



Interlocutor-dependent intraspeaker speech rate variability in interaction: a pilot study on four conversations in modern Hebrew

Yshai Kalmanovitch

University of Zurich

Yshai.kalmanovitch@sunrise.ch

Abstract

The effect of interacting with different interlocutors on intraspeaker speech variability is researched since the 1970s. While the traditional position of the Speech Accommodation Theory argued that speakers accommodate their speech features according to their changing social goals in interaction, more recent predictive code models such as the Interactive Alignment Model assume that speakers converge to their interlocutors over time to facilitate cognitive processing of speech in interaction. The two positions make different assumptions regarding the nature of and the motivation underlying interlocutor-dependent intraspeaker variability. Nevertheless, researchers of phonetic convergence in interaction use mostly the same designs to collect evidence for both theories.

The current paper presents a study designed specifically to test one of the hypotheses of the Interactive Alignment Model. A single female speaker of Modern Israeli Hebrew was recorded in dyadic interactions with four close acquaintances. 20 utterances were extracted for each speaker from each recording, controlling for syllable count. A positive correlation was found between the speech rate of the observed single speaker and that of her interlocutors in each individual interaction, supporting the Interactive Alignment Model. Simultaneously, the results point to possible integration of social parameters affecting speech variability on the local level of conversation.

Index Terms: Speech Accommodation Theory, Interactive Alignment Model, speech rate, convergence in interaction.

1. Introduction

The effect of perceived speech patterns of interlocutors on intraspeaker speech variability became a major interest in linguistics mainly since the 70s and the 80s of the last century, with the emergence and development of theories and models explaining mutual effects between interactants on linguistic and communicational behavior [1,2]. Probably the most prominent of those theories became the Communication Accommodation Theory [2]. The Communication Accommodation Theory (CAT) explained shifts in interlocutors' speech patterns in interaction as motivated by the changing interactional goals of interactants, expressing and/or reflecting social solidarity or distance between the interacting parties by respectively converging to or diverging from each other [2,3]. Studies on accommodation of phonetic and prosodic features under these premises often showed mixed and highly interpretative results, not allowing clear predictions [4,5,6,7,8]. Perceived similarity in speech production, although not supported by acoustic evidence, was nevertheless observed to correlate with perceived closeness between interactants [2,9,10,11]. Other studies

showed that linguistic similarity between interacting parties can assist and/or predict better cooperation between them [12,13].

A different hypothesis explaining interlocutors' effect on intraspeaker speech variability was brought forward by models based on episodic learning [14,15,16]. Such models predict that speakers will converge to their interlocutors with the aim of reducing variability in the immediate perceptual environment, consequently leading to greater predictability and to the facilitation of processing of communicational acts and gestures [15]. Behaviourally, this hypothesis was supported mostly by evidence from spontaneous phonetic imitation in close shadowing experiments, showing positive correlation between reaction time (as a measure of cognitive processing costs) and successful imitation [17,18].

Regarding the trend explaining interlocutor-dependent intraspeaker variability, the predictions of models such as Pickering and Garrod's influential Interactive Alignment Model [14] are much stronger than that of CAT, claiming that a default convergence tendency is automatic and not directly influenced by social goals of the accommodating speaker towards his/her interlocutors. Support for this argument can be drawn also from shadowing experiments that tried to implement some social parameters in the non-social context of their design [19,20]. Subjects were seemingly showing greater convergence towards speaker they found more socially attractive, but the general tendency towards convergence was still valid for all cases.

As for the actual nature of the process of converging towards other interlocutors, models based on exemplar learning such as the Interactive Alignment Model make considerably different assumptions than those of CAT [4]. To begin with, successful convergence is not intended to be based on an intentional (even if not necessarily conscious) imitation of one's interlocutors as supposed by CAT. Instead, convergence is conceived of as a spontaneous and automatic process, enabled and enhanced by cumulated interactional experience [15]. Furthermore, especially considering non-content features of speech, there is no evidence that adjustment and attenuation to a perceived stimulus is a linear process. Indeed, several studies suggest that convergence involves sensorimotor adaptation through non-conscious calibration and recalibration of the articulatory behavior based on the evaluation of the acoustic feedback of one's own speech, compared to that of other interlocutors [15,21].

This implies that successful convergence, especially of non-content features of linguistic behavior, requires a much longer interactional experience than the one allowed by many studies on speech accommodation in interaction, often observing a 20-40 minutes of interaction between speakers completely unfamiliar to each other prior to their recording. Indeed, several

studies showed the advantage of long-term speaker with long-term acquaintance in converging to each other [4,22,23].

