

Nautilus Speaker Characterization (NSC) Corpus

Laura FERNÁNDEZ GALLARDO

Quality and Usability Lab, Technische Universität Berlin, Germany

October 12, 2017

Purpose:	speaker characterization
Number of speakers:	300 speakers; 126 male, 174 female
Speakers' age:	18 to 35 years old
Speakers' language:	German (mother tongue, no marked dialect/accents)
Speech type:	conversational; scripted, semi-spontaneous, and spontaneous
Size and duration:	13.1 GB, 155 hours of speech
File format:	1-channel audio/flac, 48 kHz, 16 bit speakers: microphone AKG C 414B-XLS interlocutor: headset Sennheiser HMD 46
Metadata:	demographic information and self-assessed personality
Labels:	34 interpersonal speaker attributions (continuous numeric) 34 voice descriptions for selected speakers (continuous numeric)
Data owner:	Laura Fernández Gallardo
ISLRN	157-037-166-491-1

1 Corpus Overview

A total of 300 German speakers (126 males and 174 females) have participated in the recordings of clean microphone conversational speech. This document describes the speech material released in the Nautilus Speaker Characterization (NSC) corpus as well as the procedures to record and to label the speech. The recordings were performed in the acoustically-isolated room *Nautilus*, which gives name to this database.

The main motivation for the compilation of this speech corpus was to support the current research on the detection of speaker social characteristics, such as personality, charisma, and voice attractiveness. More precisely, it is investigated how different degradations introduced by channel transmissions affect the subjective perception and automatic recognition of the speakers' social traits. A sampling rate of at least 32 kHz is required for super-wideband (SWB, 50–14,000 Hz) conditions to be studied.

The NSC corpus comprises clean microphone recordings of four scripted and four semi-spontaneous dialogs that simulate telephone call inquiries. Additionally, spontaneous neutral, emotional (predominantly excitement or frustration), question and statement utterances have been produced by the speakers casually interacting with their interlocutor. The language of all recordings is German. Section 2 describes the speech uttered by the recruited participants, whose demographic data is presented in Section 3 along with information regarding the recording sessions. Details of the setup employed for voice recording can be read in Section 4.

The speech corresponding to one of the semi-spontaneous dialogs was later evaluated by 114 independent listeners with respect to 34 interpersonal speaker characteristics (SC), such as *likable*, *attractive*, *competent*, *childish*, etc. The labeling process is detailed in Section 5. After a dimensionality reduction of the 34 perceived speaker attributes to five factors [1], a set of 20 “extreme” speakers have been selected based on their evaluated warmth–attractiveness and labeled according to 34 naïve voice descriptions (VD), such as *bright*, *creaky*, *articulate*, *melodious*, etc., as indicated in Section 6.

The entire corpus material (13.1 GB of data) comprises 9192 minutes of speech (flac), the files employed as stimuli for SC- and VD-labeling (flac, 115 minutes), csv files with speakers' turn tags, listeners' ratings using

the SC and VD questionnaires, SC and VD items–dimensions information, factor scores derived from the conducted factor analyses, speakers’ metadata, and database documentation (this document). The folder structure and naming conventions of this corpus can be found in Section 7. The NSC data have been publicly released under license agreement for exclusive use with non-commercial scientific research and teaching purposes. All speakers gave their consent to participate in the recordings by reading and signing the document *Einwilligungsform_Sprachaufnahmen* (see Appendix A).

The material provided in the NSC corpus is expected to be of broad interest to phoneticians and speech scientists working on the perceptual and acoustic correlates of personal attributes. Speech and prosody production and conversational behavior in human-human interactions can be studied by analyzing speaker’s and interlocutor’s turns of semi-spontaneous and spontaneous speech. The NSC data may also be adequate for other speech-related research requiring high-quality clean recordings in German.

2 Speech Contents

The recorded speech material, from each speaker individually, comprises eight dialogs (four scripted and four semi-spontaneous) and additional spontaneous interactive speech uttered during the recording session: neutral and emotional excerpts, and questions. The speakers were also asked to sustain the vowel ‘a’ for at least three seconds repeated times. This sustained vowel ‘a’ was recorded with the purpose of conducting acoustic-phonetic analyses.

The dialogs simulated telephone calls held with an interlocutor, who conducted the recording session with the recruited speaker. Two student workers undertook this task and performed as well as speakers in other recording sessions. Their pseudonyms are: *amsterdam* (interlocutor in 279 sessions) and *aden* (interlocutor in 21 sessions); both are female. Hence, each session corresponds to the material from one individual speaker and his/her interlocutor. There were 24 sessions in which the interlocutor *amsterdam*’s speech was, regrettably, not recorded.

– *Scripted dialogs*

The speakers were asked to read the exact given text as naturally as possible, yet without emotions or exaggerated friendliness. The interlocutor always played the role of a contact person or agent, while the speaker’s turns involved different inquiries about some information. The dialogs topics were:

1. Health insurance inquiry
2. Mobile phone rate plan inquiry
3. Car rental inquiry
4. Real estate agency inquiry

The full scripted dialog texts are provided in *Dialogskripte* of Appendix A. Several attempts were needed for each dialog turn, in order to maintain the required wording. This fact elicited sometimes spontaneous emotional speech (e.g. excitement, frustration).

All speakers’ utterances are provided, named by the corresponding dialog turn and numbered by their order of occurrence. Most utterances are incomplete or muddled but there is at least one with correct wording and satisfactory naturalness for each dialog turn.

– *Spontaneous dialogs*

The speakers were confronted with four different conversational scenarios employed to elicit semi-spontaneous dialogs. These scenarios are extracted and adapted from the Short Conversation Tests of [2, 3] and can be found in *Szenarien_Spontane_Konversationen* of Appendix A. Again, the speakers played the client’s role and the interlocutor the agent’s role. The topics were:

5. Car rental booking
6. Pizza order
7. Book from the library
8. Doctor’s appointment

For each semi-spontaneous dialog, individual audio files are provided with sorted speaker turns, with inserted between-turns silences of duration 0.7 s. Unsatisfactory turns (unnatural, wrong, or noisy), have been removed and only the amended speaker turns are kept in these files. The removed turns and the interactive speech are kept in the *Interactions*, as detailed in next paragraph.

– *Interactions*

For the sessions where the interlocutor’s speech was recorded, the complete speaker and interlocutor speech is provided in two separate audio files from the start of Dialog 5 until the end of Dialog 8. This passage includes all dialog turns, interlocutor’s indications and, for most of the speakers, spontaneous neutral, emotional, speech and questions, and/or turn repetitions/amendments. Timestamps and tags that define each speaker’s turn have been written in an accompanying csv-file as specified in Table 1. The spontaneous utterances, which do not belong to the dialogs, are subject to background noises (mostly papers rustling) and might not be directed to the microphone.

tag	description
5	Dialog 5 (semi-spontaneous, car rental booking)
6	Dialog 6 (semi-spontaneous, pizza order)
7	Dialog 7 (semi-spontaneous, book from the library)
8	Dialog 8 (semi-spontaneous, doctor’s appointment)
d	(repeated) semi-spontaneous dialog turn. The unnatural, wrong, or noisy dialog turns that had to be repeated or amended.
s	Neutral spontaneous speech. Utterances with neutral prosody that don’t belong to dialogs, produced with no script.
e	Emotional spontaneous speech. Excerpts high arousal where the speaker is excited and amused by the recording situation or by having uttered wrong speech turns. In some cases, the speaker is frustrated of needing many turn repetitions.
q	Spontaneous questions about dialog contents or recording procedure. The interlocutor’s response is included in the <i>Interaction</i> .
f	Spontaneous short feedback. The speaker’s acknowledgments indicating his/her understanding of the the interlocutor’s guidance. Generally short segments: ‘ok’, ‘ja’ (yes), ‘in Ordnung’ (alright), etc.

