INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

Individual differences in implicit attention to phonetic detail in speech perception

Natalie Lewandowski[1], *Daniel Duran*[2]

[1] Independent scientist, Stuttgart, Germany
[2] Albert-Ludwigs-Universität Freiburg, Germany

natalie.lewandowska@hotmail.com, daniel.duran@germanistik.uni-freiburg.de

Abstract

We present a study on the interactions between implicit attention to acoustic-phonetic detail in speech and individual differences (IDs). Attention to phonetic detail was assessed with acoustically manipulated speech stimuli within a computer game, an alternative to regular highly-controlled categorization tests. Twenty-two native German speakers (11f) completed the game and further tests including individual attention test measures (e.g. Simon Test), the BFI-10 (short version of the Big Five Inventory), and a Self-monitoring Test (need for social approval). With this study, we contribute to the understanding of the processes underlying human speech perception and the impact of cognitive and personality features on the attention to phonetic detail. Our results show that the general (non-verbal) attention capacity (mental flexibility, inhibition), interacts with implicit attention to phonetic detail. Furthermore, IDs in personality, such as *sensitivity to social cues* or *conscientiousness* significantly add to the effects. Understanding these interactions, especially arising in an intuitive and non-explicit study design, is an important step on the way towards explaining not only the influence of IDs on attention to phonetic detail, but also the dynamics of speech interaction (e.g. phonetic convergence).

Index Terms: speech perception, gamification, attention, personality

1. Introduction

Individual cognitive differences are an important factor in speech processing which has been found to impact speech perception and production, in L1 as well as L2 settings. A considerable amount of literature is concerned with the influence of attention (switching) skills on speech (e.g. see [1, 2, 3, 4, 5]).

Lev-Ari and Peperkamp [3], for instance, employed a manual version of the Stroop Test [6] and a Simon Test [7] (see section 2.2) within their study and correlated the results to lexical decisions for VOT-manipulated target words in French and a sentence reading task. A stronger priming effect in perception was found for subjects with lower inhibitory skills, but no correlation was identified for production. However, Lev-Ari and Peperkamp [2] did show an influence of inhibition on production. Late bilinguals with lower inhibitory skill (measured with a Retrieval-Induced-Inhibition Test) in their study tended to produce voiceless plosives in their dominant language with similar VOTs to their nondominant language. Darcy and colleagues [1] suggested that motor-based tests, such as the Simon Test, might be well suited to capture production differences in pronunciation. For phonetic convergence in dialogs, where both perception and production are involved, the amount of adaptation was found, a.o., to be negatively correlated to the switch costs in a Simon Test [4, 8]. The smaller switch costs in the Simon Test, indicating less of a reaction time drop when switching between the congruent and the incongruent trials [7], proved to increase the degree of phonetic convergence of L1 German speakers to their English L1 conversational partners. The more mentally flexible a speaker was, the more convergence she or he exhibited within the L2 dialog setting. This is also in line with [9] who report a correlation between the amount of imitation and a test variable from the Autism Spectrum Quotient (AQ) [10] which can be related to attention-switching. Scharenborg and colleagues' study [5] demonstrated a link between a perceptual learning effect and attention switching control. Lower attention-switching skills were related to an increased perceptual learning effect, where subjects tended to rely more on lexical information in a word than on the ambiguous sound of the target word they were presented with. This allows to speculate that listeners with higher attention control directed more of their attention towards sound information and were not adapting their phoneme categories to match the ambiguous sound as readily as the lower skilled group in their study [5].

Regarding the impact of psychological (personality-related) individual differences (IDs), work has been focused rather on IDs as predictors of successfull L2 learning in general or the presence of a special phonetic aptitude (see e.g. [11, 12, 13]). Only a couple of studies have investigated personality in relation to phonetic adaptation [9, 4]. Yu and colleagues have found a link between openness and imitation [9], whereas Lewandowski and Jilka [4] report a positive effect on the amount of phonetic convergence for *openness* and – quite counterintuitively – *neuroticism*, and a negative impact of the *Behavior Inhibition Score* (BIS, from the BIS/BAS scale, see [14]).