A second assumption, on which many studies on speech accommodation are seemingly based on, is the existence of a speaker-characteristic neutral mode of speech. This is reflected in the common use of a baseline collected prior to the interaction, to which the speech of the observed speakers during and/or after the interaction is compared. Instead, within an exemplar framework such as the Interactive Alignment Model the notion of neutral speech is by itself conceptually contradictory and cannot be considered as representing any objectivity of unaccommodated speech [4].

A study that seeks to support or disprove the predictions of the Interactive Alignment Model needs to take its unique assumption in account. Furthermore, it needs to consider the concrete function the Interactive Alignment Model attributes to intraspeaker speech variability in interaction. Thus, while CAT and similar approaches assume that intraspeaker speech variability serves to reflect the speaker's changing social and communicational interest in interaction, the Interactive Alignment Model argues that the long-term and reoccurring interactional experiences of a speaker with the same interlocutor will result in the emergence of distinct interlocutor-dependent speech patterns as part of a comprehensive conceptual-behavioral priming mechanism [14]. The following sections proposes such a design used in a pilot study of speech rate convergence.

2. Method

The goal of the current study is to investigate intraspeaker speech variability with respect to speech rate. Following the assumptions and predictions of the Interactive Alignment Model [14,15], distinct speech rate patterns are expected to be found in the same speaker when interacting with different interlocutors. The different speech rate patterns of the observed speaker are expected to converge to those of his/her changing interlocutors.

2.1. Subjects

A 23 years old female native speaker of Modern Israeli Hebrew served as the main speaker in this study. The main speaker, a writer and editor in a local newspaper living in Tel Aviv, was recorded in four one-hour sessions of free dyadic conversation with four speakers well acquainted to her and with whom she interacts on a regular basis. Table 1 gives some basic details about the four secondary speakers (S1, S2, S3 and S4); all are native speakers of Modern Israeli Hebrew.

Table 1: *Details about the secondary speakers and their acquaintance with the main speaker*

Speaker	Age	Gender	Place of residency	acquaintance
S1	23	Male	Kiryat Ono	5 years
S2	28	Male	Central Tel Aviv	Ca. half a year
S3	21	Female	Central Tel Aviv	1.5 years
S4	24	Female	Central Tel Aviv	1 year

S1, S3 and S4 were defined by M as "close friends". S2 was described by her as "a friend and sometimes more than that". Some of the secondary speakers knew each other to some extent, but none of them constituted a close social group

with each other. The secondary speakers were paid 120 NIS each. The main speaker was paid 600 NIS.

2.2. Recordings

All the conversations were recorded in stereo using a ZOOM H4n handy recorder and two external omni-directional lavalier microphones (Sennheiser MKE 2 PC) attached to the participants' collar, at a sample rate of 44.1 KHz. All the recordings were performed by the main speaker who was instructed by the author about operating the recording device. The recordings were then transferred and saved as WAV-files on an external drive (Seagate Technologies).

The recordings took place between 17th and 27th of October 2013 in the speakers' natural surroundings at the home of either the main speaker or the secondary speaker participating in the respective recording session, in the later evening after the usual working time. Thus, potential environmental or fatigue effects are assumed to be similar for all the speakers in all the recordings. The speakers were unaware of the research question of the study and the experimental context did not seem to have affected the openness and the naturalness of the conversations, as became embarrassingly evident from the content of some interactions.

2.3. Extracting the utterances for analysis

Initially, all the utterances made by the speakers were extracted and annotated using TextGrids aligned with the speech signal in PRAAT [24]. For this study, only utterances containing statements were considered. Utterances containing perceivable pauses, hesitation or otherwise non-fluent speech as well as utterances overtly deviating from the speaker perceived characteristic speech patterns (including utterances containing direct reported speech, see [25,26]) were excluded from the analysis.

To simplify the statistical analysis, 20 utterances for each speaker in each session were blindly extracted to create an equal normal distribution of canonical syllable count in each utterance [27,28,29], which is considered a strong predictor of intraspeaker variability in speech rate [30,31,32]. The equal distribution of syllable count for each speaker-in-session was tested by a one-way ANOVA with 'speaker' as independent variable and 'syllable count' as dependent variable ($F(7,152)=0.142, n.s.$), as is shown in table 2.

