Talking to a system and oneself: A study from a Speech-to-Speech, Machine Translation mediated Map Task

Hayakawa Akira[†], *Fasih Haider*[†], *Saturnino Luz*[‡],
Loredana Cerrato[†], *Nick Campbell*[†]

[†]ADAPT Centre, School of Computer Science and Statistics, Trinity College Dublin, Ireland
[‡]Usher Institute of Population Health Sciences & Informatics, University of Edinburgh, UK
{campbeak, haiderf, cerratol, nick}@tcd.ie[†], S.Luz@ed.ac.uk[‡]

Abstract

The results of a comparison between three different speech types — *On-Talk*, speaking to a computer, *Off-Talk Self*, speaking to oneself and *Off-Talk Other*, speaking to another person — uttered by subjects in a collaborative interlingual task mediated by an automatic speech-to-speech translation system, are reported here. The characteristics of the three speech types show significant differences in terms of speech rate ($F_{2,2719} = 101.7; p < 2e-16$), and for this reason a detection method was implemented to see if they could also be detected with good accuracy based on their acoustic and biological characteristics. Acoustic and biological measures provide good results in distinguish between *On-Talk* and *Off-Talk*, but have difficulty distinguishing the sub-criteria of *Off-Talk*: *Self* and *Other*.

Index Terms: speech recognition, human-computer interaction, computational paralinguistics

1. Introduction

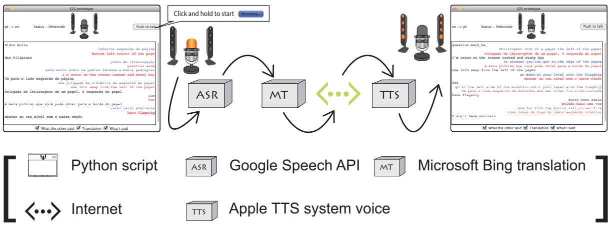
People talking to a computer can sometimes speak aside to themselves. This may be to repeat what is displayed on the computer screen, think out load, vent out frustration, or personify the computer and give it a hypothetical pat on the back. This behaviour has been observed before in speakers interacting with elaborate automatic dialogue systems [1], [2], and has been referred to as *Off-Talk* following Oppermann et al. [3], where *Off-Talk* is defined as comprising "every utterance that is not directed to the system as a question, a feedback utterance or as an instruction". Batliner et al. [2] referred to *On-Talk* as a default register for interaction with computers. How people talk to computers has been proven, in several studies, to be different from how they talk to other humans [4], [5]. In this study, we do not try to define Computer Talk, but to simply differentiate *On-Talk* and *Off-Talk* in a Speech-to-Speech (S2S) Machine Translated (MT) task oriented interaction. In a previous study [6] with the ILMT-s2s corpus, we concluded that subjects preferred a communication setup where they could not see the interlocutor. A possible side effect of this setup is the reduction of back channeling — metadata related to the understanding/completion of an instruction is not transmitted to the interlocutor, since facial cues and gestures usually carry this information. We think that a system trained to recognise these utterances could enhance its performance by either not reacting to these utterances, or process them in a special way, for instance, on a meta level, as an indication of the (mal-) functioning of the system or as an additional feedback channel to the interlocutor. Previous studies contrasting *On-Talk* and *Off-Talk*, focussing on the phonetic and prosodic delivery of utterances [2], have shown that generally Computer Talk (i.e. *On-Talk*) is similar to talking to someone who is hard of hearing: more hyperarticulated with higher energy. Branigan et al. [4] even mentions that communication to a computer is more exaggerated when compared to a fellow human. Automatic speech recognition (ASR) systems do not always work as they should and this can trigger different repair strategies from speakers. These strategies are meant to increase the understanding for the system, but actually end up being even more difficult to process for ASR systems, causing an increase of the recognition error rates. What has been less investigated is the speaker reaction in terms of production of *Off-Talk* consisting of comments about the mal-functioning of communication — due to the system or the difficulty of the task. In our study we look at *On-Talk* and 2 variants of *Off-Talk* produced by users of a computer system that mediates their interlingual S2S interactions in a collaborative task.