To further establish the links between cognitive and psychological factors and phonetic adaptation, Schweitzer and colleagues [15] designed the GECO2 database to include not only dialog recordings and a new form of testing participants' perception skills (the GDX game framework, see [16, 17]) but also provide data on the subjects' personalities, self-monitoring behavior, and cognitive skills – including attention. GECO2 also allows us to test whether a person's perception skills, here given as the successful attention paid to fine phonetic detail within the GDX game, can be linked to general cognitive and personality features. We hypothesize that the better the general attention skill of the subjects, the better their implicit attention to fine phonetic detail. Given its reported links to phonetic adaptation, certain personality features, as, for instance, *openness*, might also be expected to impact attention to fine phonetic detail.

2. Method

A number of psychological and cognitive tests was carried out within GECO2 [15] in addition to the perception test (GDX).

  http://dx.doi.org/10.21437/Interspeech.2019-2989

2.1. Personality tests

A self-monitoring test [18] and the short version of the Big Five Personality Questionnaire – the BFI-10 [19] – were administered. The self-monitoring scale is a German adaptation of earlier work of Snyder (1974) and assesses the need for social approval. It provides a total score and additional scores on the four subscales *Acting*, *Extraversion*, *Other-Directedness* and *Sensitivity for Social Cues and Expressive Behavior*. The BFI-10 encompasses a ten-item questionnaire and provides scores on the well-known standard personality features *extraversion, agreeableness, conscientiousness, neuroticism*, and *openness*. Both tests were filled in on screen by the participants.

2.2. Cognitive tests

The cognitive tests within this study were two tests for attention (inhibition): a Simon and a Stroop Test. The first is also called a test for *mental flexibility* [7], which is a nonverbal test for inhibition, a cognitive component much discussed for its influence on phonological processing skills (e.g. [2, 1]). The Simon Test is based on spatial competition arising from the stimuli and requires clicking on the target appearing either on the congruent or opposite (= incongruent) side of the screen. Clicking on the incongruent targets is usually tied to a higher error rate and longer reaction times, since it requires overcoming (i.e. inhibiting) the existing spatial interference. The second was a Stroop Test [6, 20], a verbal test for attention, usually based on naming congruent or incongruent ink colors of printed color names. Here as well, the correct answer follows the inhibition of the habitual response, i.e., the reading of the color names, which is the more automatic task. Further typical inhibitory tests are the Antisaccade Test [21, 22] and the Stop-Signal-Test [23].

2.3. Gamified Discrimination Experiment Engine – GDX

In order to test implicit attention to fine phonetic detail, we employ the **G**amified **D**iscrimination E**x**periment Engine (GDX) framework which was used as one of the test scenarios in the GECO2 database[1] [24, 16]. The subjects are not instructed to pay attention to the phonetics explicitly – neither within the game nor by the experimenters prior to the test session. Rather, responding to differences in phonetic features is crucial for success within the first-person shooter game scenario. Participants in the game move through an alien invasion scenario and are supposed to tell the difference between humans and aliens in order to rescue the former and freeze (immobilize) the latter. The aliens, however, are disguised as humans, so targets can only be told apart by paying attention to the sounds they make.

The game-like scenario yields more naturalistic data as it imposes a less artificial laboratory setting upon the subjects in comparison to traditional perception tests. The latter require conscious decisions focusing all attention explicitly on the task. With respect to perception of phonetic detail in every day situations, that is a highly unnatural situation making any generalizations on human speech perception from these tests questionable. We address this problem by employing a game paradigm in which *access fluidity* (the speed and/or automaticity of connecting words to their meaning) and *attention control* (the ability to focus and refocus attention on different semantic levels, i.e. local vs. global meaning relations or sound vs. meaning) [25] are a natural part of the game, and acoustic/auditory targets can easily be tied into the scenario. The utility of GDX has been validated against a classical perception experiment [26, 27]. This framework provides a good balance between spontaneous data and full experimental control, and allows for a better distinction of participants' individual differences in a natural scenario resembling everyday speech perception.