Table 2: *mean syllable count per utterance per session for the main speaker (M) and the secondary speakers (S).*

Session	M	S
1	17.4 (± 1.96) range: 13-19	17.75 (± 2.29) range: 11-21
2	17.8 (± 2.57) range: 13-18	17.45 (± 2.61) Range: 11-24
3	17.35 (± 1.42) range: 13-21	17.75 (± 2.02) range: 13-25
4	17.8 (± 3.43) range: 14-19	17.75 (± 1.55) range: 13-18

2.4. Calculating speech rate

Mean syllable duration was calculated for each utterance by dividing the overall duration of the utterance by the total number of the canonical syllables. To reduce utterance onset and offset effects, the first syllable and the last stressed syllable of each utterance (and any syllable following it) were omitted

from the total measured duration of the utterance, as well as from the total syllable count. The mean syllable duration for each speaker in each interaction was then calculated as the mean value of the mean syllable duration for each of the utterances he/she contributed to the data. The speech rate of the main speaker was calculated for each interaction separately.

3. Results

Table 3 lists the mean syllable duration and the standard deviation for the main speaker (M) and her interlocutors (S) for each session.

Table 3: *mean syllable duration of the main speaker (M) and her interlocutors (S) in each interaction.*

Session	M	S
1	107ms (± 14) range: 81-134ms	121ms (± 16) range: 95-156ms
2	106ms (± 16) range: 73-134ms	120ms (± 14) Range: 96-148ms
3	95ms (± 23) range: 63-146ms	103ms (± 16) range: 79-141ms
4	95ms (± 13) range: 77-124ms	96ms (± 18) range: 71-150ms

A two-way ANOVA conducted on the overall data revealed a significant main effect for 'speaker' ($F(1,152)=11.39$, $p<.001$) and for 'session' ($F(1,152)=33.96$, $p<.001$), but only a marginal effect for the interaction between the two ($F(1,152)=3.43$, $p=.066$), reflecting the fact that the main speaker (M) shows a constant tendency to speak faster than her interlocutors. While the main speaker's interlocutors in each interaction are different, the main speaker herself is always the same person. However, from a descriptive point of view, it appears that the main speaker demonstrates different speech rates which in general correlate positively with the speech rate demonstrated by the individual interlocutors she interacts with.

This observation is confirmed by a complementary analysis in two separate one-way ANOVAs, which yield significant main effects for 'session' on the speech rate of the main speaker ($F(3,76)=3.1$, $p=.031$) as well as of her interlocutors ($F(3,76)=11.11$, $p<.001$). Thus, the main effect for 'session' observed in the main analysis did not result only from interpersonal difference in speech rate between the different secondary speaker but reflected also the intraspeaker variability of the main speaker itself.

Further analysis correlated the speech rate of M with the mean speech rate demonstrated by each of her interlocutors in each session to show a medium size positive correlation ($r=.321$, $p=.004$), further confirming the prediction in the current study.

3.1. The effect of syllable count and time of production

Although the distribution of syllable count was regulated in advance to eliminate its possible effect on the results of the main analysis, a linear model was fitted to each speaker in each session individually to verify that this assumption was indeed valid. The model also included the effect of the stage in the session, in which each utterance was produced. The stage-in-session parameter was included to look for possible effects of changes of the individual speakers over time that might reflect a within-session convergence accounting for the results of the

main session. None of the models was found significant, supporting the interpretation of the main analysis that the differences in speech rate observed in the main speaker when talking to different interlocutors were of a long-term nature rather than the result of within-session speech accommodation.

It should be noted that the lack of effect of syllable count is in no contradiction to past studies, showing it to be a strong predictor in speech rate variability [27,28,29,30]. It can be easily explained by the normal distribution of syllable count per utterance achieved during the process of extracting the utterances for analysis (see section 2.3) and by the fact that most of the values of this parameter were centered around the mean value in each recorded interaction (see table 2).

4. Discussion

The aim of this paper was to propose a possible design for the study of (phonetic) convergence in interaction. The study was based on the prediction of Pickering and Garrod's Interactive Alignment Model, claiming that speakers develop over time distinct speech patterns which are activated in reoccurring interactional contexts to reflect and enhance conceptual and communicational experience [14]. The interpretation of the main analysis of the pilot study reported here indeed indicates that the observed main speaker shows different and distinct patterns of speech rate, and that this intraspeaker variability is determined by convergence to her different interlocutors.