2. Material

The data used in this study is part of the ILMT-s2s corpus [7] and includes the speech of 30 subjects, with 15 annotated and recorded dialogues between speaker of 2 different languages (English and Portuguese) and biological signals recorded by means of biosignal tracking devices.

2.1. The ILMT-s2s System

Two subjects, seated in two different rooms, used the ILMT-s2s system (Figure 1) to communicate to each other. The ILMT-s2s system, is a system that uses off-the-shelf components to perform Speech-To-Speech Machine Translation. It is activated by a "Push-to-talk" button that the subject will click-and-hold for the duration of the utterance and release once the subject has finished. Neither subject can hear the other's voice since the output of the ASR and MT is provided by a synthetic voice.



Sinto muito
inferior esquerdo da página
bottom left corner of the page
das Filipinas
ponto de interrogação
question mark
seta estou sobre as pedras lavadas e bahia pedregoso
I'm arrow on the stones washed and stony Bay
Vá para o lado esquerdo da página
uma polegada de distância da esquerda do papel
one inch away from the left of the paper
Polegada de Christopher de um papel, à esquerda do papel
sim
Yes
é mais próximo que você pode obter para a borda do papel
tenho navio almirante
have flagship
descer ao seu nível com o carro-chefe

question mark_ba_
Christopher inch of a paper the left of the paper
Polegada de Christopher de um papel, à esquerda do papel
I'm arrow on the stones washed and stony Bay
is closest you can get to the edge of the paper
é mais próximo que você pode obter para a borda do papel
one inch away from the left of the paper
go down to your level with the flagship
descer ao seu nível com o carro-chefe
Yes
go to the left side of the mountain until your level with the flagship
Vá para o lado esquerdo da montanha até seu nível com o carro-chefe
have flagship
furry once again
peludo mais uma vez
how far from the bottom left corner fire
como longe do fogo do canto esquerdo inferior
I don't have mountain

Figure 1: *ILMT-s2s system used to collect the data*

doi: 10.21437/SpeechProsody.2016-159

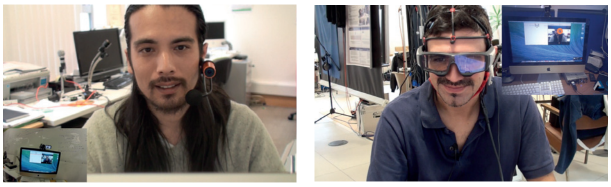
2.2. Audio, Video and Biosignal Recordings

Two audio and five video sources are included in the ILMT-s2s corpus. Of these, the audio from the two video cameras that captured the images in Figure 2 were used for this study, since they recorded the whole dialogue from start to end.

To record the biosignals, a Mind Media B.V., NeXus-4 was used to collect the Heart Rate (HR) using the Blood-Volume Pulse (BVP) readings, Skin Conductance (SC) and the brains electrical activity through Electroencephelography (EGG). The BVP sensor was placed on the index finder, with the SC sensor placed on the middle and ring finger. EEG sensors were placed in the F4, C4, P4 with a ground channel placed at A1 of the 10 – 20 location system. The sampling frequency for the SC, HR and EEG were 32 kHz, 32, kHz and 1,024 kHz respectively.

2.3. The Subjects and Recording Environment

The subjects were recruited from the Trinity College Dublin digital noticeboard or via personal connections. Fifteen recordings of fifteen native English speakers (♀5, ♂10), and fifteen native Portuguese speakers (♀11, ♂4), between the ages of 18 and 45 were collected. Each recording session was conducted in a working office and they last between 20 and 74 minutes and contains between 43 and 219 transcribed utterances. One subject during each recording session was fitted with the biosignal recording device, while the other subject was not (Figure 2).