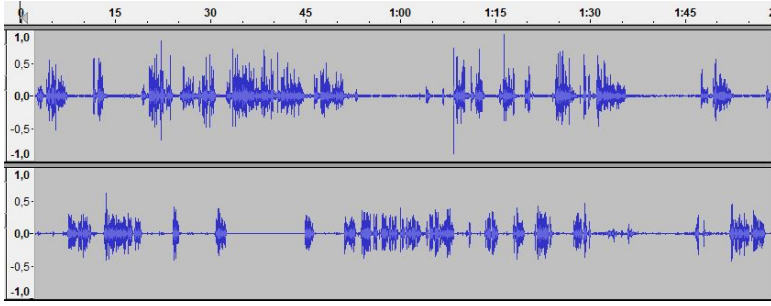
Table 1: *Tags given to speech segments of the recorded speaker. They can be found in the csv-file of interactions.*

For the remainder of this document, the set of three files (interlocutor speech (flac), speaker speech (flac), and tags+timestamps (csv)) will be referred to as *interaction*. All provided interactions can be found in a dedicated folder of the database structure for each speaker (Section 7). An example of an interaction can be seen in Figure 1, which shows the start of the flac- and csv-files corresponding to the interactions when recording spontaneous dialogs in one of the sessions.

Additional (shorter) interactions have been created in the events of spontaneous neutral and emotional speech and questions occurring before the start of Dialog 5 and after the end of Dialog 8 (segments tagged as ‘s’, ‘e’, ‘q’, or ‘f’).

When the interlocutor’s recording was not available (for 24 sessions), the corresponding segments tagged as ‘d’, ‘s’, ‘e’, or ‘q’ have been concatenated in audio files, inserting silences of duration 0.7 s between segments. This is referred to as *concatenation*, and more details can be read in Section 7.

A semi-automatic voice segmenting tool was developed and used to delimit and to arrange speech segments. Only speaker utterances with energy over a a certain level threshold set empirically were tagged as indicated in Table 1 or as scripted dialog turns. Non-speech events such as noises or clearing up throat were left out from the final set of segments. The tagging result was reviewed but may be subject to human error.



(a) First track: ‘m243_saotome_spontaneousdialogs_interlocutor.flac’. Second track: ‘m243_saotome_spontaneousdialogs_speaker.flac’.

#	tag	timestart_s	timeend_s
1	5	7.329	11.676
2	5	13.058	19.602
3	s	23.980	25.606
4	5	30.780	33.060
5	5	44.737	46.808
6	5	51.009	53.433
7	5	53.679	64.028
8	s	64.028	68.978
9	5	73.159	75.942
10	5	77.635	80.046
11	5	81.325	84.458
12	5	87.137	90.534
13	f	106.631	108.604
14	6	111.893	117.951

(b) Content of the csv-file (timestamps and tags).

Figure 1: Start of the interaction in spontaneous dialogs for speaker with pseudonym ‘saotome’ (male, ID = 243). The 3rd and 8th turns are tagged as ‘s’, indicating that this speech does not belong to the final version of the dialog. Upon finishing Dialog 5, the speaker uttered ‘ok’ (13th turn) and Dialog 6 started.

The speech file durations of scripted and semi-spontaneous turns, interactions or concatenations (depending on whether the interlocutor’s speech was recorded), and sustained vowel ‘a’ have been summed up only for the speaker’s speech. Figure 2 presents an histogram of the amount of speech from the different recording sessions. It has to be noted that this calculation of the amount of speech does not include the interlocutor’s speech nor the stimulus files employed for labeling.

The summed duration of all speakers’ speech is, approximately: 18.2 hours of scripted speech, 14.5 hours of (sorted) semi-spontaneous speech, 59.5 hours of interactions, 1.2 hours of concatenations, 0.75 hours of sustained vowel ‘a’, and 1.9 hours of speech stimuli employed for labeling - these are shorter versions of one sorted semi-spontaneous dialog (see Section 5).

The recordings of the interlocutors’ speech, provided as part of the interactions, have the same total duration as the speakers’ speech: 59.5 hours, from which 54.2 hours correspond to the interlocutor with pseudonym *amsterdam* and 5.3 hours correspond to the interlocutor with pseudonym *aden*.

Thus, approximately 115 hours of speech are provided in total in the NSC corpus.

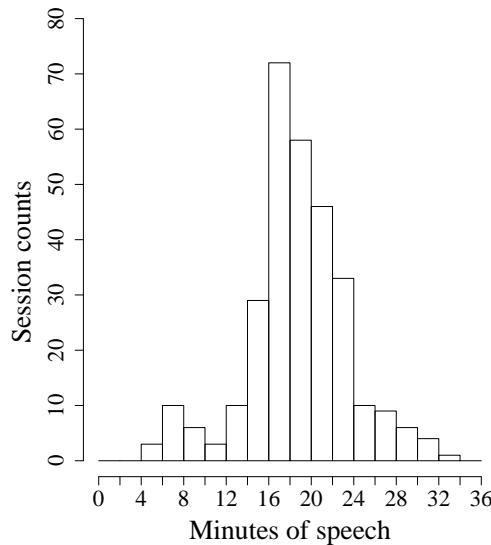


Figure 2: Distribution of the session’s speech (not including interlocutor’s speech or stimuli files).

3 Speakers’ Metadata

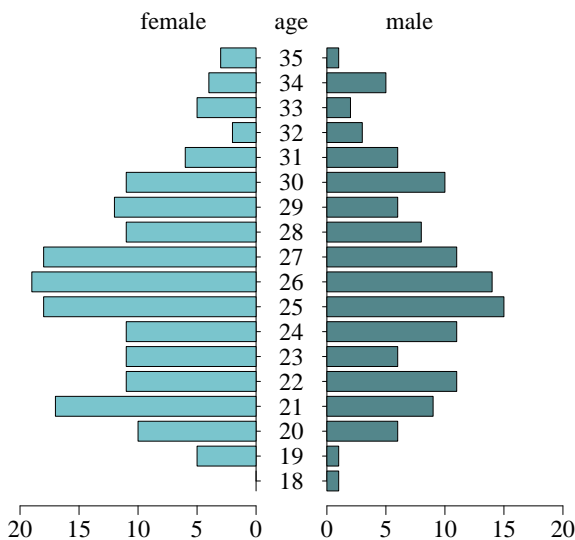
The recording sessions spanned from January 27th 2016 until March 17th 2017. Speakers were recruited via the ‘Probandenportal’ of the Technische Universität (TU) Berlin, mailing lists and paper announcements. The requirements for participation were:

- Native German speaker without marked regional dialect or accent
- 18–35 years old
- No hearing or speech disorders
- No sickness or cold at the time of the recording

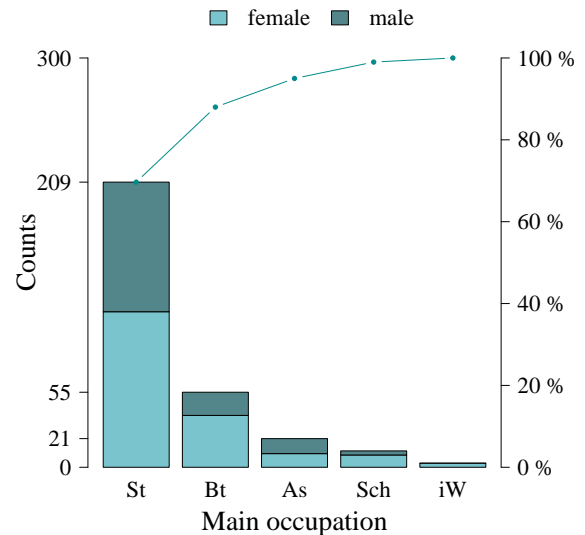
Attention was paid to only recruit young adult speakers (age: 18–35) whose mother tongue is German and who do not have a marked accent dialect distinctive of the region they come from. In this manner, low bias in the subjective impressions of the speakers was pursued.

