screen and set it to "extend" (access settings by pressing Windows key+p). The external screen is automatically set as the experimental screen.

> **Note**
> The **Scale and Layout** under **Display Settings** should be set to 100%.

> **External screen connections currently known to work**
> VGA, Display Port.

3.6. Cables and audio splitter

You will need the following cables:

- 3.5 stereo jack to 2x RCA male cable.
- 3.5 mm stereo headphone splitter (1 male to 2 female)

3.6.1. If the Experimenter- and participant channel splitter is used

If you are using the Experiment- and participant channel splitter (see section 3.8), you will not need the 3.5 mm stereo headphone splitter (1 male to 2 female), but you will instead need:

- 3.5 stereo jack to 3.5 stereo jack (ca 30 cm).

3.7. Gamepad (optional)

A gamepad is optional as you can control RSE directly with the laptop keyboard (see section 6.2). You can find a suggestion here. We find that a gamepad offers more flexibility and swifter interactions with RSE, but ultimately it is a matter of taste. To customize your own gamepad button mapping, see section 4.1.7.

3.8. Experimenter- and participant channels splitter (optional)

This splitter needs to be built from scratch and besides the parts detailed below, you will need a soldering station, solder and wire stripper. The splitter splits the experimenter's headphones so that, when the Experimenter

Channel is turned on under Settings (see section 6.1.2.1), the left headphone relays the same audio heard by the participant, while the right headphone relays the following additional information:

- The word actually uttered by the participant during exchange trials.
- A 1.5 ms 3520 Hz sine tone (perceived as a click) as a marker of each stimuli onset.
- A 73 ms 440 Hz sine tone at the end of [*or at stimuli onset of*] the trial preceding a word which is to be exchanged.
- A 290 ms 220 Hz sine tone at the end of [*or at stimuli onset of*] the trial preceding a word which is to be recorded.
- A 29 ms 220 Hz sine tones bookending a recording when it is listened to (see section 6.2).

Note that, if you are using the low-pass filter for the participant's feedback, this will be relayed through the left, but not the right, headphone.

> **Note**
> An issue with **feedback between the participant- and experimenter channels** has been noted, meaning that there could potentially be feedback from the experimenter to the participant channel during experiments (both sine tones and non-manipulated speech during exchanges). As this can lead to serious confounds in the experiment, we recommend that you make sure there is no such feedback before starting your experiment.

You can build e.g., either a simple splitter, or a sturdier box. Building instructions for both are detailed below.

3.8.1. Building instructions - Simple splitter [coming soon]

The splitter needs to be built from scratch and requires the following parts:

- 3 x 3.5 mm cable mount stereo jack sockets.
- About 20 cm of equipment wire.

> **Warning!**
> If you solder the splitter incorrectly, your laptop could be severely damaged.

3.8.2. Building instructions – Splitter box [coming soon]

> **Warning!**
> If you solder the splitter incorrectly, your laptop could be severely damaged.

If you wish to build a sturdier box, the following parts are required:

- Potting box around 7 x 5 x 3 cm with a screw-on lid.
- 3 x 3.5 mm panel mount stereo jack sockets.
- About 20 cm of equipment wire.

Figure 2. Experiment- and participant channels splitter box.

3.9. Connections

Figure 3 shows the experimental hardware set up and connections. Table 1 lists the same hardware connections.

> **Note**
> Always connect your laptop, mixer, and any other equipment to **grounded power outlets**.

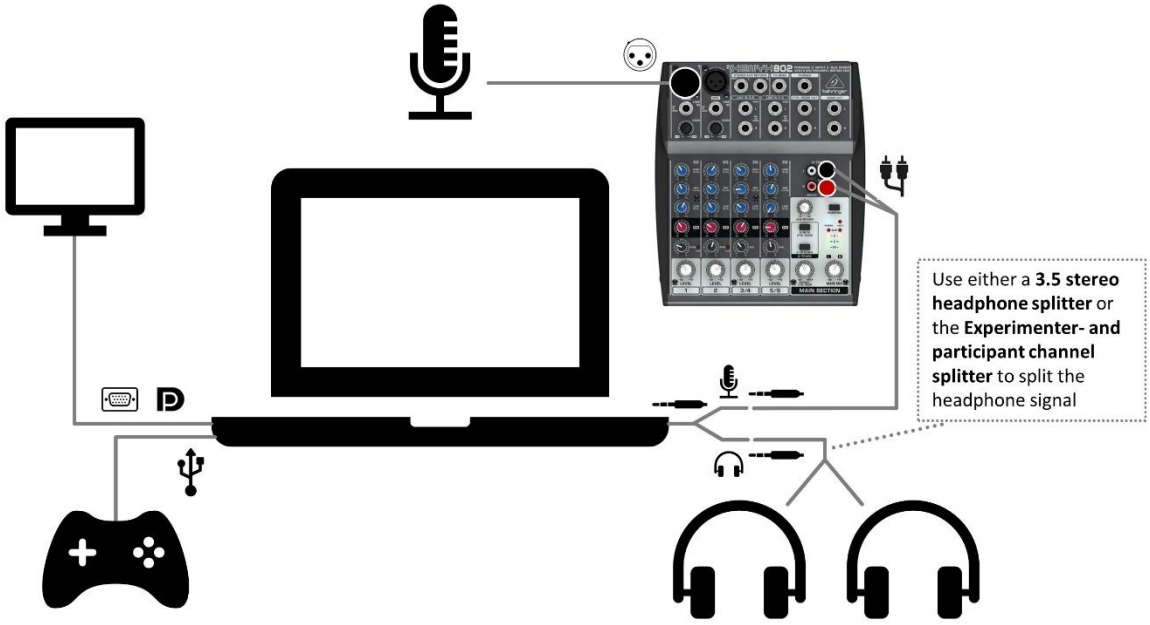

Figure 3. Hardware set up and connections including connecting the microphone via a Xenyx 802 mixer.