Although the game implements a forced-choice categorization test, the evaluation of its results is more complex in comparison to classic computer-based test frameworks such as DMDX [28] or Praat's "MFC" [29], for example. There are multiple ways to evaluate responses during experimental trials because the player not only has to respond with the left of right mouse button, but has to aim at a moving target at the same time. If the player does not hit the target with the first click she/he has to click again until the target is hit[2]. The game produces a log of all relevant events, such as mouse clicks, the position and orientation of the target agents (which trigger experimental trials), and the player in world space and on screen. It also captures data relevant for timing of each event (like frame counts as well as absolute time using low level system time measurements, or the current frame rate at which the game is rendered). For this study, we extracted the following data from the game logs:

- The onset of each trial: the time, the acoustic stimulus and its category.
- The first click after trial onset by the player: the time and the response category (mouse button).
- The *hit* (i e. final click within a trial): the time and the response category (mouse button).

Based on these data, we computed two measures: (1) the reaction time, i.e. the time difference between trial onset and hit, and (2) the relative reaction time, i.e. the time difference between the offset of the audio stimulus and the hit (taking into account the variable length of our stimuli). In total 7264 trials were evaluated. Out of these, only 202 did not hit the target with the first click (mean click number per trial = 1.032) and only in one case did the subject respond with a different mouse button on the first click in comparison to the final hit. Thus taking the hit events as the subjects response seems appropriate.

Twenty-two adult native speakers of German participated in the experiments (11 female). No participant reported any known hearing impairments. All test subjects gave their written informed consent to participate in the study. The participants' test data are part of the second **Ge**rman **Co**nversations database (GECO2) [15]. The computer game and the psychological and cognitive tests took place in a sound-attenuated psycholinguistics laboratory, with additional head-mounted AKG headphones where sound was involved (i.e. during the computer game).

2.4. Speech material

Speech stimuli for GDX were created from utterances spoken by an adult female native speaker of Standard High German. Various utterances, ranging from single words to entire short sentences were recorded in a sound attenuated booth. Their content was related only in part to the contents of the game, and most stimuli were devoid of any meaningful relation to the game storyline or the game level appearance. This was motivated by the desire to create a plausible game environment, on the one hand, and the expectation that a number of semantically unrelated stimuli may imply to the players that the *meaning* is

[1] The version of the game employed in this study was called ΨX 732.

[2] A trial may also fail, if the player does not hit the target or does not react at all to a stimulus and moves on with the game. Overall, such events were extremely rare in our experiments: 7 out of 7274 trials in this study. We omitted such trials from the present analysis.

not the essential cue, on the other hand. Four phonetic features have been manipulated using Praat [29]: increased range of the fundamental frequency (f0), heightened second vowel formant (F2), increased voice onset time of plosives (VOT) and removed lower frequencies in the spectrum of fricative sounds using a band-pass filter. The manipulated stimuli were associated within the game with one category (A) and the original utterances with the alternative category (B). The four different phonetic cue categories correspond to four experiment levels of the game, each of which implements an A–B categorization task. This means that the subjects always had to figure out only one specific difference in pronunciation per game level and never encountered more of them mixed together.

2.5. Procedure

Participants were seated for the perception test (the game) in a quiet, window-less psycholinguistics laboratory in front of a computer screen. A standard office keyboard and a mouse were used as input devices. Participants were wearing headphones. The game was run on a Windows platform. The display resolution of the game used for this study was 1680×1050 pixels. At the beginning of the session, participants were instructed to adjust the sound volume to a comfortable level, such that they can "hear well". It was not pointed out to the participants that they had to rely on the speech stimuli in order to distinguish the two target categories. The time limit for each experiment level was set to 12 minutes. A session with the game thus lasted for approximately one hour in total (including a training level where participants were familiarized with the game controls, not the stimuli). The trials within each experiment level have been randomized without repetitions of stimuli for each participant. During the entire session, the player was alone in the room and the experimenter waited in the room next door.

3. Analysis

All analyses were carried out with R version 3.5.0 [30] and all data preprocessing and visualization was performed with packages from the tidyverse environment [31]. The linear mixed modeling is based on the *lmerTest* package [32], which automatically provides the significance values (p-values) reported in the datasets. Outliers were removed from the data by setting cut-off points of 2 SD away from the mean of the distribution. Due to extreme values in the two cognitive tests, one participant had to be excluded from the analysis.