Each recording session had an approximate duration of 45 minutes, including introduction, reading and signing the consent agreement, reading the general recording instructions (Anleitungen.Sprachaufnahmen, see Appendix A), recordings of the eight dialogs, and providing socio-demographic data and self-assessments of personality. Of course, there was no time pressure for the participants to finish the recordings within the allocated 30 minutes. The speakers were monetary compensated after their participation.

The socio-demographic data collected for each speaker comprise: age; gender (age–gender distribution is shown in Figure 3a); place of birth; chronological places of residence and duration of stay; place of birth of the mother; place of birth of the father; highest education level; educational background; main occupation (distribution shown in Figure 3b); past occupations (if any); years of work experience (if any).



(a) Age distribution for male and female speakers.



(b) Main occupation for male and female speakers. “St”: Student; “Bt”: Berufstätig (employed); “As”: Arbeitssuchend (seeking employment); “Sch”: Schüler / Auszubildender (trainee); “iW”: in Weiterbildung (continuing education).

Figure 3: Population pyramid and Pareto chart illustrating the speakers’ age distribution and main occupation. “Student” and “employed” make up for 88% of the test participants. Other main occupations are “seeking employment” (7.0%), “trainee” (4.0%) and “continuing education” (1.0%).

Most of the speakers completed a questionnaire for self-assessment:

- 92 speakers filled in the Big-Five inventory (BFI)-10 personality questionnaire in German [4] employing continuous sliders from 0 to 100 to answer each of the 10 questions

- 181 speakers filled in the BFI-44 personality questionnaire in German [5] using 5-point Likert scales to answer each of the 44 questions (plus the additional 45th question of the German version of BFI)
- All these speakers (92 + 181 = 273) indicated on a continuous slider from 0 to 100 the self-assessment of their vocal attractiveness by answering to “Wie sympathisch könnten andere Ihre Stimme finden?” (“How likable could others find your voice?”)

Before the recording sessions started, the speakers were asked to carefully read information about the project and to sign and keep a copy of the consent form *Einwilligungserklärung_Sprachaufnahmen* (see Appendix A). All speakers were informed that:

- their voices will be recorded and that socio-demographic data will be collected in the session
- the data are stored and pseudo-anonymized for scientific analyses which can be published
- the data will be treated confidentially and be part of a database that may be publicly released under license agreement for exclusive use in scientific studies
- the participation is voluntary and it is possible to withdraw from the study at any time without penalty

Accordingly, all data were pseudo-anonymized: every speaker was assigned a unique city name as pseudonym¹. The mapping between the speakers’ true names and their speech files and metadata is securely stored, with the solely purpose of being able to delete a speaker’s contribution (all data files) if he/she wishes at any time.

4 Recording Setup

During the recording session the speaker sat in the acoustically-isolated room *Nautilus*, in the TEL-building of the TU Berlin (Figure 4). The room’s dimensions are 2.75 m x 2.53 m x 2.10 m, and RT60 = 0.08 s at 2 kHz. The interlocutor sat in the adjacent room *Belafonte*, subject to background noises. She listened to the speaker, gave the pertinent instructions, and acted as dialog partner by using headsets. The connections for speech recordings are depicted in Figure 5.



Figure 4: *Recording setup in the room Nautilus. In the position of a speaker is Laura Fernández Gallardo.*

The speaker’s speech was acquired by three different microphones, yet only the recordings made with the stand-up microphone AKG C 414B-XLS (frequency range 20–20,000 Hz, mounted on a boom stand) were selected for the *share* version of the NSC Corpus. The other microphones are the table microphone Audio-Technica U851R and the headset Sennheiser HMD 46. The approximate distances from each microphone to the speaker’s mouth are 35 cm, 50 cm, and 3 cm for the stand-up, table, and headset microphone, respectively. A fourth talkback headset microphone was dedicated to capture the interlocutors’ speech.

As sketched in Figure 5, the microphone signal is amplified with the RME QuadMic II Mic Preamp and then digitalized employing the RME Fireface UCX Audio Interface, conducted to room *Belafonte* via a USB port, and recorded using the software Cubase 4 with 48 kHz sampling frequency and 32-bit quantization.

After the tagging procedure (Section 2), fade-in and fade-out were applied over 0.3 s to all speaker segments, except for the speech in *interactions*.

¹the city name was randomly taken from a list of most and second most populated cities of every country in the world.

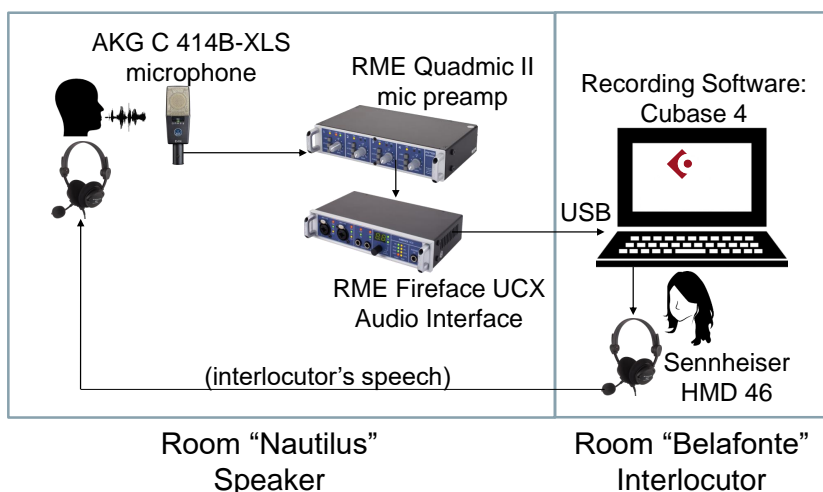


Figure 5: *Diagram of device connections between room Nautilus and room Belafonte used for the speech recordings.*

5 Labels: Interpersonal Speaker Characteristics

The sorted speaker turns of Dialog 6 (spontaneous, pizza order) were selected for speech labeling in terms of perceived speaker characteristics. Some turns were left out for some speakers in order to avoid biases in the ratings, since their wording was thought to be distinct in comparison to other dialogs with neutral content. In order to shorten the speech stimulus duration, the turns in which the speaker provides address and telephone number are removed. For the female speaker *avarua* no Dialogs 6, 7, 8 were recorded and hence her Dialog 5 was used instead. These stimuli are provided in a dedicated folder of the database structure (see Section 7).

A questionnaire involving a 34-item semantic differential rating scale has been employed to collect continuous numeric labels. The questionnaire items are based on previous research on interpersonal traits [6, 7]; the three dimensional evaluations *valence*, *activity*, *potence* [8]; frequent social and physiological attributions [9]; and aspects of longer-term interpersonal attraction [10]. A first version of this questionnaire was validated in [11], and later applied [12] employing only a small set of 15 male voices. The version used for labeling of NSC and its results are examined in [13, 1]. We refer to this questionnaire to as the speaker characteristics (SC)-Questionnaire. Its items (antonym pairs) are given in Table 2.

A series of listening tests were conducted in the frame of the *Usability Engineering (UE)* course, offered by the Quality and Usability Lab of the TU Berlin in Summer Semester 2017. A group of 114 naïve normal-hearing test participants (70 males and 44 females) performed the listening test for evaluating speaker characteristics. They were aged 24.5 years on average ($SD = 3.4$). 93 out of the 114 listeners spoke German as mother tongue, whereas the rest spoke other 10 different languages and claimed to have very good knowledge of German. Each participant listened to and rated 23.2 female and 16.4 male speakers on average. This results on an average of 15.1 ($SD = 1.2$) 34-dimensional continuous interpersonal ratings have been collected from the different listeners for each of the 300 speakers. That is, the questionnaire has been completed for each speaker by 15 different raters on average.