Table 1. Hardware connections.

Experimenter- and participant channel splitter is not used (Experimenter Channel turned off)	*Experimenter- and participant channel splitter is used (Experimenter Channel turned on)*
1. Xenyx 802 AC Power In <> Power outlet 2. Laptop <> Power outlet 3. USB Gamepad > laptop 4. Laptop <> Mic/headphone splitter 5. Microphone > Xenyx 802 Mic Preamp 1 6. Xenyx 802 2-Track Output L/R > mic/headphone splitter (mic) 7. Mic/headphone splitter (headphones) > 3.5 stereo headphone splitter 8. 3.5 stereo headphone splitter > Experimenter headphones 9. 3.5 stereo headphone splitter > Participant headphones	1. Xenyx 802 AC Power In <> Power outlet 2. Laptop <> Power outlet 3. USB Gamepad > laptop 4. Laptop <> Mic/headphone splitter 5. Microphone > Xenyx 802 Mic Preamp 1 6. Xenyx 802 2-Track Output L/R > mic/headphone splitter (mic) 7. Mic/headphone splitter (headphones) > Home-built splitter box 8. Experiment- and participant channel splitter > Participant headphones 9. Experiment- and participant channel splitter > Experimenter headphones

3.10. Laptop and mixer settings

Phantom power must be turned on on the mixer. You need to measure the input/output so that the sound in the headphones is e.g., 10dB above what goes into the microphone.

4. Setting up an experiment

To design and run an experiment using RSE, you use several definition files and multimedia files, organized in folders. These were mentioned under section 2.1. but are described in detail here.

4.1. Experiment folder

An experiment folder can be named after the experiment (e.g., "Stroop Experiment") and should contain a folder named Multimedia (described in section 4.2) in addition to the following files:

- "Experiment.txt"
- "StimulusDefinitions.xml"
- "IntroductionText.rtf"
- "CompletionText.rtf"
- "BreakText.rtf" [*to be replaced*]

> **Note**
> TXT-files and XML-files should be edited in a text editor, such as *Notepad* (XML-files can also be edited in specific XML-editing programs). RTF-files are most easily edited in *WordPad*.

4.1.1. Experiment.txt

This file defines the test sequence. It consists of five columns. Each row defines one trial. There are two ways of using visual stimuli in RSE: (1) each visual stimuli is a JPG- or PNG-file, (2) each visual stimuli is a text, e.g., a word. Table 2 describes the function of each column for both alternative approaches. Example files for both approaches are given in figure 4A, showing Demo 1, and 4B, showing Demo 2.

Table 2. Description of the functions of the six columns in the Experiment.txt file.

Column	*Stimuli is a JPG- or PNG-file*	*Stimuli is text*
1	**Defines recorded and inserted words** This column fills two functions: (1) It names the word to be recorded during R-trials (and consequently which words appear in Recording Manipulation (unless pre-recorded auditory stimuli is used, in which case the name of each recording will elicit an appearance in Recording Manipulation)); (2) it defines which word is to be inserted during exchange trials, meaning you only need to write something in column 1 during recording- and exchange trials (otherwise, write "-").	
2	**Defines JPG- or PNG-image** Defines which JPG- or PNG-image shall be shown. The name of each JPG- or PNG-image needs to be defined in StimulusDefinitions.xml.	**Defines text colour** Defines the colour of the text of the visual stimuli. The colours are defined in StimulusDefinitions.xml (in the demo package there are six colours defined: black, red, green, blue, brown and grey).
3	**Not used** Always mark "-" (if something is written here, then it is ignored by RSE)	**Defines text** Defines the text of the visual stimuli. Case sensitive.
4	**Defines word recording, test breaks, and prelayed insertions** Defines if the word should be recorded (mark "R"), if there should be a break (mark "B" ["BN" i.e., "B1", "B2" etc.]) after the trial (see "5.1.4. BreakText.txt" below), or, in case the playback is timed to the stimuli presentation rather than triggered by the voice trigger (i.e., prelay mode), the time in ms after stimuli presentation that the recording should be played back (mark e.g., "P800").	
5	**Defines stimulus display time and trial length [*to be added*]** The first XXX defines the stimulus display time and the second XXX defines the trial length (and, consequently, the interstimulus interval). E.g., 1500/2500 gives a stimulus display time of 1500 ms and a trial length of 2500 ms. *Note:* If this is not defined in the Definition file, then the values given in Settings is used, i.e., the definitions in the definition file overrides the settings in Settings.	
5	**Defines trigger threshold for each trial [*to be added*]** Allows you to set the trigger threshold for each individual trial. If it is marked "-" the standard trigger threshold (set in Settings > Sound Trig Threshold 1) is used instead.	
6	**Allows you to make a note (e.g. "REC" or "PLAY")** If the comment is started with a semicolon it is not visible in the RSE Test Manager. Otherwise it appears there.	