Although the design of the reported study differs from most studies on speech accommodation and imitation in interaction by using well acquainted interactants and by avoiding the use of a baseline, it is not completely novel. At least two well-known studies already used a similar design in the past to investigate questions related to CAT.

Coupland observed speech accommodation of a travel agent in conversation with eight of her clients [33]. He found that the travel agent converged towards the regional varieties of her respective interlocutors. Coupland however was careful not to conclude that his main speaker converged to the actual speech of her interlocutors as she perceived it. Instead, he argued that her convergence could have been based on her knowledge of her clients' origin and her own expectations regarding the regional varieties she attributed to them [31, p. 65]

Gregory and Webster looked at the level of convergence of the TV host Larry King to his celebrity interviewees [5]. They assumed that King would converge more towards interviewees with higher public status than towards those with a lower one. Analyzing Kings voice's long-term average spectrum, they indeed confirmed their hypothesis. Yet, a different interpretation is also possible to Gregory and Webster's results that in fact support the assumption discussed earlier, predicting that successful convergence needs more time and experience.

In particular, Gregory and Webster noted that King converged the most towards his fellow moderator Mike Wallace who was ranked only in the lower half of the celebrity scale, but with whom Larry King should have most likely been personally well acquainted as a colleague. Since King, as a matter of practice in his profession, is expected to have had better acquaintances with higher profile celebrities than with lesser ones, we may hypothesize that the different degree of King's convergence towards his guests is in fact a matter of experience rather than of social attractiveness. Therefore, the

apparent exception of King's convergence towards Wallace seems to be more coherent than the authors wanted to conclude.

4.1. How to define convergence?

It is important to note that the main interest of investigation in the study reported in the current paper is not the actual degree of similarity between the main speaker and her different interlocutors. Thus, the fact that the main speaker shows greater similarity in speech rate to S4 than to S1 in their respective interactions doesn't influence the general interpretation of the results. The main speaker is described as converging to her interlocutors because she clearly speaks slower with the slower speakers (S1 and S2) and faster with the faster speakers (S3 and S4).

This approach to convergence is significantly different from the approach taken in many studies using a baseline, which is considered as the starting point for any process of convergence. These studies seem to assume that convergence is absolute so that a shift from the baseline is expected regardless of the initial distance between the baseline and the target speech pattern [4]. The Interactive Alignment Model in contrast claims that convergence is necessary only to the point that the difference between the auditory input from other speakers and the auditory feedback from self-production does not add to processing costs [15]. Furthermore, due to natural distortion in the acoustic perception of the self-produced speech signal (e.g. due to bone conducting), it is unlikely that converging speakers can perfectly converge towards the speech signal acoustically perceived from others. Even with respect to temporal characteristics of the speech signal (such as speech rate), which unlike spectral features are not assumed to be distorted in perception, a certain tolerance band should be assumed [10,34].

Some flexibility is also essential to allow quick and effortless adjustment to changing environmental factors that might influence speech production and perception in the short or long term. For instance, Kalmanovitch [22] demonstrated how a group of speakers shifted their general characteristics regarding the voicing feature of Standard German consonants due to influence of their new social environment while maintaining a very stable overall degree of convergence between them with respect to those features over a year.

4.2. Speech accommodation – an integrative approach

It is very easy to come to the conclusion that CAT and similar approaches contradict models such as the Interactive Alignment Model. However, as was pointed out in the previous section of this discussion, the two approaches investigate different aspects of speech accommodation and intraspeaker speech variability in interaction and can therefore be integrated. As mentioned earlier, the differences in speech rate between the main speaker and her four interlocutors are rather diverse. In fact, the speech rate of the main speaker in her interaction with S2 is closer to that of S3 than to that of S2, and closer to the speech rate of S3 than the speech rate of the main speaker in her actual interaction with S3.

It is noteworthy however that in all the recorded interactions the main speaker speaks faster than her interlocutors. Although it is possible that this tendency is characteristic to the main speaker, it also echoes some findings in line with CAT, relating speech rate and perceived personality. Indeed, some studies suggest that faster speech rate is perceived as reflecting more competence [9,11]. Consequently, one might assume that by speaking faster, the

main speaker attempts to be perceived as more competent, either in the eyes of her interlocutors or in the eyes of the spatially non-present experimenter. The latter hypothesis could be supported also by the fact that the difference between the main speaker's speech rate and that of her interlocutors is greater at the earlier recordings than in the later ones, especially in the last one. This might be explained by the main speaker's adaptation to the recording conditions.