The dependent variable in the first model was the mean reaction time for a hit (click) in the game (mean.RT.hit). The model selection procedure was manual and included step-wise anova tests to verify the contribution of the respective fixed factor. Visual inspection of the distribution proved no obvious deviations from normality. *Subject* was included as a random factor. None of the two cognitive variables proved to significantly impact the reaction time, however the type of hit (i.e., correct or incorrect hit) and the personality feature *agreeableness* influenced the mean reaction times for hits in our subjects. The factor *agreeableness* interacted with *correct.hit* (AIC 652.2, BIC 662.9, logLik -320.1, deviance 640.2, df.resid 38) lengthening reaction times considerably in case of correct hits (correct.hitT stands for correct.hitTRUE). The influence of the interaction was tested and could be confirmed in an anova: adding the factor *agreeableness* into the interaction has a significant impact on the model (Chi square 4.2975, Pr(>Chisq) 0.03817*) compared to treating it as a simple fixed factor. Table 1 shows the estimates of the first analysis. The analysis shows that correct hits were in general quicker than wrong hits, however for subjects who scored high on agreeableness, the reaction time was slightly prolonged.

Table 1: *Fixed factor analysis of the linear mixed model on mean reaction times for hits.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	418.81	378.06	23.14	1.11	0.28
correct.hitT	-267.30	120.10	22.00	-2.23	0.04
Agreeable	-19.90	81.11	23.14	-0.25	0.81
correct.hitT: Agreeable	56.13	25.77	22.00	2.18	0.04

In the second analysis we estimated the influence of the cognitive and personality factors on the number of correct hits in the game (see Table 2). The proportion of correct hits was used as the dependent variable for a linear model. The highest R squared value was achieved when fitting a full model, including all features as simple factors and two interactions (guided by visual inspection of the plots and subsequent model comparison). The full model (maximal model), given in Table 2, shows a negative effect for *Sensitivity for Social Cues and Expressive Behavior*, *Conscientiousness* and an effect for the switch costs in the Simon Test (model parameters: Residual standard error 0.03601 on 7 degrees of freedom, Multiple R-squared 0.8749, Adjusted R-squared 0.6247). The seemingly present interaction of *Conscientiousness* and *Gender* visible in Figure 1 (page 4) is not significant and its inclusion did not improve model fit. Figure 2 shows the relationship of *Sensitivity* from the Self-monitoring Test and the proportion of correct hits in the GDX game. Figure 3 illustrates the effect for the switch costs in the Simon Test. Here, the negative effect is driven by female participants. Those who showed a higher attention switching capacity (by displaying lower switch costs in the Simon Test), landed more correct hits in the perception test (the game).

Table 2: *Results for the linear model with the proportion of correct hits as the dependent variable.*

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.77	0.11	6.83	0.00**
Acting	0.00	0.01	0.80	0.45
Otherdirect	0.01	0.01	1.02	0.34
Extraversion	0.01	0.01	1.20	0.27
Sensitivity	-0.04	0.01	-4.19	0.00**
Extroverted	0.01	0.01	1.57	0.16
Agreeable	-0.01	0.01	-1.00	0.35
Conscientious	-0.03	0.01	-2.93	0.02*
Open	-0.01	0.01	-1.83	0.11
RT_Simon	-0.00	0.00	-2.58	0.04*
genderm	-0.11	0.10	-1.13	0.30
RT_Stroop	0.00	0.00	1.12	0.30
Neurotic	0.01	0.02	0.62	0.56
RT_Simon:genderm	0.00	0.00	1.63	0.15
genderm:Neurotic	0.02	0.02	1.10	0.31