RSETestDefinition.txt

```
1  -   brown   blue    -   -
2  -   green   blue    -   -
3  -   brown   brown   -   -
4  -   blue    green   -   -
5  -   red blue    B   -
6  -   blue    grey    -   -
7  -   brown   red -   -
8  green   green   brown   R   REC
9  -   grey    brown   -   -
10 -   brown   green   -   -
11 green   grey    green   B   PLAY
12 -   brown   grey    -   -
13 -   red green   -   -
14 grey    grey    red R   REC
15 blue    brown   B   -   NM
16 -   grey    blue    -   -
17 -   red brown   -   -
18 -   green   red -   -
19 -   green   green   -   -
20 -   blue    red -   -
21 grey    green   grey    B   PLAY
22 -   blue    blue    -   -
23 -   grey    grey    -   -
24 -   red red -   -
25 -   red grey    -   -
26
```

A

RSETestDefinition.txt

```
1  -   moon    -   -   -
2  pan pan -   R   REC
3  -   tent    -   -   -
4  bat bat -   R   REC
5  -   mouse   -   -   -
6  pan pen -   B   PLAY
7  deer    deer    -   R   REC
8  train   train   -   R   REC
9  -   chair   -   -
10 bat cat -   B   PLAY
11 key key -   R   REC
12 deer    pear    -   B   PLAY
13 -   eye -   -
14 boat    boat    -   R   REC
15 -   clock   -   -   -
16 -   pig -   -
17 -   house   -   -   -
18 -   bus -   -
19 train   truck   -   B   PLAY
20 boat    goat    -   B   PLAY
21 -   ear -   -
22 key tree    -   B   -
23 -   fly -   -
24 -   hat -   -
25 -   horse   -   -   -
26
```

B

Figure 4. [*to be updated with 6 columns, and will include all three demos*] Experiment.txt-files for Demo 1 (A), where the stimuli is text, and Demo 2 (B), where the stimuli are JPG- or PNG files.

4.1.2. StimulusDefinitons.xml

This file defines the visual stimuli used for your experiment (figure 5). Figure 5 contains the stimulus definitions needed for both Demo 1 and Demo 2. Rows 5-8 defines the RGB colour codes for the text stimulus in Demo 2 (row 4 defines the colour code for black text) as defined in row 2 of the Experiment.txt file, when text is used as visual stimuli. Rows 10-34 defines the names of the image files used in Demo 1, as defined in row 2 of the Experiment.txt file, when images are used as visual stimuli.

StimulusDefinitions.xml

```
1  <?xml version="1.0" encoding="utf-8"?>
2
3  <StimulusList>
4    <Stimulus Name="black"> <Color> 0 0 0     </Color> </Stimulus>
5    <Stimulus Name="red">   <Color> 255 0 0     </Color> </Stimulus>
6    <Stimulus Name="green"> <Color> 0 255 0     </Color> </Stimulus>
7    <Stimulus Name="blue">  <Color> 0 0 255     </Color> </Stimulus>
8    <Stimulus Name="brown"> <Color> 139 69 19   </Color> </Stimulus>
9    <Stimulus Name="grey">  <Color> 128 128 128 </Color> </Stimulus>
10   <Stimulus Name="moon"> <Picture> moon.jpg </Picture></Stimulus>
11   <Stimulus Name="pan"> <Picture> pan.jpg </Picture></Stimulus>
12   <Stimulus Name="tent"> <Picture> tent.jpg </Picture></Stimulus>
13   <Stimulus Name="bat"> <Picture> bat.jpg </Picture></Stimulus>
14   <Stimulus Name="mouse"> <Picture> mouse.jpg </Picture></Stimulus>
15   <Stimulus Name="pen"> <Picture> pen.jpg </Picture></Stimulus>
16   <Stimulus Name="deer"> <Picture> deer.jpg </Picture></Stimulus>
17   <Stimulus Name="train"> <Picture> train.jpg </Picture></Stimulus>
18   <Stimulus Name="chair"> <Picture> chair.jpg </Picture></Stimulus>
19   <Stimulus Name="cat"> <Picture> cat.jpg </Picture></Stimulus>
20   <Stimulus Name="key"> <Picture> key.jpg </Picture></Stimulus>
21   <Stimulus Name="pear"> <Picture> pear.jpg </Picture></Stimulus>
22   <Stimulus Name="eye"> <Picture> eye.jpg </Picture></Stimulus>
23   <Stimulus Name="boat"> <Picture> boat.jpg </Picture></Stimulus>
24   <Stimulus Name="clock"> <Picture> clock.jpg </Picture></Stimulus>
25   <Stimulus Name="pig"> <Picture> pig.jpg </Picture></Stimulus>
26   <Stimulus Name="house"> <Picture> house.jpg </Picture></Stimulus>
27   <Stimulus Name="bus"> <Picture> bus.jpg </Picture></Stimulus>
28   <Stimulus Name="truck"> <Picture> truck.jpg </Picture></Stimulus>
29   <Stimulus Name="goat"> <Picture> goat.jpg </Picture></Stimulus>
30   <Stimulus Name="ear"> <Picture> ear.jpg </Picture></Stimulus>
31   <Stimulus Name="tree"> <Picture> tree.jpg </Picture></Stimulus>
32   <Stimulus Name="fly"> <Picture> fly.jpg </Picture></Stimulus>
33   <Stimulus Name="hat"> <Picture> hat.jpg </Picture></Stimulus>
34   <Stimulus Name="horse"> <Picture> horse.jpg </Picture></Stimulus>
35 </StimulusList>
```

Figure 5. StimulusDefinitions.xml file. See text for description.