4.3. Speech accommodation and conversational intensity

Another possible account for the diversity of the level of convergence between the main speaker and her interlocutors in the different interactions which comes from within the framework of the Interactive Alignment Model itself is the effect of interactional or rather conversational intensity on the need of the interacting parties to converge to each other. The degree of conversational intensity – in terms of higher or lower frequency of turn taking, greater or lesser equality in the contributions of the different interacting parties etc. – accounts for conflicts in the immediate perceptual environment between the speech signal produced by interlocutors and the acoustic feedback from the speaker's own speech production [15,35]. The more frequent such conflicts are, the greater are the needs – and the chances – for successful convergence [36], leading to the stabilization of speech patterns and enhancement of predictions about the acoustic input from all the interacting parties [1].

Indeed, the conversational involvement of S1 and S2, both contributing only 25%-30% of the spoken material in their conversation with the main speaker, was notably lower than that of S3 and especially S4, whose contribution was practically equal to that of the main speaker in their respective interactions with her. Accordingly, the observed pattern, where the main speaker converges more with S3 and S4 than with S1 and S2, is well anticipated.

5. Conclusions

In the current paper I argue that designs used in multiple studies of speech accommodation in interaction are not suitable for the investigation of the assumptions and hypotheses of more recent models such as the Interactive Alignment Model. By contrast, I present an alternative design which reflects a different approach towards convergence in interaction based on interlocutor-dependent intraspeaker speech variability. Instead of investigating local changes in the speech patterns of interacting parties, this approach looks at the differences between speech patterns produced by the same speaker in interaction with different interlocutors. A pilot study using the proposed design indicates the convergence of a single speaker's speech rate to that of her different interlocutors. The interpretation of the results, however, allows the integration of socially motivated approaches like CAT with cognitively oriented theories such as the Interactive Alignment Model.

6. Acknowledgements

This study was made available thanks to the funding of the doctoral program in linguistics (DPL) of the University of Zurich.

7. References

- [1] J. N. Cappella, J. O. and Greene, "A discrepancy-arousal explanation of mutual influence in expressive behavior for adult