The test participants employed continuous sliders to indicate the extent to which they perceived each speaker characteristic. No pre-defined indicator, marks, numeric values or labels other than the antonym-pair were shown on the sliders. The listeners were asked to first complete the SC questionnaire for the set of female voices and then for the set of male voices. They could listen to each speaker dialog as many times as they wished and breaks were controlled in order to avoid tiredness. The raters wore Shure SRH240 headphones (diotic listening, frequency range 20–20,000 Hz) and performed the test in a quiet office room. The students of the UE course obtained extra credit for the course completion, yet no monetary reward was offered. The consent form and instructions of the listening test can be read in *Einwilligungserklärung_Hoerversuch* and *Anleitungen_Hoerversuch*, respectively, of Appendix A

The collected ratings are provided in the SC_ratings.csv file (see Section 7). The name of each scale rating (from 0 to 100) corresponds to the German right adjective of the antonym pair. The dimensions and item loadings determined by factor analysis as presented in [1] can be found in the SC_dimensionloadings.txt and the derived factor scores in SC_factorscores_malespk.csv and SC_factorscores_femalespk.csv, for male and female

speakers, respectively (see Section 7).

item	Antonyms (German)	English translation
1	sympathisch / unsympathisch	likable / non-likable
2	unsicher / sicher	insecure / secure
3	unattraktiv / attraktiv	unattractive / attractive
4	verständnisvoll / verständnislos	sympathetic / unsympathetic
5	entschieden / unentschieden	decided / indecisive
6	aufdringlich / unaufdringlich	obtrusive / unobtrusive
7	nah / distanziert	close / distant
8	interessiert / gelangweilt	interested / bored
9	emotionslos / emotional	unemotional / emotional
10	genervt / nicht genervt	irritated / not irritated
11	passiv / aktiv	passive / active
12	unangenehm / angenehm	unpleasant / pleasant
13	charaktervoll / charakterlos	characterful / characterless
14	reserviert / gesellig	reserved / sociable
15	nervös / entspannt	nervous / relaxed
16	distanziert / mitfühlend	distant / affectionate
17	unterwürfig / dominant	conformable / dominant
18	affektiert / unaffektiert	affected / unaffected
19	gefühlskalt / herzlich	cold / hearty
20	jung / alt	young / old
21	sachlich / unsachlich	factual / not factual
22	aufgeregt / ruhig	excited / calm
23	kompetent / inkompetent	competent / incompetent
24	schön / hässlich	beautiful / ugly
25	unfreundlich / freundlich	unfriendly / friendly
26	weiblich / männlich	feminine / masculine
27	provokativ / gehorsam	offensive / submissive
28	engagiert / gleichgültig	committed / indifferent
29	langweilig / interessant	boring / interesting
30	folgsam / zynisch	compliant / cynical
31	unaufgesetzt / aufgesetzt	genuine / artificial
32	dumm / intelligent	stupid / intelligent
33	erwachsen / kindlich	adult / childish
34	frech / bescheiden	bold / modest

Table 2: *SC-Questionnaire: 34 semantic-differential items for interpersonal speaker characteristics. The 300 speakers have been labeled on each item on a continuous scale from 0 to 100.*

6 Labels: Naïve Voice Descriptions

A set of 20 “extreme” speakers have been selected based on their salient scores on the first two SC factors (*warmth* and *attractiveness* [1], as mentioned in the previous section). The 20 speakers correspond to 5 speakers x 2 genders x 2 high/low score.

For these 20 speakers, another 34-item semantic differential questionnaire with items corresponding to naïve voice descriptions (VD) has been completed by 26 naïve normal-hearing listeners (13 males and 13 females, 26.6 years old on average, SD = 4.8 years). They were German native speakers and none of them participated in the previous speaker characteristics study. As for collecting the SC labels, the (modified) sorted speaker turns of Dialog 6 were employed as speech stimuli. Each participant listened to and rated the 20 speakers.

The voice descriptions (VD)-Questionnaire is provided in Table 3. Its items are based on previous work [14, 15, 16, 17], revised after [11, 9], and validated in [18].

As for the previous listening tests, the listeners indicated their ratings by employing continuous sliders with the antonym pairs at their ends. They could listen to each speaker as many times as they desired and also wore

item	Antonyms (German)	English translation
1	klangvoll / klanglos	sonorous / flat
2	tief / hoch	low / high
3	nasal / nicht nasal	nasal / not nasal
4	stumpf / scharf	blunt / sharp
5	gleichmäßig / ungleichmäßig	even / uneven
6	akzentfrei / mit Akzent	accented / without accent
7	dunkel / hell	dark / bright
8	leise / laut	quiet / loud
9	knarrend / nicht knarrend	creaky / not creaky
10	variabel / monoton	variable / monotonous
11	angenehm / unangenehm	pleasant / unpleasant
12	deutlich / undeutlich	articulate / inarticulate
13	rau / glatt	coarse / not coarse
14	klar / heiser	clear / hoarse
15	unauffällig / auffällig	not remarkable / remarkable
16	schnell / langsam	quick / slow
17	kalt / warm	cold / warm
18	unnatürlich / natürlich	unnatural / natural
19	stabil / zittrig	stable / shaky
20	unpräzise / präzise	imprecise / precise
21	brüchig / fest	brittle / firm
22	unmelodisch / melodisch	not melodious / melodious
23	angespannt / entspannt	tense / relaxed
24	holprig / gleitend	bumpy / smooth
25	lang / kurz	long / short
26	locker / gepresst	lax / pressed
27	kraftvoll / kraftlos	powerful / powerless
28	flüssig / stockend	fluent / halting
29	weich / hart	soft / hard
30	professionell / unprofessionell	professional / unprofessional
31	betont / unbetont	emphasized / not emphasized
32	sanft / schrill	gentle / shrill
33	getrennt / verbunden	disjointed / jointed
34	nicht behaucht / behaucht	not breathy / breathy

Table 3: *VD-Questionnaire: 34 semantic-differential items for naïve voice descriptions. The 20 selected “extreme” speakers have been labeled on each item on a continuous scale from 0 to 100.*

Shure SRH240 headphones.

The VD ratings can be found in the `VD_ratings.csv` file (see Section 7). As for the SC evaluations, each scale was named as the German right adjective of the antonym pair. The dimensions and item loadings determined by factor analysis, conducted analogously as for SC, can be found in the `VD_dimensionloadings.txt` and the derived factor scores in `VD_factorscores_malespk.csv` and `VD_factorscores_femalespk.csv`, for male and female speakers, respectively (see Section 7).

7 Database Folder Structure

The structure of the *share* version of the NSC Corpus is shown in Figure 6, where [g] is the speaker gender (‘m’ for male and ‘w’ for female); [id] is the speaker ID, from ‘001’ to ‘300’; [pseudonym] is the speaker pseudonym (city name); [ii] is the number of interaction (speakers generally have around 10 different interactions); [dd] is the dialog number, from ‘d1’ to ‘d8’; [tt] is the turn ID (as specified in *Dialogskripte* (see Appendix A)); [nn] is the order of occurrence of the scripted turn in the dialog [dd].

There are 300 folders under the *speech* folder, each of them corresponding to one speaker.

As previously mentioned, *Interactions* are only provided when the interlocutor’s speech was recorded. This is indicated in the *speech.interlocutor* column of *speakers_metadata.csv*. The interlocutor’s speech was not recorded for 24 sessions: *Interactions* are not created in this case. Instead, the other spontaneous dialog turns, emotional, and neutral spontaneous questions and statements are concatenated in separated audio files and provided in the *concatenations* folder, in place of the *interactions* folder.

The sustained vowel ‘a’ was recorded for the 181 speakers who also completed the BFI-44 personality questionnaire (from speaker ID 120 until speaker ID 300, both included). This is indicated in the *selfassessment_questionnaire* column of *speakers_metadata.csv*.