4.1.3. IntroductionText.rtf

The introduction text is shown on the experiment screen when RSE is opened. Here, test instructions can be written. If no instructions are needed, simply leave the file blank.

4.1.4. CompletionText.rtf

The completion text is shown when the experiment is finished and can read e.g., "Thank you for participating". If no completion text is needed, simply leave the file blank.

4.1.5. BreakTextN.rtf

[*This will be changed to a JPG-file (in Multimedia) so that you can define "B1", "B2", etc. and each one shall call a JPG-file*]
The break text is shown when a break is inserted in the experiment. The number corresponds to the number in the fourth column of the Experiment.txt file: BreakText1 – B1; BreakText2 – B2 etc [*presently there can be only one BreakText-file*]. The break text can read e.g. "What did you just say?".

4.1.6. RSEKeyboardMapping [coming soon]

This file allows the user to specify which keys on the keyboard fills what function.

4.1.7. RSEGamepadMapping [coming soon]

This file allows the user to specify which buttons on the gamepad fills what function.

4.2. The "Multimedia" Folder

This folder contains the image- and sound files used as visual stimuli and, if the experiment design requires, as pre-recorded sounds used for specified voice exchanges, respectively. It also contains the break-files [*coming soon*].

4.2.1. Visual stimuli files

Use JPG- or PNG-files for visual stimuli. Each file should be given a name defined in StimulusDefinitions.xml.

Presently, RSE can malfunction if the combined size of visual stimuli files is too large. Therefore, we recommend creating files of smaller size. For example, if you create your stimuli in Power Point, set the Custom Slide Size to 5 cm width and 3.5 cm height (one such slide is included in the starter pack), insert any text or images you want to use for your experiment, and then save in desired format. Remember to name each file.

If you use the "break" function (column 4 in the Experiment.txt file, see section 4.1.1.) you need to include one or more files named e.g., "BN.jpg", i.e., "B1.jpg", "B2.jpg", "B3.jpg" etc. [*this function is coming soon*]

> **Note**
> Image files currently known to work with RSE are: JPG, PNG.

4.2.2. Auditory stimuli files

You can name a WAV-file e.g., "pan.wav" to call it at a specific trial where you have specified that an exchange with a recording with that name is to be made. So, for example, row 6 in figure 4B, specifies that "pan" is to be inserted on trial 6. If no WAV-file is added to Multimedia with that name, you need to record that word (at row/trial 2). If there is a file with that name added to Multimedia (and you do not make a recording on trial 2),

the WAV-file will be inserted on trial 6. If you are planning on using recordings that are pre-loaded in Multimedia (e.g., of the participant's or someone else's voice), you do not need to include a trial to record that word (in this case trial 2).

> **Note**
> Audio files currently known to work with RSE are: WAV.

4.2.3. Break files [coming soon]

4.3. Creating substitution files with RSE

[*this is not currently working properly*] To create new substitution files of e.g., someone else's voice to insert during experiments, you open RSE using a definition file containing all the words you want to record as REC-words. You then place the marker on the appropriate word in the Recording Manipulation-window and make a recording. It is recommended that you record words using the same set up, with the same settings, and under the same circumstances as will prevail during the experiment, to ensure equivalent sound quality, and that ambient noise etc. is identical. If a specific recording is not satisfactory for some reason, simply do it again, and the previous recording will be over-written. Once all words are recorded, press Tools [*File*] > Write All Substitutions To Wav-file. The recordings are written as WAV-files to a folder "SubstitutionWavFiles_YYYYMMDD_HHMMSS", with each recording named after the word on which it was recorded. These WAV-files can then either be moved as they are to the Multimedia folder for use in a future experiment, or they can be processed offline in any way desirable before being used in the experiment.

4.4. Log files

Files are written as WAV-files to the "LogFiles" folder with the name "LogFile_YYYYMMDD_HHMMSS". Each file has two channels. Channel 1 is the Participant Channel, i.e., the channel heard by the participant during the experiment (i.e., including the inserted words), while Channel 2 is the Experimenter Channel (see section 3.8), containing the word the participant is actually saying during manipulated trials, as well as the four types of sine tone described under section 3.8.

For suggestions on how to perform data analysis, see section 9.1.

4.5. Debug files [*coming soon*]

5. Starting RSE

5.1. Screen and sound settings

Before starting RSE, there are some screen and sound settings to be made on your computer:

- The **Scale and Layout** under **Display Settings** should be set to 100%.
- [*necessary?*] Under Control panel -> Sound -> Playback -> Properties [for the active unit] -> Improvements -> System Effects Configuration Effects -> De-activate all sound effects.
- [*necessary?*] Under Control panel -> Sound -> Recording -> Properties [for the active unit] -> Advanced -> Signal improvements -> De-activate "activate improved sound"

5.2. Starting RSE

Start RSE by double clicking "RSE.exe".