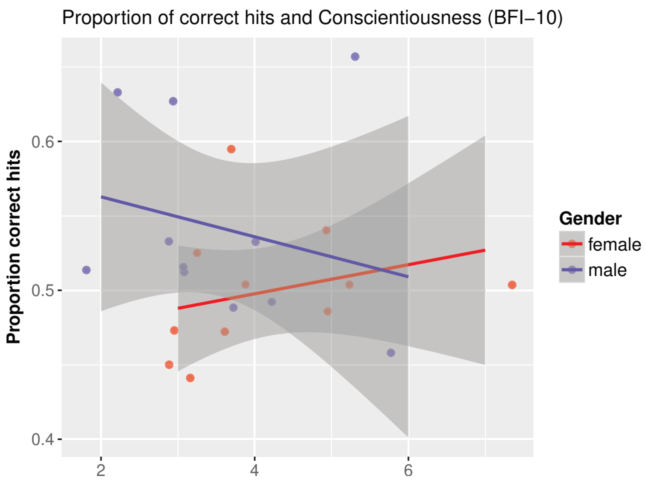

Figure 1: *Proportion of correct hits and Conscientiousness.*

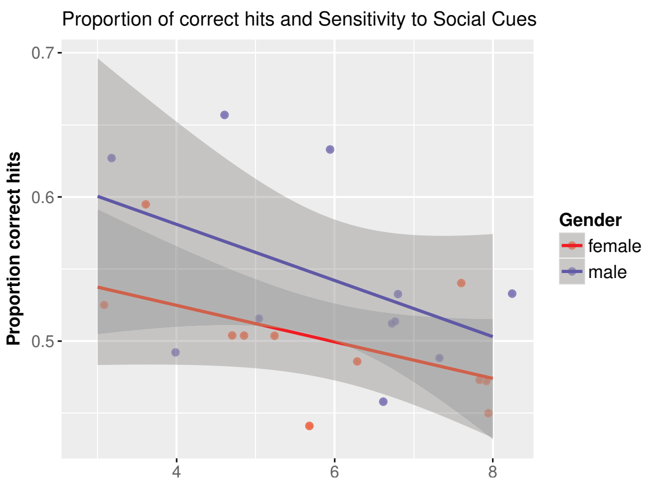

Figure 2: *Proportion of correct hits and Sensitivity to Social Cues and Expressive Behavior.*

4. Discussion & Conclusion

The results show an interesting interaction between the personality factor *agreeableness* and the reaction times for correct hits in the gamified perception test. It seems as if subjects who tend to be more agreeable, harmony-driven and easier to get on with, took slightly longer to decide to which category the sound belongs, however only in cases where they actually decided correctly.

The analysis of the proportion of correct hits in the gamified perception experiment showed that attention switching skills indeed mattered for perception in the L1. Test subjects with higher attention switching capacity, i.e. with a higher mental flexibility, proved to score more often correctly in the game. This advantage within the perception test held largely true only for the female subjects. The male subjects did not seem to benefit from higher mental flexibility in the same way. It is advisable that future studies take gaming experience into account here, since the observed gender difference might be related to disparities in gaming proficiency/ previous gaming experience. We suspect that in case of highly proficient gamers the potential benefit from generally higher attention switching skill might in

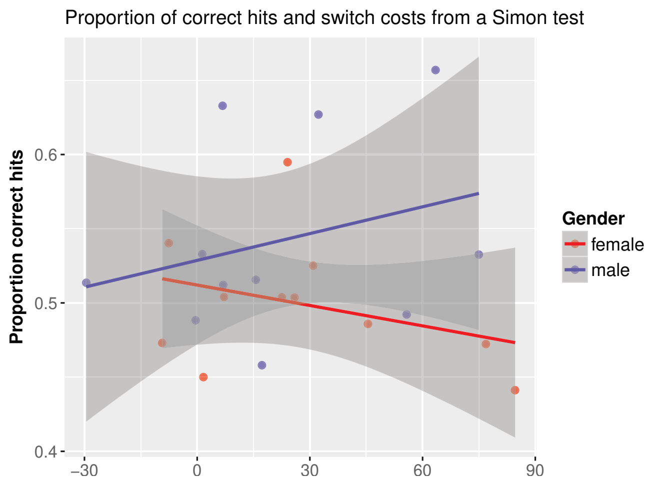

Figure 3: *Proportion of correct hits and the interaction of switch costs in the Simon Test and gender.*

fact disappear. A skilled attention controller is generally said to be better able to switch between bottom-up and top-down signals or weigh them more situation-appropriately [33, 34, 5]. Our participants with a higher mental flexibility might have figured out the better strategy and listened more attentively to the sound (the acoustic signal) and paid less attention toward the (potentially confusing) meaning of the words being uttered. The impact of attention switching was also previously found for phonetic convergence [4].