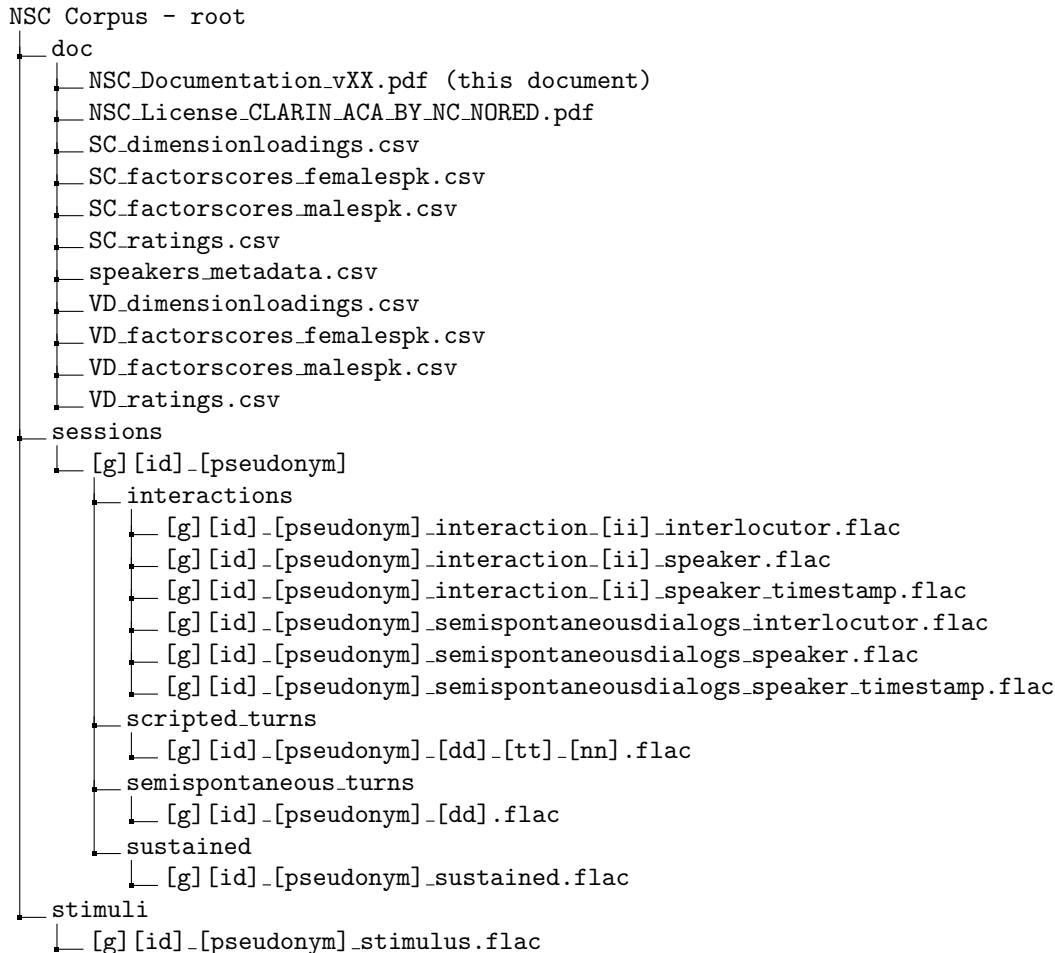


Figure 6: *NSC data folder structure.*

8 Acknowledgments

The collection of this database would not have been possible without the Grant FE 1603/1-1 by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) to Laura Fernández Gallardo and the valuable support from Quality and Usability Lab colleagues. The author is thankful to:

- Benjamin Weiss, for his advice on speech contents to be recorded, practice sessions, and contribution with the semantic differential rating scales of the SC-Questionnaire and the VD-Questionnaire
- *amsterdam* and *aden*, for conducting the recording sessions and for segmenting part of the audio data
- Christian Krüger and Christoph Hold, for their assistance with the technical setup of the recordings
- Lars-Erik Riechert, for his initial implementation of the semi-automatic voice segmenting tool
- Rafael Zequeira Jiménez for implementing the listening test setup for evaluating interpersonal speaker characteristics

References

- [1] L. Fernández Gallardo and B. Weiss, “The Nautilus Speaker Characterization Corpus: Speech Recordings and Labels of Speaker Characteristics and Voice Descriptions,” submitted to International Conference on Language Resources and Evaluation (LREC), 2018.
- [2] S. Möller, *Assessment and Prediction of Speech Quality in Telecommunications*. Boston, MA, United States: Kluwer Academic Publishers, 2000, vol. 244.
- [3] ITU-T Recommendation P.805, *Subjective Evaluation of Conversational Quality*, International Telecommunication Union, CH-Geneva, 2007.
- [4] 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.
- [5] B. Rammstedt and D. Danner, “Die Facettenstruktur des Big Five Inventory (BFI),” *Diagnostica*, vol. 67, pp. 70–84, 2017.
- [6] J. S. Wiggins, P. Trapnell, and N. Phillips, “Psychometric and Geometric Characteristics of the Revised Interpersonal Adjective Scales (IAS-R),” *Multivariate Behavioral Research*, vol. 23, no. 4, pp. 517–530, 1988.
- [7] I. Jacobs and W. Scholl, “Interpersonale Adjektivliste (IAL),” *Diagnostica – Zeitschrift für Psychologische Diagnostik und Differentielle Psychologie*, vol. 51, no. 3, pp. 145–155, 2005.
- [8] C. Osgood, G. Suci, and P. Tannenbaum, *The Measurement of Meaning*, ser. Illini Books, IB47. University of Illinois Press, 1957.
- [9] B. Weiss, D. Estival, and U. Stiefelhagen, “Studying Vocal Perceptual Dimensions of Non-experts obtained from Speaker (Dis-)Similarities assessed by Direct Comparisons,” *Acta Acustica united with Acustica*, 2017.
- [10] E. Aronson, T. D. Wilson, and R. M. Akert, *Social Psychology*, 7th ed. Prentice Hall, 2009.
- [11] B. Weiss and S. Möller, “Wahrnehmungsdimensionen von Stimme und Sprechweise,” in *Elektronische Sprachsignalverarbeitung (ESSV)*, 2011, pp. 261–268.
- [12] L. Fernández Gallardo and B. Weiss, “Towards Speaker Characterization: Identifying and Predicting Dimensions of Person Attribution,” in *Interspeech*, 2017, pp. 904–908.
- [13] —, “Perceived Interpersonal Speaker Attributes and their Acoustic Features,” in *Phonetik und Phonologie im deutschsprachigen Raum (PundP13)*, 2017.
- [14] K. R. Scherer, “Voice Quality Analysis of American and German Speakers,” *Journal of Psycholinguistic Research*, vol. 3, pp. 281–298, 1974.
- [15] W. D. Voiers, “Perceptual Bases of Speaker Identity,” *The Journal of the Acoustical Society of America*, vol. 36, pp. 1065–1073, 1964.
- [16] F. W. P. F., L. W. A. Van Herpt, and L. Boves, “Analysis of the Perceptual Qualities of Dutch Speakers Voice and Pronunciation,” *Speech Communication*, vol. 1, no. 4, pp. 315–326, 1983.
- [17] L. Boves, *The Phonetic Basis of Perceptual Ratings of Running Speech*. Dordrecht, Holland ; Cinnaminson, U.S.A. : Foris Publications, 1984.
- [18] B. Weiss, “Voice Descriptions by Non-Experts: Validation of a Questionnaire,” in *12. Tagung Phonetik und Phonologie im deutschsprachigen Raum*, 2016, pp. 228–231.

A Appendix

- **Einwilligungserklaerung_Sprachaufnahmen:** Consent form read and signed by the speakers before the recording session (in German, 3 pages).
- **Anleitungen_Sprachaufnahmen:** Instructions for the speakers about the recording session (in German, 2 pages).
- **Dialogskripte:** Texts of the scripted dialog recordings (in German, 1 page).
- **Szenarien_Spontane_Konversationen:** Scenarios for holding the spontaneous dialogs. Extracted and adapted from [3] (in German, 9 pages).
- **Einwilligungserklaerung_Hoerversuch:** Consent form read and signed by the participants of the listening test (evaluating speaker characteristics) (in German, 1 page).
- **Anleitungen_Hoerversuch:** Instructions for the participants of the listening test (evaluating speaker characteristics) (in German, 1 page).