5.3. ASIO settings

Once you have opened RSE for the first time, and before using it, you need to make some settings in ASIO. To access the ASIO settings from RSE, press Tools > Show ASIO Dialog. The ASIO settings will depend on your laptop's internal sound card. A general rule is that only your internal sound card should be activated in ASIO (in figure 6, this is Conexant ISST Audio). Only two things should be activated within your sound card (from within ASIO): Headphones and Audio capture (it might be necessary to try some different settings before getting it right). The ASIO buffer size must be set to a minimum (64 Samples). The Latency Compensation should be set to a minimum for both In and Out (0 Samples). Allow Pull Mode (WaveRT) should be activated (Buffer Offset should be set to 1 ms [*1 or 0 ms?*]). Always Resample 44.1kHz <-> 48kHz should be activated. Force WDM Driver To 16 Bit should be de-activated.

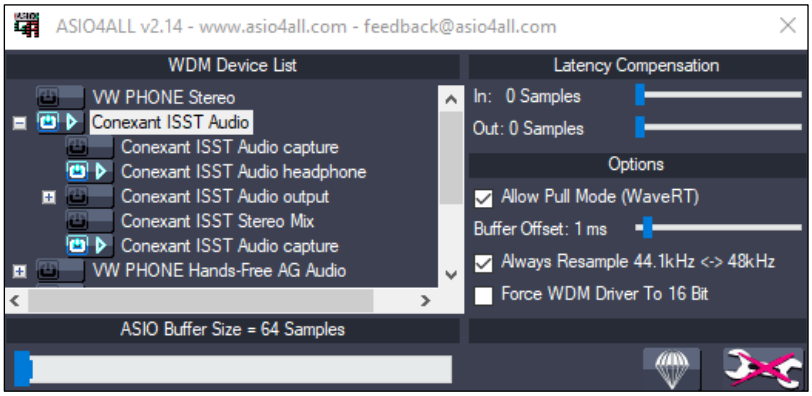

Figure 6. ASIO settings.

6. User interface

RSE is controlled via a visual interface and via the computer's keyboard and/or a computer game gamepad.

6.1. Visual interface

The visual interface (figure 7) contains a command bar and several windows used by the experimenter to change settings in, and control, RSE. The computer's mouse and keyboard and/or a gamepad (see section 6.2) is used to interact with the different windows.

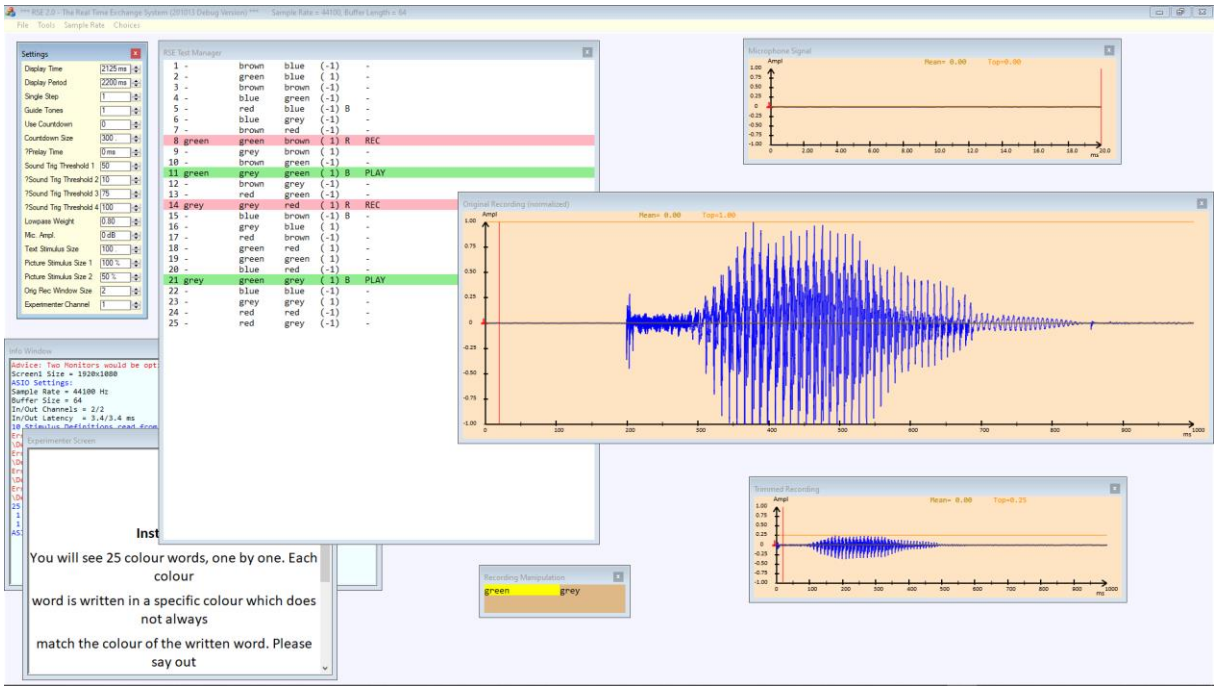

Figure 7. Visual interface.

6.1.1. Command bar

The functions of the command bar buttons are described in table 3.

Table 3. Command bar functions.