w/o biosignal monitor

w/ biosignal monitor

Figure 2: *Subjects during recordings*

2.4. The Map Task Technique

Maps from the HCRC Map Task corpus [8] were used to elicit the task oriented conversation between the subjects. Of the sixteen HCRC Map Task maps, map 01 and map 07 were used – with a copy translated into Portuguese for the Portuguese speaker. As with the HCRC Map Task, the subjects in each recordings were given a role of either Information Giver (IG) or Information Follower (IF), where the IG has a map with a route drawn on it. The IG has to instruct the IF to draw the route on his/her unmarked copy of the map. Each map contains a number of landmarks (e.g., “white mountain”, “baboons”, “crest falls”) which may or may not be common to both maps (Figure 3). This difference between the IG’s and IF’s map, combined with the fact that neither subject can see the other’s map adds to the complexity of the task.

2.5. On-Talk, Off-Talk Labels

We used the dedicated annotation tool ELAN [9] to label the transcription with *On-Talk*, *Off-Talk* (*Self* and *Other*). As mentioned in § 1, *On-Talk* are locations within the dialogue where the subject is talking to the ILMT-s2s system to communicate to the other subject and *Off-Talk* are utterances that are not directed at the ILMT-s2s system. *Off-Talk* was further subcategorised into *Self* and *Other*. *Self* being *Off-Talk* to oneself and *Other* being *Off-Talk* to another person in the room – the other person could be the technician of the experiment who entered the room on the few occasions when the system crashed, or a university member using the office for other purposes.

HCRC Map Task map 01

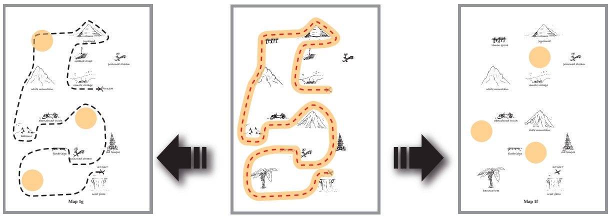


Figure 3: *Map 01, with differences highlighted*

On-Talk locations were retrieved from the ILMT-s2s system’s log and all other utterances were annotated manually for *Off-Talk Self* and *Off-Talk Other* (Table 1).

Table 1: *Total No. of utterances types in ILMT-s2s corpus.*

Utterances	w/o Biosignals	w/ Biosignals	Total
On-Talk	1,110	1,329	2,439
Off-Talk	579	610	1,189
Self	370	478	848
Other	209	132	341
Total	1,689	1,939	3,628

3. Method and Results

Based on the following speech rate comparison of the data, a significant difference between *On-Talk* and *Off-Talk* was observed from the speech rate of the subjects (§ 3.2). Since there was an overlap for all 3 talk type speech rates, we experimented with the data to see if *On-Talk* and *Off-Talk* can be automatically detected or not with other means (§ 3.3).

3.1. Method: Speech rate comparison

For this analysis, a 180 wpm TTS output of all the utterances was made using the same synthetic voice as the ILMT-s2s system, and then segmented using Praat [10] to obtain a reference utterance duration, as used in our previous study [11]. The reference utterance duration was then used to calculate a percentage difference with the original subject utterance $(1 - S/T)$, where S is the duration of the speaker’s utterance and T is the duration of the TTS output, with a positive result indicating speech faster than the ILMT-s2s system TTS output and a negative result indicating slower speech. However, due to the higher ratio of single word utterances (e.g., “umm”, “ok”, “yes”, “what?”, “ah”, etc.) in *Off-Talk Self*, single word utterances have been removed from the data to reduce the standard deviation difference that it causes (All utterances’ *sd* w/ 1 word: 74.14, w/o 1 word: 47.09). This resulted in 2,093 *On-Talk*, 629 Off-Talk (395 *Self* and 243 *Other*) utterance speech rates being used for this analysis.

Preliminary tests of the dialogue show that within the first thirty seconds of the dialogue, there is no significant difference between the speech rate difference of *On-Talk* and *Off-Talk* $(F_{1,44} = 0.031; p = 0.862)$. Even when the data is expanded to the first one hundred seconds, the significance is still small ($F_{1,121} = 4.005; p = 0.048$). This suggests that the subjects started the dialogues with similar speech rates.