Informationsblatt: Sprachaufnahmen zur Unterstützung sprachwissenschaftlicher Forschung

Ziel dieses Versuches ist es, eine Sammlung von Sprachaufnahmen für wissenschaftliche Zwecke zu erstellen. Es handelt sich hier um ein von der Deutschen Forschungsgemeinschaft gefördertes Projekt mit dem Titel „Menschliche Wahrnehmung und automatische Erkennung von Persönlichkeit und Sympathie eines Sprechers – Einfluss moderner Telekommunikationskanäle“. Zu einem späteren Zeitpunkt werden Versuchsteilnehmer die Sprachaufnahmen anhören und nach verschiedenen Kriterien annotieren. Anhand dieser Aufnahmen und Bewertungen sollen diese und in Zukunft noch weitere sprachwissenschaftliche Fragestellungen erforscht werden. Die gewonnenen Erkenntnisse können dann für die Entwicklung von sprachtechnologischen Anwendungen verwendet werden. Weitere Informationen zum konkreten Ziel dieser Studie geben wir Ihnen gerne nachdem die Sprachaufnahmen abgeschlossen sind.

Im Folgenden möchten wir Ihnen einige Informationen zum heutigen Termin geben.

Sprachaufnahmen

Es werden heute Sprachaufnahmen von Ihnen gemacht. Hierbei werden Ihnen zum einen Texte bzw. Textstücke vorgegeben, die Sie nachsprechen sollen. Zum anderen sollen von Ihnen freie Texte produziert werden. Es werden weiterhin soziodemographische Daten von den Sprechern erfasst. Darunter fallen Angaben zum Alter, Geschlecht, Bildungsgrad, Arbeitserfahrung, Geburtsort, Geburtsorte der Eltern und Ihre bisherigen Wohnorte.

Die Studie wird 50 Minuten in Anspruch nehmen, die Vergütung beträgt 10 €.

Vertraulichkeit

Die personenbezogenen Daten werden von den Projektmitarbeitern pseudonymisiert, bevor sie für Auswertungen zugänglich gemacht werden. Gemachte Angaben werden so verfremdet, dass kein Rückschluss auf Personen, Orte und Institutionen bzw. Organisationen möglich ist.

Die Gesprächsaufzeichnungen, die soziodemographischen Daten (Metadaten), sowie Daten, die aus zukünftigen Annotationen der Gesprächsaufzeichnungen hervorgehen, werden über eine Datenbank auch anderen anerkannten wissenschaftlichen Institutionen zur Verfügung gestellt. Wissenschaftler, welche die Datenbank verwenden möchten, müssen sich schriftlich verpflichten, wie die Versuchsleiterin, die Daten ausschließlich zu Zwecken der wissenschaftlichen Forschung und Lehre zu verwenden und nicht an Dritte weiterzugeben. In folgenden Kontexten dürfen die Daten verwendet werden:

- als Datengrundlage für Forschungsprojekte
- für die Ausbildung des wissenschaftlichen Nachwuchses z.B. bei Lehrveranstaltungen
- als Anschauungsmaterial auf wissenschaftlichen Veranstaltungen und in wissenschaftlichen Publikationen

Möglichkeiten der Rücknahme

Die Teilnahme ist selbstverständlich freiwillig. Sie werden in diese Studie also nur dann einbezogen, wenn Sie dazu schriftlich Ihre Einwilligung erklären. Sofern Sie nicht an der Studie teilnehmen möchten, während des Versuchs aus ihr ausscheiden möchten oder zu einem späteren Zeitraum die Löschung Ihrer Daten aus der Datenbank wünschen, entstehen Ihnen dadurch keine Nachteile. Bitte wenden Sie sich in diesem Fall an die Versuchsleiterin Laura Fernández Gallardo, Ph.D, (Tel: +49 175 2629100, E-Mail: laura.fernandezgallardo@tu-berlin.de) oder an den Leiter des Instituts, Prof. Dr.-Ing. Sebastian Möller (Tel: +49 30 8353 58465, sebastian.moeller@tu-berlin.de).

Bei weiteren Fragen

Sollten Sie weitere Fragen zur Studie und zur Verwendung der Daten haben, können Sie sich gerne an uns wenden (Ansprechpartner: Laura Fernández Gallardo, Ph.D, Tel: +49 175 2629100, E-Mail: laura.fernandezgallardo@tu-berlin.de).

Einwilligung des Teilnehmers

Ich erkläre, dass

- ich die Informationen zur Studie „Sprachaufnahmen zur Unterstützung sprachwissenschaftlicher Forschung“ gelesen habe und bereit bin, an der Studie teilzunehmen.
- meine Fragen ausreichend und verständlich beantwortet wurden.
- ich darüber informiert wurde, dass von mir heute folgende Daten erhoben werden: Audioaufnahmen und soziodemographische Daten.
- ich darüber informiert wurde, dass sämtliche Daten gespeichert und pseudonymisiert für wissenschaftliche Auswertungen genutzt werden.
- ich mir bewusst bin, dass die Audioaufnahmen und die Metadaten Teil einer öffentlichen Datenbank werden, die auch anderen Wissenschaftlern zu Forschungszwecken zur Verfügung gestellt werden soll. Alle Daten, die von mir für diese Studie erfasst werden, werden vertraulich behandelt. Personenbezogene Daten werden vor dem Zugriff Unbefugter geschützt.
- ich mir bewusst bin, dass die Teilnahme freiwillig ist und ich sie jederzeit ohne Angaben von Gründen widerrufen kann, ohne dass mir daraus irgendein Nachteil entsteht.
- ich eine Kopie der Informationen zur Studie und dieser Einwilligung enthalten habe.

Zudem ist mir bekannt, dass alles, was ich während der heutigen Sitzung gesehen und gehört habe, vertraulich ist und erkläre, dass ich diese Informationen nicht weitergeben werde.

Name: _____

Datum: _____

Unterschrift: _____

Versuchsleiterin:

Laura Fernández Gallardo, Ph.D
Quality and Usability Lab, Telekom Innovation Laboratories, Technische Universität Berlin
Ernst-Reuter-Platz 7, D-10587 Berlin, Germany
+49 175 2629100
laura.fernandezgallardo@tu-berlin.de

Erläuterung zu den Sprachaufnahmen

Danke für Deine Teilnahme an diesen Sprachaufnahmen.

Es ist von großer Wichtigkeit, dass Deine Stimme in sehr guter Qualität aufgenommen wird, da diese Aufnahmen die Datengrundlage für ein breites Forschungsprojekt werden sollen und die sprachwissenschaftliche Forschung für viele Jahre unterstützen sollen.

Darum haben wir strikte Voraussetzungen für die Teilnahme an dieser Studie aufgestellt: Deutsch ist deine Muttersprache, du sprichst akzentfrei, bist im Alter von 18 bis 35, und hast zur Zeit keine gesundheitlichen Beeinträchtigungen, die die Stimme betreffen.

Damit alle Sprachaufnahme der Datenbank konsistent sind, ist es wichtig, dass Du die folgenden Anweisungen beachtest. Sprich während der Aufnahmen bitte:

- Mit korrekter Aussprache
- Ohne Emotionen oder Müdigkeit in der Sprache
- Ohne übertriebene Freundlichkeit
- Lies nicht nur vor, sondern sprich so, als ob es spontane Sprache wäre
- Spiele dabei keine Rolle, sei Du selbst

Es ist von großer Wichtigkeit, dass die Sprachaufnahmen ohne jegliche Störungen oder Verzerrungen aufgenommen werden. Darum wirst Du gebeten, den gleichen Text nochmal zu wiederholen wenn:

- Einen Punkt der vorherigen Liste nicht erfüllt ist
- Irgendwelche Störungen oder Geräusche da sind (z.B. Geraschel vom Papier)
- Deine Stimme komisch klingelt (z.B. heisere oder raue Stimme)
- Bei Verzögerungen oder Unterbrechungen (husten, kichern, räuspern)

Es wird zwei verschiedene Arten von Sprachaufnahmen geben:

- Wiedergeben eines vorgeschriebenen Textes: Im Dialog mit dem Versuchsleiter gibst du einen vorgeschriebenen Dialog in natürlicher Sprache (nicht Ablesen) wieder.
- Spontane Sprache: Du führst mit dem Versuchsleiter einen kurzen Dialog, der sich an einem Rollenspiel orientiert. Du sprichst frei und in Deinen eigenen Worten.