Key	*Sub-key*	*Function*
File	***Open Experiment File***	Allows you to open an experiment. Press the button, find the experiment you wish to open and mark the folder containing the experiment definition files and the Multimedia folder.
	Write All Substitutions to WAV-file	Allows you to write all the recordings you have made to separate WAV-files (see section 4.3 for more details)
Tools	***Show RSE Test Manager […] Show The Three Oscilloscopes [to be renamed "Waveform Displays"]***	Allows you to open each of the visual interface windows (these are all open when you open RSE).
	Show Latency Measurement Oscilloscope [to be renamed "Waveform Display"]	[*not currently working*]
	Show ASIO Dialog [to be updated: "Dialog" to "Settings"]	Allows you to access the ASIO settings.
	Show Settings Dialog [to be updated "Dialog" to "Window"]	Allows you to open the Settings Window (this is open when you start RSE).
	Write All Substitutions to WAV-file [to be moved to "File"]	Allows you to write all the recordings you have made to separate WAV-files (see section 4.3 for more details) [*to be moved to "File"*].
	Debug Mode	See section XXX. To set RSE to debug mode, press Debug Mode and restart RSE.
	Latency Measurement Mode	See section XXX. To set RSE to latency measurement mode, press Latency Measurement Mode and restart RSE.
	Reset Default Settings [to be inserted]	Resets the default settings of RSE.
Sample Rate		*Set this to 44100Hz *Buffer Size: 64
Choices		

6.1.2. Visual interface windows

The visual interface contains a number of different windows. Below, the Settings window is first described separately, and then the windows used to control RSE during experiments are described.

6.1.2.1. Settings window

In the Settings window, a number of settings can be made. These are detailed in Table 4.

Table 4. The functions of the Settings window.

Setting	*Function*
Display Time	Sets the time period for displaying each visual stimulus.
Display Period	Sets the time period starting from the display start of one visual stimuli until the display start of the next visual stimuli.
Single Step	If Single Step is set to 0, the test will run through automatically once started. If it is set to 1, the experimenter can feed forward each trial themselves (using button 1, see **Keyboard and/or gamepad** below).
Guide Tones	If set to 1, the experimenter channel (see section 6.1.2.15 below) automatically plays back the following sine tones during an experiment: a 1.5 ms 3520 Hz sine tone (perceived as a click) marking each stimuli onset; a 290 ms 220 Hz sine tone at the end of [*or at stimuli onset of*] the trial directly preceding a trial where a recording is to be made; and a 73 ms 440 Hz sine tone at the end of [*or at stimuli onset of*] the trial directly preceding a trial where an exchange is to be made. These tones are recorded to channel 2 of the LogFile, irrespective of if Guide Tones is set to 0 or 1 [*correct?*], in order to simplify data analysis (see section XXXXX).
Use Countdown	If set to 1 there will be a countdown (5, 4, 3, 2, 1) before the test starts, irrespective of whether the experiment uses Single Step or not.
Countdown Size	Sets the size of the countdown numbers on the screen.
?Prelay Time	
Sound Trig Threshold 1-4	Sets the trigger threshold for sound recording and triggering. Sound Trig Threshold 1 is the default threshold, while Sound Trig Threshold 2-4 sets the threshold for buttons 2-4 on the keyboard and gamepad. To use these thresholds, hold button 2, 3, or 4 while recording (button 5) or triggering an exchange (button 8) (see section 6.2). [*the usefulness of the trigger settings should be explained in more detail somewhere*] [*needs to be explained in more detail what a setting, e.g., 50, means*]
Lowpass Weight	Applies a low-pass filter [*describe*]. We recommend that it is set to 0.80 [*explain*].
Mic. Ampl.	Should always be set to 0 dB when RSE is used in experimental settings. It can be increased in order to try RSE for e.g., practice purposes with a microphone that does not give off a strong signal.
Text Stimulus Size	Sets the size of the text stimulus.
Picture Stimulus Size 1 [to be renamed "Stim Size Part Scrn"]	Sets the size of the visual stimuli in the external, participant's, screen. [*dubbelkolla så det inte är tvärtom*]
Picture Stimulus Size 2 [to be renamed "Stim Size Exp Scrn"]	Sets the size of the visual stimuli in the Experimenter Screen [*döps om till Participant Screen*] window. [*dubbelkolla så det inte är tvärtom*]
Orig Rec Window Size	Sets the size of the Original Recording window. The setting in figure 7 is 2, making this window larger than the other waveform displays, which makes trimming (described under section 6.1.2.2. **"Original Recording (normalized)"**) easier. Each time you change the size, you must restart RSE for it to take effect.
Experimenter Channel	

6.1.2.2. RSE control windows

A number of windows are used to control RSE during experiments. These are described in Table 5.

Table 5. The functions of the RSE control windows.