Furthermore, as previously studied in [12], [11], a correlation between Word Error Rate (WER) and hyperarticulation has been identified. However, it was observed that of the fourteen subjects that start with one hundred percent accurate ASR results, the onset of hyperarticulation precedes the ASR result error. If it was a reaction of WER, then hyperarticulation should start after the first ASR error. This is an indication that communication through the ILMT-s2s system was not the only cause of hyperarticulation for the subjects.

To indicate that there is a difference between the talk types the following null hypothesis is tested on all the individual subjects, and the various categories that they can be divided into within the corpus settings.

H_0: The means of utterance speech rate differences are the same for talk types.

3.2. Result: Speech rate comparison

Of the 30 subjects, 15 subjects have less than 12.22% of *Off-Talk* utterances in their dialogues, of which 3 subjects have no *Off-Talk* utterances at all — 1047 *On-Talk*, 52 *Off-Talk* (mean % of *Off-Talk* within each dialogue: 5.03%, sd: 4.2%) (43 *Self* and 9 *Other*). The remaining 15 subjects have between 12.24% to 61.95% of *Off-Talk* utterances within the dialogues — 1046 *On-Talk*, 577 *Off-Talk* (mean % of *Off-Talk* within each dialogue: 34.83%, sd: 14.0%) (352 *Self* and 225 *Other*).

The ANOVA test results show that 15 subjects out of the 30 subjects have a significant difference between the 2 types, *On-Talk* and *Off-Talk*. Of the 15 subjects with a significant difference, 2 subjects only have 8.05% and 10.81% of *Off-Talk* within the dialogues, while the other 13 subjects have between 12.24% and 61.95% (mean 36.33%, sd: 14.2%).

When the test is run for the 3 types, *On-Talk*, *Off-Talk Self*, *Off-Talk Other*, 17 subjects out of the same 30 subjects have a significant difference in the speech duration. Compared to the 2 type comparison above, a subject with 12.22% and 32.53% of *Off-Talk* have shown significant differences.

Following the ANOVA test of individual subjects, to clarify that this difference is not merely a characteristic of a specific category within the corpus the test was performed with the subjects divided by categories. The results show a significant difference being observed in all categories (Table 2 and Figure 4). The removal of 2 word utterances revealed similar results.

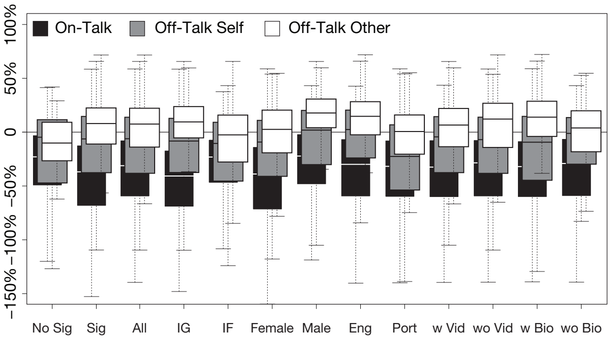


Figure 4: *The 3 talk types plotted with 0 indicating the same speech rate as the TTS reference output, positive % points as faster than the TTS output, and negative % points as slower*

Table 2: H_0 of the 3 talk types in each category

Category	ANOVA results	
H_0 – All	$F_{2,2719} = 101.7; p < 2e-16$	(∗∗∗)
H_0 – IG	$F_{2,1475} = 86.59; p < 2e-16$	(∗∗∗)
H_0 – IF	$F_{2,1241} = 21.30; p < 8.06e-10$	(∗∗∗)
H_0 – ♀	$F_{2,1465} = 84.24; p < 2e-16$	(∗∗∗)
H_0 – ♂	$F_{2,1251} = 41.17; p < 2e-16$	(∗∗∗)
H_0 – En	$F_{2,1457} = 89.27; p < 2e-16$	(∗∗∗)
H_0 – Pt	$F_{2,1259} = 29.96; p < 1.94e-13$	(∗∗∗)
H_0 – Pt-Pt	$F_{2,1131} = 8.15; p < 0.000305$	(∗∗∗)
H_0 – w/ Video	$F_{2,1574} = 79.15; p < 2e-16$	(∗∗∗)
H_0 – w/o Video	$F_{2,1142} = 23.46; p < 1.03e-10$	(∗∗∗)
H_0 – w/ Bio	$F_{2,1397} = 48.78; p < 2e-16$	(∗∗∗)
H_0 – w/o Bio	$F_{2,1127} = 54.13; p < 2e-16$	(∗∗∗)