Diese Aufnahmesession besteht aus vier Teilen:

- Teil 1: Vier vorgegebene Dialoge (Kunde und Berater): „Krankenversicherung“, „Telefontarif“, „Mietwagen“ und „Wohnungsgesellschaft“.
- Teil 2: Vier spontane Dialoge (Kunde und Berater): Pizzaservice, Arzttermin, Autovermietung, Bibliothek. Dialoge führen (mit freier Sprache) und Informationen Sammeln

Im Vorfeld erhältst Du die vorgeschriebenen Texte (Teil 1) und die Richtlinien für die spontanen Dialoge (Teil 2).

Solltest Du noch Fragen haben, melde Dich bitte bei der Versuchsleiterin, auch gerne während der Aufnahme: der Ablauf der Aufnahmesession wird sich Dir anpassen.

Wir hoffen auf gutes Gelingen und dass Dir die Aufnahmen Spaß bereiten.

Laura Fernández Gallardo

Nautilus Speaker Characterization (NSC) Corpus: Dialogskripte

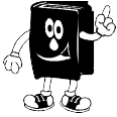
Dialog 1: Krankenversicherung	<p>Guten Tag, DKS Versicherungen. Mein Name ist Heinmüller, was kann ich für Sie tun?</p> <p>1a Guten Tag, mein Name ist Schmidt. Ich hab eine Frage bezüglich meiner Krankenversicherung. Geben Sie mir bitte Ihre Vertragsnummer durch.</p> <p>1b Meine Vertragsnummer ist die 4035. Einen Moment bitte. Alles klar, wie lautet Ihre Frage?</p> <p>1c Ich hab von meinem Hausarzt ein Rezept über 10 Physiotherapie-Anwendungen bekommen. Deshalb wollte ich mal fragen, ob Sie die Kosten in vollem Umfang übernehmen. Ich verstehe. Ich leite Sie eben zu Ihrem persönlichen Sachbearbeiter weiter.</p> <p>1d Alles klar, Dankeschön!</p>
Dialog 2: Telefontarif	<p>Die Firma Handytel, Frau Schneider am Apparat. Was kann ich für Sie tun?</p> <p>2a Guten Tag, mein Name ist Eckhard. Ich hab eine Frage zu meinem Handytarif. Alles klar, Herr/Frau Eckhard, ich öffne mal eben Ihre Vertragsdaten. Einen Moment. So, wie lautet Ihre Frage?</p> <p>2b Ich hab in meinem Vertrag viele Frei-SMS, die benutze ich aber kaum. Ich würde auf die SMS gern verzichten und meine Frei-Minuten dafür erhöhen. Darum meine Frage: Ist das möglich? Und zu welchen Konditionen? Das ist gar kein Problem. Ich werde diese Anfrage gleich weiterleiten und die Kollegen werden sich zeitnah mit einem Angebot bei Ihnen melden. In Ordnung?</p> <p>2c Ja, vielen Dank. Auf Wiederhören.</p>
Dialog 3: Mietwagen	<p>Die Firma AutoRent, mein Name ist Krüger. Wie kann ich Ihnen weiterhelfen?</p> <p>3a Guten Tag, mein Name ist Weinmann. Ich habe eine Frage zu meiner Reservierung bei Ihnen. Alles klar. Wie lautet bitte Ihre Auftragsnummer?</p> <p>3b Einen Moment. Das ist die 3423. Einen Moment bitte. So, wie lautet Ihre Frage?</p> <p>3c Ich habe mir für nächste Woche einen Umzugstransporter reserviert. Es gab jetzt Änderungen in der Planung und darum müsste mein Sohn den Transporter fahren. Er ist allerdings erst 22. Ich gehe davon aus, dass ich das zusätzlich versichern muss. Können Sie mir sagen, was das kostet? Für Fahranfänger haben wir gesonderte Tarife. Falls Ihr Sohn den Führerschein jedoch schon länger als 3 Jahre hat, fallen keine zusätzlichen Versicherungskosten an.</p> <p>3d Ach super, dann passt das wohl. Vielen Dank für die Auskunft.</p>
Dialog4: Wohnungsgesellschaft	<p>Die WoGeBa GmbH, Schiller am Apparat. Was kann ich für Sie tun?</p> <p>4a Guten Tag, mein Name ist Hoffmann. Ich habe eine Frage zu meinem Mietvertrag. Ok. Können Sie mir bitte Ihre Mietvertragsnummer sagen?</p> <p>4b Die Nummer lautet 4011. Einen Moment bitte. Alles klar, wie lautet Ihre Frage?</p> <p>4c Wie Sie wahrscheinlich sehen, hab ich das Mietverhältnis schon zum nächstmöglichen Termin gekündigt. Ich möchte aber schon früher raus und habe einen Nachmieter gefunden. Daher meine Frage: Wo müssen wir uns jetzt melden, um das unter Dach und Fach zu bringen. Ok. Am besten Sie melden sich bei Ihrer zuständigen Sachbearbeiterin. Das ist Frau Becker. Mit der ihr können Sie alles Weitere besprechen. Sie erreichen Sie unter 030-102403.</p> <p>4d Alles klar. Vielen Dank für die Auskunft.</p>

Erläuterung der Symbole:

Als Anrufer haben Sie folgende Symbole:



Neben diesem Symbol ist der Grund ihres Anrufes eingetragen (z.B.: Ich möchte ein Ticket kaufen!)



Neben diesem Symbol steht eine Bedingung, die beim Informationsaustausch berücksichtigt werden soll (z.B.: Ich möchte ein Ticket kaufen ↗ ABER möglichst günstig!)



Neben diesem Symbol sollen Sie alle Informationen eintragen, die Sie von Ihrem Gesprächspartner benötigen.



Neben diesem Symbol stehen alle Informationen, die ihr Gesprächspartner benötigt und die Sie weitergeben sollen.



Neben diesem Symbol steht eine Frage, zu der weder Sie noch Ihr Gesprächspartner eine Information vorliegen hat. Sie sollen diese Frage kurz diskutieren und zu einer einvernehmlichen Lösung kommen.

Autovermietung

Ihr Name: Pfeifer



1 Kleintransporter für einen Umzug



für ein Wochenende (Samstag und Sonntag),
mindestens 1.80 m Ladehöhe,
möglichst günstig



Modell : _____

Ladehöhe : _____

Preis : _____ €/Tag



Reservierung : Eurocard

Kreditkartennummer: 8823 7453 6592 0023

gültig bis 12/19

Führerschein Klasse 3



Ab wieviel Uhr kann ich den Wagen morgens abholen?



Szenario 7: Autovermietung

Ihr Name: *Allround Autovermietung*

Belegungsplan Kleintransporter



Tag	Ford Transit	Mercedes Sprinter	VW Transporter
Montag	<i>belegt</i>	<i>belegt</i>	
Dienstag	<i>belegt</i>	<i>belegt</i>	
Mittwoch		<i>belegt</i>	<i>belegt</i>
Donnerstag	<i>belegt</i>		<i>belegt</i>
Freitag		<i>belegt</i>	
Samstag			
Sonntag	<i>belegt</i>		
Führerschein	Klasse 3	Klasse 3	Klasse 3
Freie Kilometer	100 km	100 km	100 km
Preis pro Tag	129,- €	169,- €	109,- €
Ladehöhe	2 m	2.30 m	1.90 m



Reservierung: NAME : _____
KREDITKARTE : _____
KREDITKARTE-NR. : _____
GÜLTIG BIS : _____
FÜHRERSCHEIN : Klasse 1 Klasse 2 Klasse 3



Pizzaservice

Ihr Name: *Clemens*



1 große Pizza

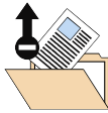


für 2 Personen
vegetarische Pizza bevorzugt



Belag : _____

Preis : _____ €



Lieferung an : Glücksburger Str. 41,
Bochum
Tel:
0334175011



Wie lange dauert es bis die Pizza geliefert wird?