Window	*Function*
Participant Screen	This is a small insert showing the same thing that is being shown on the external (participant's) screen.
Test Sequence	This shows the test structure with each row showing one trial (i.e., it is based on the Experiment.txt file). If a recording is supposed to be made on a specific trial, that trial is here marked in red, while if a trial is supposed to include a voice exchange, that trial is marked in green. Trials are fed forward, so that the current trial is always the trial on the top of the window.
Info Window	The Info window gives advice about optimum RSE functionality. For example, if no second monitor is connected, there will be an "advice" in red letters saying "Two monitors would be optimal!". It gives the sizes of Screen1 and Screen2. It gives a summary [?] of the ASIO settings (Sample Rate, Buffer Size, nr of In/Out channels, and an estimate [?] of the In/Out latency. It tells you how many stimuli are defined in StimulusDefinitions.xml, and how many Stimuli are read from Experiment.txt. It also tells you which words are to be recorded and/or are pre-recorded and will be used during voice exchanges. [*does it give continuous info about what is going on in the test?*]
Word Recordings	This window contains all the words marked in the Experiment.txt file as words to be recorded, and it allows you to mark a specific word, do a recording and edit the recording (see **6.1.8. Original Recording (normalized)** below). You can always go back to a recording later, and edit it.
Microphone Signal	Waveform display of the microphone signal.
Original Recording (normalized)	This window shows a recording with a green vertical line showing where the recording is automatically trimmed [*however, this seems off now, even though recordings are correctly trimmed. Will be fixed*], and a red line where the recording ends. The latest recorded word is displayed either until a new word is recorded, or until another recording is chosen in Recording Manipulation. If a recording is not properly trimmed automatically (which can be seen when the green vertical line is not placed at the start of the word), manual trimming can be performed by holding the mouse indicator at the start of the word, holding down the left mouse button, dragging the mouse indicator to the place where you want the recording to end, and letting go of the left mouse button.
Start Stop Computing Magic…	[*will be removed*]
Trimmed Recording	Shows the trimmed recording.

6.2. Keyboard and/or gamepad

Recordings, exchanges etc. are performed either with the laptop keyboard or with a computer game gamepad, or with both. The buttons used to control RSE are described below. The same buttons are used for both the keyboard and the gamepad. The gamepad is pictured in figure 8, while table 6 lists the use of each key 0-9 and ↑, ↓, ←, and →. The functions described are the standard settings. However, the "RSEKeyboardMapping" file allows you to customize the functions of the keys on the keyboard (see section 4.1.6), while the "RSEGamepadMapping" file allows you to customize the functions of the buttons on the gamepad (see section 4.1.7). There are additional functions that can be performed with the keyboard only. These are described below, in section 6.2.1.

Figure 8. Gamepad.

Table 6. Functions of keys/buttons 1-9, ↑, ↓, ←, and → on the keyboard and gamepad.

Key	*Function*
0	Starts LogFile-recording [to be removed and moved to 9]
1	- If *Single Step* mode = 0: Press to start the test. - If *Single Step* mode = 1: Press to start the test and then to feed forward one trial.
2	Hold while pressing 5 or 8 to activate the trigger level set in "Settings" (see section 6.1.2.8).
3	Hold while pressing 5 or 8 to activate the trigger level set in "Settings" (see section 6.1.2.8).
4	Hold while pressing 5 or 8 to activate the trigger level set in "Settings" (see section 6.1.2.8).
5	Hold to record an utterance. Hold button for as long as you would like the recording to last. - If *Single Step mode* = 0*:* Place the indicator in Recording Manipulation window on the desired word and hold button. - If *Single Step mode* = 1: Hold button on the trial where the word is to be recorded. - Always make sure there are 50-150 ms of silence recorded after each recorded utterance.
6	Press to play back recording without using voice trigger. [*does not work anymore, will be reinstated*]
7	Press to listen to recording (note: only works if the Experimenter- and participant channel splitter is used). The extracted word is played in the experimenter's headphones bookended by sine tones signifying beginning and end of recording.
8	Press to activate the voice trigger.
9	[*Key 10 to be moved here*]
10	Press to activate/de-activate LogFile-recording. [*to be moved to 9*]
↑	Press to move up.
↓	Press to move down.
←	Press to move left.
→	Press to move right.

7. Performing an experiment

Now you will be ready to start using RSE. This section allows you to familiarise yourself with RSE. Table 8 gives the settings used. Table 9 illustrates how RSE is operated by going through the different manoeuvres used in the Demo1 experiment included in the RSE download pack. This short experiment consists of 25 line drawings (taken from the Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005, set). Rather than using an actual participant, you can act as both participant and experimenter when doing this.

Table 7. Settings for test experiment.

Setting	*Value*
Display Time	2000
Display Period	3000
Single Step	0
Guide Tones	0
Use Countdown	1
Countdown Size	300
?Prelay Time	0
Sound Trig Threshold 1	50
Sound Trig Threshold 2	10
Sound Trig Threshold 3	75
Sound Trig Threshold 4	100
Lowpass Weight	0.80
Mic. Ampl.	0 dB
Text Stimulus Size	100
Stim Size Part Scrn	100
Stim Size Exp Scrn	50
Orig Rec Window Size	2
Experimenter Channel	0

Table 8. How to use RSE.