The negative effects for *Sensitivity to Social Cues and Expressive Behavior* and *Conscientiousness* are partially surprising. Whereas a higher conscientiousness could have in principle induced higher uncertainty and potentially more errors as a consequence, it would be interesting to include a further measure to corroborate these results. The addition of a BIS/BAS scale (Behavior Inhibition/Behavior Activation), as reported in [4], could shed more light onto the exact mechanisms. The decrease in correct hits for subjects scoring high on the Sensitivity Scale could be caused by the aforementioned higher reliance on the lexical content of the acoustic signal or a focus on more global sound properties as vocal tone. If these listeners expected to find relevant cues within the meaning of the words and phrases in the game or within the tone, their attention could have been removed from the important dimension here – the fine-grained spectral properties of the sound.

5. Acknowledgements

The experiments have been carried out within the collaborative research centre SFB 732 (A4) funded by Deutsche Forschungsgemeinschaft (DFG) from 2014–2018.

6. References

[1] I. Darcy, J. C. Mora, and D. Daidone, "The role of inhibitory control in second language phonological processing," *Lang Learn*, vol. 66, no. 4, pp. 741–773, 2016.

[2] S. Lev-Ari and S. Peperkamp, "Low inhibitory skill leads to non-native perception and production in bilinguals' native language," *J Phon*, vol. 41, no. 5, pp. 320–331, 2013.

[3] ——, "The influence of inhibitory skill on phonological representations in production and perception," *J Phon*, vol. 47, pp. 36 – 46, 2014.

[4] N. Lewandowski and M. Jilka, “Phonetic convergence, language talent, personality and attention,” *Frontiers in Communication*, vol. 4, p. 18, 2019. [Online]. Available: https://www.frontiersin.org/article/10.3389/fcomm.2019.00018

[5] O. Scharenborg, A. Weber, and E. Janse, “The role of attentional abilities in lexically guided perceptual learning by older listeners,” *Atten Percept Psychophys*, vol. 77, no. 2, pp. 493–507, 2015.

[6] J. R. Stroop, “Studies of interference in serial verbal reactions.” *J Exp Psychol*, vol. 18, no. 6, pp. 643–662, 1935.

[7] J. L. Craft and J. R. Simon, “Processing symbolic information from a visual display: Interference from an irrelevant directional cue.” *J Exp Psychol*, vol. 83, no. 3, Pt.1, pp. 415–420, 1970.

[8] N. Lewandowski, “Talent in nonnative phonetic convergence,” Doctoral Dissertation, Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, Germany, 2012.

[9] A. C. L. Yu, C. Abrego-Collier, and M. Sonderegger, “Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and “autistic” traits,” *PLoS ONE*, vol. 8, no. 9, p. e74746, 2013.

[10] S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin, and E. Clubley, “The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians,” *J Autism Dev Disord*, vol. 31, no. 1, pp. 5–17, 2001.

[11] R. Ellis, “Individual differences in second language learning,” in *The Handbook of Applied Linguistics*. Blackwell Publishing Ltd, 2004, pp. 525–551.

[12] Z. Dörnyei and P. Skehan, “Individual differences in second language learning,” in *The handbook of second language acquisition*, ser. Blackwell handbooks in linguistics, C. Doughty and M. H. Long, Eds. Malden, Mass.: Blackwell, 2003, vol. 14, pp. 589–630.

[13] X. Hu and M. S. Reiterer, “Personality and pronunciation talent in second language acquisition,” in *Language Talent and Brain Activity*, ser. Trends in Applied Linguistics, G. Dogil and S. M. Reiterer, Eds. Berlin: de Gruyter, 2009, pp. 97–129.

[14] A. Strobel, A. Beauducel, S. Debener, and B. Brocke, “Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White,” *Zeitschrift fr Differentielle und Diagnostische Psychologie*, vol. 22, no. 3, pp. 216–227, sep 2001.