Bibliothek

Ihr Name: *Hüllstrung*



Buch ausleihen



Ein Krimi von Agatha Christie
falls das Buch ausgeliehen ist, bitte vormerken



Buchtitel : _____
vorgemerkt für : _____
den : _____Tage

Ausleihdauer



Bestellung : Tel: 0175 2084552
Kundennummer: H 1092



Wieviel kostet das Ausleihen?

Szenario 13: Bibliothek

Ihr Name: *Stadtbücherei Bochum*



Lieferzeiten für bestellte Bücher

Auswahl von Autoren	Auswahl von Titeln	ausgeliehen bis nächsten	Ausleihdauer
Agatha Christie	Mord im Orientexpress	Freitag	10 Tage
	Tod am Nil	Montag	10 Tage
	Die Mausefalle	Dienstag	10 Tage
Astrid Lindgren	Pippi Langstrumpf	<i>erhältlich</i>	8 Tage
	Die Kinder von Bullerbü	Mittwoch	8 Tage
	Karlsson vom Dach	<i>erhältlich</i>	8 Tage
Michael Ende	Jim Knopf und Lukas der Lokomotivführer	Donnerstag	12 Tage
	Momo	<i>erhältlich</i>	12 Tage
	Die unendliche Geschichte	Freitag	12 Tage



Bestellung : NAME : _____

KUNDENNUMMER : _____

TELEFON : _____





Arzttermin

Ihr Name: *Rheinhardt*



Termin beim Augenarzt



möglichst vormittags vor 9.00 Uhr

Mittwoch oder Donnerstag



Termin : Tag : _____

Datum : _____

Uhrzeit : _____



Angaben zur Person : Krankenkasse UNIKA,

regelmäßiger Patient, letzter Besuch vor 6 Monaten,

Kontrolltermin



Wie lange muß ich bei Ihnen warten?



Szenario 17: Arzttermin

Ihr Name: *Augenarztpraxis Dr. Otto*



Terminplan

1. Juniwoche

	Montag, 02.06.97	Dienstag 03.06.97	Mittwoch 04.06.97	Donnerstag 05.06.97	Freitag 06.06.97
8.30		Klos			Sting
8.45	Kohl			Kree	Gabriel
9.00	Blüm	Sammer	Chapuisat	Feiersinger	Jackson
9.15	Waigel	Kohler	Riedle		
9.30	Seehofer	Cesar		Tanko	Ami Grant
9.45	Rau	Sousa	Ricken	Tretschok	Bon Jovi
10.00	Lafontaine	Zorc	Herrlich	Bout	Clapton
10.15	Schroeder	Möller	Heinrich		Grönemeyer
10.30	Scharping	Lambert	Freund		
10.45		Reuter		De Beer	Domingo



Tag, Datum & Uhrzeit : _____

Welche Krankenkasse? : _____

Waren Sie schon einmal bei uns? : JA NEIN

Wann waren Sie das letzte Mal da? : _____

Worum geht es? : Brille Kontrolle Beschwerden



Einwilligungserklärung



QU Usability Engineering Exercise 2017 - Group 3

Ich (Name des Teilnehmers /der Teilnehmerin in Blockschrift)

bin schriftlich über die Studie und den Versuchsablauf aufgeklärt worden. Ich willige ein, an der Befragung im Rahmen der Studie teilzunehmen. Sofern ich Fragen zu dieser vorgesehenen Studie hatte, wurden sie von Herrn/Frau _____ vollständig und zu meiner Zufriedenheit beantwortet.

Ich bin darüber informiert, dass mein Name, meine Anschrift und meine Telefonnummer nur auf dieser Einwilligungserklärung stehen.

Ich hatte genügend Zeit für eine Entscheidung und bin bereit, an der o.g. Studie teilzunehmen. Ich weiß, dass die Teilnahme an der Studie freiwillig ist und ich die Teilnahme jederzeit ohne Angaben von Gründen beenden kann.

Freiwilligkeit und Anonymität

Die Teilnahme an der Studie ist freiwillig. Sie können jederzeit und ohne Angabe von Gründen die Teilnahme an dieser Studie beenden, ohne dass Ihnen daraus Nachteile entstehen. Die im Rahmen dieser Studie erhobenen Daten werden vertraulich behandelt. So unterliegen diejenigen Projektmitarbeiter, die durch direkten Kontakt mit Ihnen über personenbezogene Daten verfügen, der Schweigepflicht. Des Weiteren wird die Veröffentlichung der Ergebnisse der Studie in anonymisierter Form erfolgen, d. h. ohne dass Ihre Daten Ihrer Person zugeordnet werden können.

Datenschutz

Die Erhebung Ihrer oben beschriebenen persönlichen Daten erfolgt vollständig anonymisiert, d. h. an keiner Stelle wird Ihr Name erfragt. Ihre Antworten und Ergebnisse werden unter einem persönlichen Codewort gespeichert. Das heißt, es ist niemandem möglich, Ihre Daten mit Ihrem Namen in Verbindung zu bringen.

Vergütung

Für die Teilnahme an der Untersuchung erhalten Sie keine Vergütung.

Ort, Datum & Unterschrift des Teilnehmers:

Ort, Datum & Unterschrift des Versuchsleiters:

Anleitungen Fragebogen Sprecher Bewertungen

Sehr geehrte Teilnehmerin, sehr geehrter Teilnehmer,

In diesem Experiment werden Sie sich Ausschnitte eines Dialogs einer Pizzabestellung anhören und die Sprecher nach verschiedenen Kriterien bewerten (per Zufall kann auch einer der Dialoge vom Thema Autovermietung handeln).

Sie hören nacheinander 23 verschiedene weibliche Sprecherinnen und im Anschluss 17 verschiedene männliche Sprecher. Um den Dialog abzuspielen, klicken Sie auf den Play-Button. Sie können die Sprachaufnahme mehrere Male anhören.

Während Sie die Aufnahme hören, müssen Sie anhand verschiedener Kriterien bewerten, welchen Eindruck der Sprecher auf Sie macht. Es geht hierbei nicht darum, wie die Stimme des Sprechers auf sie wirkt, sondern wie Sie den Sprecher selbst einschätzen würden. Bitte ignorieren Sie bei Ihrer Bewertung Hintergrundgeräusche, Störungen und/oder die Qualität der Aufnahmen. Der Wortlaut des Dialogs sollte bei der Bewertung nicht berücksichtigt werden.

Bewerten Sie den Sprecher indem Sie mit der Maus an die ihrer Meinung nach angemessene Stelle auf der Skala klicken. Es gibt hierbei keine richtigen oder falschen Antworten, es geht allein um Ihre subjektive Wahrnehmung. Die Wortpaare sind immer genau gegensätzlich (z.B. sympathisch/unsympathisch). Achten Sie auch darauf, dass eine positive Eigenschaft nicht immer nur auf der linken oder auf der rechten Seite steht, sondern wechselt.

Wenn Sie alle 34 Bewertungen für einen/eine Sprecher/in abgegeben haben, gelangen Sie mit „Next“ zum nächsten Dialog. Sollten Sie mit „Next“ nicht zu nächsten Aufgabe gelangen, haben Sie vermutlich eine Bewertung übersehen oder den Dialog nicht mindestens einmal vollständig angehört.

Bitte schalten Sie Ihr Handy auf stumm bevor es losgeht.

Sollten Sie Fragen haben, wenden Sie sich bitte jetzt an den/die Versuchsleiter/in.

Wir bedanken uns für Ihre Teilnahme!