Action	*How to*
Start LogFile recording	**Press key 9**.
Start experiment	**Press key 1**.
On trials 2, 4, 7, 8, 11, and 14 you will record the words to be used during word exchanges.	
Record word	**Press key 5** just before speech starts, to start recording, and hold the key for as long as the recording should last. The amount of trailing silence is determined by the experimenter who releases the record button at the appropriate time. About 100-150 ms of trailing silence is recommended. This trailing silence should be added to ensure that the ending of the participant's own speech will not be heard because the switchback to the participant's microphone would otherwise potentially be too early, i.e. while the participant is still speaking. The amount of trailing silence can also be adjusted using the Original Recording (normalized) waveform display (see figure 7), but do not add trailing silence going beyond the actual recording of silence.
Automatic cropping of the recording. *The recording is then automatically cropped along the time axis so that the onset of the recording matches word-onset. When a word is recorded, it is shown in the Original Recording (normalized) window (see figure 4 and description there), and the trimmed recording is shown in the Trimmed Recording window.*	
Determine if the recording is appropriate to be used during an exchange	There are two ways to determine if the recording is appropriate to be used during an exchange in the test (while the experimenter performs either of these operations, the participant receives unaltered feedback from his speech.): - (1) By visual inspection of the waveform display in the Original Recording (normalized) window. Make sure the green vertical line is exactly at the start of the word. If it is not: **Place the computer mouse indicator just at the start of the recording > hold the left mouse button > drag the indicator to the place you want the recording to stop > let go of the left mouse button.** The part of the recording you have now chosen is saved, it is shown in the Trimmed Recording window, and it can be used during a voice exchange. - (2) If Experiment channel is set to 1 and the **Experimenter and participant channel splitter** is used: **Press key 7** to listen to the recording. The recording is played back to the experimenter's left headphone, surrounded by short (29 ms) sine tones which allow the experimenter to determine that the timing of the recording is accurate.
On trials 6, 10, 12, 19, and 20 you will perform voice exchanges.	
Insert recording – Perform a voice exchange	**Press key 8** just before the participant starts uttering the word to activate the voice trigger. When the participant begins to pronounce the word, the feedback of the participant's voice is suppressed and the relevant recording is instead inserted. *The semi-automatic recording and playback procedure in the RSE system allows the experimenter a great freedom in adapting to the participant's reactions as the experiment unfolds. For example, should the participant break out laughing or be suddenly tongue-tied during a pre-specified manipulated trial, the experimenter can simply abort the voice exchange, thereby avoiding a mistimed exchange and a probable detection of the experimental procedure by the participant. Furthermore, since the experimenter activates the voice trigger, he can do so in as close proximity to the participant's utterance as possible, thereby diminishing the chances that the trigger will trigger on e.g. smacking noises.*
Automatic Break in the experiment. *After each exchange trial, the test stops (Break function) and the question "What did you just say?" is displayed.*	
Resume the test after a break	**Press key 1**.
Post-test interview	If a post-test interview is desirable, simply continue the recording (i.e. do not stop it by **pressing key 9**) and the interview can be recorded with the headset-microphone, which can be placed e.g. on a table close to the participant and experimenter.
Stop LogFile recording	**Press key 9** when the test is finished, to stop the log-file recording. The recording is then saved to the LogFile folder with the name LogFile_YYYYMMDD_HHMMSS.

8. Data analysis

[*coming soon*]

8.1. Analysis using Praat

Start Praat. Press Open > Read from file… (Ctrl+O). Find the file you want to analyse. It is shown under Objects. Mark it and press View & Edit to look at it.

9. Troubleshooting

[*coming soon*]

9.1. Reset default settings

[*Function to be inserted*]
See section 6.1.1. This should be tried before anything else.

9.2. Reset low-level settings using Regedit

In the Windows search bar, search for “regedit”. Open the Registry Editor app. Search for “BBDS Software”. Remove the whole folder ”RSE”. Start RSE again.

> **Warning!**
> Be careful not to delete or edit anything besides the “RSE” folder in the Registry Editor app, as this can cause major issues with your computer.

10. References

Boersma, P., and Weenink, D. (2017). "Praat: Doing phonetics by computer (version 6.0.33) [computer program]," http://www.praat.org (Last viewed December 9, 2020).

Breidegard, B. & Andersson, P. (1992). BBDS – A design tool for architectural evaluation and rapid prototyping of performance critical digital systems. *Proceedings of the Third International Workshop on Rapid System Prototyping,* (pp. 19-23). https://doi.org/10.1109/IWRSP.1992.243920

Breidegard, B. & Balkenius, C. (2003). Speech development by imitation. In C. G. Prince, L. Berthouze, H. Kozima, D. Bullock, G. Stojanov, and C. Balkenius (Eds.), *Proceedings of the Second International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* (pp. 57-64). Lund: Lund University Cognitive Studies.

Breidegard, B. (2007). Computer Based Automatic Finger and Speech Tracking System. *Behavior Research Methods, 39*, 824-834. https://doi.org/10.3758/BF03192975

Breidegard, B. (2013). The Minimeter – Enabling communication technology for people with severe brain injuries. *Technology and Disability, 25*, 37-47. https://doi.org/10.3233/TAD-130367

Caltenco, H. A., Breidegard, B., & Andreasen Struijk, L. N. S. (2014). On the tip of the tongue: Learning typing and pointing with an intra-oral computer interface. *Disability and Rehabilitation: Assistive Technology, 9*, 307-317. https://doi.org/10.3109/17483107.2013.823629

Franken, M.K., Hartsuiker, R.J., Johansson, P., Hall, L., Wartenberg, T., & Lind, A. (2019). Does passive sound attenuation affect responses to pitch-shifted auditory feedback? *Journal of the Acoustical Society of America, 146*, 4108-4121. https://doi.org/10.1121/1.5134449

Kim, K. S., Wang, H., & Max, L. (2020). It's about time: Minimizing hardware and software latencies in speech research with real-time auditory feedback. *Journal of Speech, Language and Hearing Research, 63*, 2522-2534. DOI: https://doi.org/10.1044/2020_JSLHR-19-00419

Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. Acta Psychologica, 119, 159-187. https://doi.org/10.1016/j.actpsy.2005.01.002

Shuster, L. I., & Durrant, J. D. (2003). Toward a better understanding of the perception of self-produced speech. Journal of Communication Disorders, 36, 1-11. https://doi.org/10.1016/S0021-9924(02)00132-6