3.3. Method: Detection of On-Talk & Off-Talk

For the following experiments, the start and end times of the *On-Talk*, *Off-Talk* label annotation were used to segment the synchronised audio and biosignal files. Two of the fifteen EEG recordings provided faulty readings and were excluded from the dataset. This resulted in 1,127 *On-Talk*, 554 *Off-Talk* (422 *Self* and 132 *Other*) utterance locations being used for this experiment.

For the detection of *On-Talk* and *Off-Talk* we extract features from audio and biosignals and explore the potential use of these features to identify *On-Talk* and *Off-Talk*.

Exp. 1: A 2-Class experiment where we only distinguish the difference between *On-Talk* and *Off-Talk*.

Exp. 2: A 3-Class experiment where we distinguish the difference between *On-Talk*, *Off-Talk Self* and *Off-Talk Other*.

3.3.1. Feature Extraction

The following features were used for the classification task.

Audio features: For the classification task we use the INTERSPEECH 2013 Computational Paralinguistics Challenge (ComParE) feature set [13]. This contains energy, spectral, cepstral (MFCC) and voicing related low-level descriptors, as well as other descriptors such as logarithmic harmonic-to-noise ratio (HNR), spectral harmonicity, and psychoacoustic spectral sharpness. To ignore the most irrelevant acoustic feature, K Means clustering algorithm is employed. This divides the feature set into 9 clusters and of these only the cluster with highest number of features is selected for classification. As a result, the total number of acoustic features reduces from 6,373 to 6,356.

Biosignal features: For the biosignals (HR, SC and EEG) we calculate Shannon entropy, mean, standard deviation, median, mode, maximum value, minimum value, maximum ratio, minimum ratio, energy and power. This feature set is calculated for each biosignal and its first and second order derivative. In total we have 33 features for each biosignal. The EEG gamma signals from sensor A and B (10 – 20 system: F4 – C4 and C4 – P4) are considered in this study due to their higher prediction power for mental tasks classification [14]. The minimum ratio of an observation is measured by counting the number of instances which have a lower value compared to their preceding and following instance and then dividing it by the total number of instances in that observation. Similarly, the maximum ratio of an observation is measured by counting the number of instances which have a higher value compared to their preceding and following instance and then dividing it by the total number of instances in that observation.

3.3.2. Classification Method

The classification method was implemented in MATLAB[1] using Statistics and Machine Learning Toolbox and employed discriminant analysis in 10-fold cross validation experiments. The classification method works by assuming that the feature sets of the classes to be discerned are drawn from different Gaussian distributions and adopting a pseudo-linear discriminant analysis (i.e. using the pseudo-inverse of the covariance matrix [15]).

3.4. Result: Detection of On-Talk & Off-Talk

The following results were obtained. See Table 3 and Figure 5 for details.

Table 3: *Discriminative Analysis Method Results – F Score (%)*

Signal	Experiment 1		Experiment 2		
Class (Talk type):	On	Off	On	Off Self	Off Other
EEG	79.91	34.47	79.89	26.55	5.13
HR	80.41	31.55	80.44	21.68	8.54
SC	81.61	48.30	81.39	36.31	5.19
HR + SC	80.85	49.40	81.64	40.12	11.56
All Bio (EEG+HR+SC)	80.32	50.31	80.88	42.60	13.33
Audio	94.14	87.55	94.00	77.60	36.64
Audio + EEG	94.31	87.94	94.13	78.50	33.60
Audio + HR + SC	94.87	88.91	94.17	77.17	34.62
Audio + All Bio	94.09	87.47	94.38	79.08	38.13