- and infant-adult interaction". *Communications Monographs*, vol. 49, no. 2, pp. 89-114, 1982.
- [2] H. Giles, N. Coupland and J. Coupland, "Accommodation theory: communication, context and consequences," in: H. Giles et al. (eds.), *Context of accommodation*. Cambridge: Cambridge University Press, 1991, pp.1-68.
- [3] H. Giles, H. and J. Gasiorek, "Parameters of non-accommodation: Refining and elaborating communication accommodation theory," in: J. P. Forgas et al. (eds), *Social cognition and communication*, (pp 155-172) London: Psychology Press, 2013, pp. 155-172.
- [4] Y. Kalmanovitch, "Is it all about that base? Alternative methods for the study of phonetic convergence in interaction," in *Proceedings of the conference on phonetics & phonology in German-speaking countries*, Berlin, Germany, 2017, pp. 93-96.
- [5] S. W. Gregory, Jr., and S. Webster, "A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions," *Journal of personality and social psychology*, vol. 70 no. 6, pp. 1231-1240, 1996.
- [6] J. S. Pardo, "On phonetic convergence during conversational interaction," *Journal of the Acoustical Society of America*, vol.119, pp. 2382-2393, 2006.
- [7] J. S. Pardo, I. C. Jay and R. M. Krauss, "Conversational role influences speech imitation". *Attention, Perception, & Psychophysics*, vol. 72, no. 8, pp. 2254-2264, 2010.
- [8] Pardo, J. S., Urmanche, A., Wilman, S. and J. Wiener, "Phonetic convergence across multiple measures and model talkers". *Attention, Perception, & Psychophysics*, vol. 79, no. 2, pp. 637-659, 2017.
- [9] W. B. Putman and R. L. Street, "The conception and perception of noncontent speech performance: Implications for speech-accommodation theory," *International Journal of the Sociology of Language*, vol. 1984 no. 46), pp. 97-114, 1984.
- [10] R. L. Street Jr and R. M. Brady, "Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context," *Communications Monographs*, vol. 49, no. 4, pp. 290-308, 1982.
- [11] R. L. Street, "Speech convergence and speech evaluation in fact-finding interviews," *Human Communication Research*, vol. 11, no. 2, pp. 139-169, 1984.
- [12] D. B. Buller, and R. K. Aune, "The effects of speech rate similarity on compliance: Application of communication accommodation theory". *Western Journal of Communication (includes Communication Reports)*, vol. 56 no. 1, pp. 37-53, 1982.
- [13] J. H. Manson, G. A. Bryant, M. M. Gervais and M. A. Kline, "Convergence of speech rate in conversation predicts cooperation," *Evolution and Human Behavior*, vol. 34, no. 6, pp. 419-426, 2013.
- [14] M. J. Pickering and S. Garrod, "Toward a mechanistic psychology of dialogue," *Behavioral and brain sciences*, vol. 27, no. 2, pp. 169-190. 2004.
- [15] M. J. Pickering, and S. Garrod, "An integrated theory of language production and comprehension," *Behavioral and Brain Sciences*, vol. 36 no. 4, pp. 329-347, 2013.
- [16] S. Grossberg, "Resonant neural dynamics of speech perception," *Journal of Phonetics*, vol. 31, no. 3, pp. 423-445, 2003.
- [17] S. D. Goldinger, "Echoes of echoes? An episodic theory of lexical access". *Psychological Review*, vol. 105, no. 2, pp. 251-279, 1998.
- [18] C. A. Fowler, J. M. Brown, L. Sabadini, and J. Wehling, "Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks". *Journal of Memory and Language*, vol. 49, no. 3, pp. 396-413, 2003.
- [19] M. Babel, *Phonetic and social selectivity in speech accommodation*, Doctoral dissertation, University of California, Berkeley. 2009.
- [20] M. Babel, "Dialect divergence and convergence in New Zealand English," *Language in Society*, vol. 39 no. 04, pp. 437-456, 2010
- [21] J. F. Houde and M. I. Jordan, "Sensorimotor Adaptation of Speech I: Compensation and Adaptation," *Journal of Speech, Language, and Hearing Research*, vol. 45, no. 2, pp. 295-310, 2002.
- [22] Y. Kalmanovitch, "The Role of Long Term acquaintances in speech accommodation," in: A. Leemann et al. (eds.), *Trends in Phonetics and Phonology. Studies from German speaking Europe*. Frankfurt am Main / Bern: Lang, pp. 93-107, 2015.
- [23] A. Lelong and G. Bailly, "Study of the phenomenon of phonetic convergence thanks to speech dominoes," in A. Esposito et al., (eds.), *Analysis of Verbal and Nonverbal Communication and Enactment. The Processing Issues*. Berlin, Heidelberg: Springer, pp. 273-286, 2011.
- [24] P. Boersma and D. Weenink, "Praat: Doing Phonetics by Computer," Version 5.4.10., www.praat.org, 2015.
- [25] E. Couper-Kuhlen, "Coherent voicing: on prosody in conversational reported speech," *Pragmatics & beyond. New series*, vol. 63, pp. 11-33, 1999.
- [26] Kalmanovitch 2015 "Jumping out of Context – Jumping out of Tone," *Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow, UK, 2015*, paper 804.
- [27] J. Korenan, "Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech," *the Journal of the Acoustical Society of America*, vol. 119 no. 1, pp. 582-596, 2006.
- [28] , H. R. Pfitzinger (1999, August). Local speech rate perception in German speech," in *Proc. of the XIVth Int. Congress of Phonetic Sciences, Leeds, UK, August 1999*, pp. 893-896.
- [29] P. Roach, "Myth 18: Some languages are spoken more quickly than others," in L. Bauer and P. Trudgill (eds.), *Language Myths. London: Penguin*, pp. 150-158, 1998.
- [30] H. Quené, "Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo," *The Journal of the Acoustical Society of America*, vol. 123, no. 2, pp. 1104-1113, 2008.
- [31] E. Jachevich, A. F. Robert and W. Lai, "Between-speaker and within-speaker variation in speech tempo of American English," *The Journal of the Acoustical Society of America*, vol. 128, no. 2, pp. 839-850, 2010.
- [32] S. Schwab and M. Avanzi, "Regional variation and articulation rate in French," *Journal of Phonetics*, vol. 48, pp. 96-105, 2015.
- [33] N. Coupland, "Accommodation at work: Some phonological data and their implications," *International Journal of the Sociology of Language*, vol. 46, 49-70, 1984.
- [34] H. Quené, "On the just noticeable difference for tempo in speech," *Journal of Phonetics*, vol. 35, no. 3, pp. 353-362, 2007.
- [35] Y. Kalmanovitch, "Speech in Interaction – The Zurich Tangram Corpus," in: Tagungsband der 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum, München, Germany, Oktober 2016, pp.79-81.