[15] A. Schweitzer, N. Lewandowski, and D. Duran, “Attention, Please! Expanding the GECO Database,” in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, The Scottish Consortium for ICPhS 2015, Ed., Glasgow, UK, 2015, paper number 620.

[16] D. Duran, N. Lewandowski, and A. Schweitzer, “A 3D computer game for testing perception of acoustic detail in speech,” in *Proceedings of Meetings on Acoustics*, vol. 28, 2016, p. 060004.

[17] N. Lewandowski and D. Duran, “Testing speech perception today and tomorrow: serious computer games as perception tests,” in *Tagungsband der 29. Konferenz Elektronische Sprachsignalverarbeitung (ESSV)*. Ulm: TUDpress, 2018, pp. 232–239.

[18] G. Collani and S. Stürmer, “Deutsche Selbstüberwachungsskala,” ZIS - GESIS Leibniz Institute for the Social Sciences, Tech. Rep., 2011.

[19] B. Rammstedt and O. P. John, “Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german,” *Journal of Research in Personality*, vol. 41, no. 1, pp. 203–212, 2007.

[20] G. D. Logan, N. J. Zbrodoff, and J. Williamson, “Strategies in the color-word Stroop Task,” *Bull Psychon Soc*, vol. 22, no. 2, pp. 135–138, 1984.

[21] P. E. Hallett, “Primary and secondary saccades to goals defined by instructions,” *Vision Res*, vol. 18, no. 10, pp. 1279–1296, 1978.

[22] A. Miyake and N. P. Friedman, “The nature and organization of individual differences in executive functions,” *Curr Dir Psychol Sci*, vol. 21, no. 1, pp. 8–14, 2012.

[23] G. D. Logan, *On the ability to inhibit thought and action: A users guide to the stop signal paradigm*, ser. Foundations of neuropsychology series. Academic Press, 1994, pp. 189–239.

[24] L. Lange, B. Pfeiffer, and D. Duran, “ABIMS auditory bewildered interaction measurement system,” in *Proceedings of the 16th Annual Conference of the International Speech Communication Association (Interspeech)*. Dresden: ISCA Archive, 2015, pp. 1074–1075.

[25] N. Segalowitz, “Access fluidity, attention control, and the acquisition of fluency in a second language,” *TESOL Quarterly*, vol. 41, no. 1, pp. 181–186, 2007.

[26] J. E. Krüwald, D. Duran, and N. Lewandowski, “Gamification in the phonology lab,” Lisbon, Portugal, 2018. [Online]. Available: http://labphon16.labphon.org

[27] N. Lewandowski and D. Duran, “Testing speech perception today and tomorrow: serious computer games as perception tests,” in *Elektronische Sprachsignalverarbeitung 2018: Tagungsband der 29. Konferenz*, A. Berton, U. Haiber, and W. Minker, Eds. Dresden: TUDpress, 2018, pp. 232–239.

[28] K. I. Forster and J. C. Forster, “DMDX: A Windows display program with millisecond accuracy,” *Behavior Research Methods, Instruments, & Computers*, vol. 35, no. 1, pp. 116–124, Feb. 2003.

[29] P. Boersma and D. Weenink, “Praat: doing phonetics by computer,” 2015. [Online]. Available: http://www.praat.org/

[30] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018. [Online]. Available: https://www.R-project.org/

[31] H. Wickham, *tidyverse: Easily Install and Load the 'Tidyverse'*, 2017, R package version 1.2.1. [Online]. Available: https://CRAN.R-project.org/package=tidyverse

[32] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, “lmerTest package: Tests in linear mixed effects models,” *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017.

[33] S. L. Mattys and L. Wiget, “Effects of cognitive load on speech recognition,” *J Mem Lang*, vol. 65, no. 2, pp. 145–160, 2011.

[34] S. L. Mattys, F. Seymour, A. S. Attwood, and M. R. Munafò, “Effects of acute anxiety induction on speech perception,” *Psychol Sci*, vol. 24, no. 8, pp. 1606–1608, 2013.