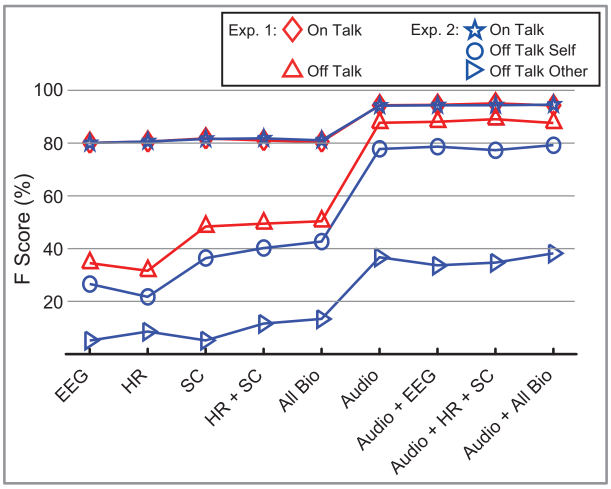


Figure 5: *Discriminative Analysis Method Results*

The results of experiment 1 show that the acoustic and biological measures significantly contribute to the prediction of *On-Talk* and *Off-Talk*. The acoustic feature set provides the optimum performance with a maximum F scores of 94.14% for *On-Talk* and 87.55% for *Off-Talk*. Also the SC feature set performs better than other biological features but a fusion of the bio feature sets cause an increase in prediction. However, a fusion of acoustic and bio features improves the performance in two cases, but has almost no effect as compared to audio feature alone when audio features are fused with all the bio features.

From the results of experiment 2 we can see that the 3-Class results for *On-Talk* are almost the same as the 2-Class *On-Talk* results. Also results for *Off-Talk Other* are poor using bio features alone (max. 13.33%) but significantly improve when combined with the acoustic feature set (38.13%) — considering that the dataset is imbalanced, with less instances for *Off-Talk Other* (7.85%) these results can be regarded as quite good. The HR is found to have more prediction power as compared to EEG and SC and the fusion of biosignals improves the prediction. A decrease in *Off-Talk Other* results is observed when audio (36.64%) feature set is combined with EEG (33.60%) and with HR and SC (34.63%) feature sets. This might be due to the lower number of bio features since when we fuse them all together (All Bio: HR, SC, and EEG) and increase the number of bio features, we get the highest F-Score (38.13%) as expected. Although the acoustic feature set performs best as compared to other signal sets, we believe there is still room for improvement from the biosignals since they currently use a limited number of features (only 33 features for each signal) and may contain some noise components (head movements of subjects etc).

4. Discussion and Conclusion

The main motivation of this study, apart from it's novelty, was to verify if there was a distinguishable difference between *On-Talk*, *Off-Talk Self* and *Off-Talk Other* for a interactive system to provide better performance and a better understanding of the interlocutor. This was achieved with the clear significant difference, moderate Cohen's d estimate and good prosodic prediction results. However the sub-finding that hyperarticulation was not initiated by the ASR WER is of interest and also the significant difference between the *On-Talk* of IG and IF ($m = -47.05, sd = 44.43$ and $m = -26.99, sd = 42.48$), and female and male ($m = -48.49, sd = 48.23$ and $m = -27.920, sd = 38.24$) in Figure 4 needs further investigation. It is easily imaginable that the perception of simplicity of the map task with the actual complexity of providing understandable instruction caused the initial hyperarticulation. Combine this with the difficulty of using the computer interaction systems may be the cause of this difference, and it will be interesting to see if the speakers of the original HCRC map task also displayed similar hyperarticulation differences.

It must also be mentioned that the method described in § 3.3, in general provides good results to predict *On-Talk* and *Off-Talk*, but results from experiment 2 leaves the need to explore other prosodic and biological discriminative feature sets (notably using the higher frequency band of the EEG signal).

5. Acknowledgements

This research is supported by Science Foundation Ireland through the CNGL Programme (Grant 12/CE/I2267) in the ADAPT Centre (www.adaptcentre.ie) at Trinity College Dublin, and FP7-METALOGUE project under Grant No. 611073.

[1] http://uk.mathworks.com/products/matlab/

6. References

[1] A. Batliner, C. Hacker, and E. Nöth, "To talk or not to talk with a computer: On-talk vs. off-talk," *How People Talk to Computers, Robots, and Other Artificial Communication Partners*, p. 79, 2006.

[2] ——, "To talk or not to talk with a computer: Taking into account the user's focus of attention," *Journal on multimodal user interfaces*, vol. 2, no. 3-4, pp. 171–186, 2009.

[3] D. Oppermann, F. Schiel, S. Steininger, and N. Beringer, "Off-talk-a problem for human-machine-interaction?" in *Proceedings of INTERSPEECH'01: the 2nd Annual Conference of the International Speech Communication Association*. Aalborg, Denmark: Citeseer, 2001, pp. 2197–2200.

[4] H. P. Branigan, M. J. Pickering, J. Pearson, and J. F. McLean, "Linguistic alignment between people and computers," *Journal of Pragmatics*, vol. 42, no. 9, pp. 2355–2368, 2010.

[5] K. Fischer, "How people talk with robots: Designing dialog to reduce user uncertainty," *AI Magazine*, vol. 32, no. 4, pp. 31–38, 2011.

[6] L. Cerrato, A. Hayakawa, N. Campbell, and S. Luz, "A speech-to-speech, machine translation mediated map task: An exploratory study," in *Proceedings of the Future and Emerging Trends in Language Technology*, Seville, Spain, 2015, in press.

[7] A. Hayakawa, S. Luz, L. Cerrato, and N. Campbell, "The ILMT-s2s Corpus — A Multimodal Interlingual Map Task Corpus," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. Portorož, Slovenia: European Language Resources Association (ELRA), 2016, in press.

[8] A. H. Anderson, M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. S. Thompson, and R. Weinert, "The HCRC Map Task Corpus," *Language and Speech*, vol. 34, no. 4, pp. 351–366, Oct. 1991. [Online]. Available: http://las.sagepub.com/content/34/4/351

[9] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes, "Elan: a professional framework for multimodality research," in *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Genoa, Italy, 2006, pp. 1556–1559.

[10] P. Boersma and V. van Heuven, "Speak and unspeak with praat," *Glot International*, vol. 5, no. 9-10, pp. 341–347, 2001.

[11] A. Hayakawa, L. Cerrato, N. Campbell, and S. Luz, "A Study of Prosodic Alignment In Interlingual Map-Task Dialogues," in *Proceedings of ICPhS XVIII (18th International Congress of Phonetic Sciences)*, The Scottish Consortium for ICPhS 2015, Ed., no. 0760. Glasgow, United Kingdom: University of Glasgow, 2015.

[12] A. J. Stent, M. K. Huffman, and S. E. Brennan, "Adapting speaking after evidence of misrecognition: Local and global hyperarticulation," *Speech Communication*, vol. 50, no. 3, pp. 163–178, 2008.

[13] B. Schuller, S. Steidl, A. Batliner, A. Vinciarelli, K. Scherer, F. Ringeval, M. Chetouani, F. Weninger, F. Eyben, E. Marchi *et al.*, "The INTERSPEECH 2013 Computational Paralinguistics Challenge: Social Signals, Conflict, Emotion, Autism," in *Proceedings of INTERSPEECH'13: the 14th Annual Conference of the International Speech Communication Association*. Lyon, France: ISCA, 2013, pp. 148–152.

[14] H. Liu, J. Wang, C. Zheng, and P. He, "Study on the effect of different frequency bands of EEG signals on mental tasks classification," in *Proceedings of the 27th Annual International Conference of Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005.* IEEE, 2006, pp. 5369–5372.

[15] S. Raudys and R. P. W. Duin, "Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix," *Pattern Recognition Letters*, vol. 19, no. 5-6, pp. 385–392, Apr. 1998.