Estimating Adaptation of Dialogue Partners with Different Verbal Intelligence

K.Zablotskaya
Inst. of Communications Engineering,
University of Ulm, Germany
kseniya.zablotskaya@uni-ulm.de

F.Fernández-Martínez
E.T.S.I. de Telecomunicacíon,
Universidad Politécnica de Madrid, Spain
ffm@die.upm.es

W.Minker
Inst. of Communications Engineering,
University of Ulm, Germany
wolfgang.minker@uni-ulm.de

Abstract

This work investigates to what degree speakers with different verbal intelligence may adapt to each other. The work is based on a corpus consisting of 100 descriptions of a short film (monologues), 56 discussions about the same topic (dialogues), and verbal intelligence scores of the test participants. Adaptation between two dialogue partners was measured using cross-referencing, proportion of "I", "You" and "We" words, between-subject correlation and similarity of texts. It was shown that lower verbal intelligence speakers repeated more nouns and adjectives from the other and used the same linguistic categories more often than higher verbal intelligence speakers. In dialogues between strangers, participants with higher verbal intelligence showed a greater level of adaptation.

1 Introduction

When two speakers are talking to each other, they try to adapt to their dialogue partner and synchronize their verbal behaviours. The adaptation may occur at different levels: lexical (Garrod and Anderson, 1987; Brennan and Clark, 1996), syntactic (Reitter et al., 2006), acoustic (Ward and Litman, 2007), articulation (Bard et al., 2000), comprehension (Levelt and Kelter, 1982), etc. Moreover, synchronization of dialogue partners at one level may cause the adaptation process at any other level (Pickering and Garrod, 2004; Cleland and Pickering, 2003). In this paper we analyse to what degree dialogue partners with different verbal intelligence and levels of acquaintance may adapt to each other during a conversation.

Verbal intelligence (VI) is "the ability to analyse information and to solve problems using language-based reasoning" (Logsdon, 2012). The ability to find suitable words and expressions may be a great help in accomplishing such goals as persuasions, encouragements, explanations, influence, etc. Moreover, there exists a dependency between an individual's verbal intelligence level and his or her success in life (Buzan, 2002).

The first hypothesis we check in this paper is that *both lower and higher verbal intelligence speakers are able to adapt to their dialogue partners; however, this adaptation is reflected by different linguistic features*.

The second hypothesis we check in this work is that *when higher and lower verbal intelligence speakers are talking to a stranger, the former ones adapt better to their dialogue partner than the latter ones*.

This investigation may be helpful for improving the user-friendliness of spoken language dialogue systems. Systems which automatically adapt to users' language styles and change their dialogue strategies may help users to feel free and comfortable when interacting with them.

2 Method

2.1 Corpus Description

For the analysis, a corpus containing 100 monologues, 56 dialogues and 100 verbal intelligence

Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), pages 126–130,
Seoul, South Korea, 5-6 July 2012.

scores of the participants was used. The corpus was collected at the University of Ulm, Germany. All the participants were German native speakers of different genders, ages, educational levels and social status. For the monologue collection, the participants were shown a short film and were asked to describe it with their own words. The candidates were not asked to follow the language style of the film; they were asked to talk as naturally as possible in order to capture their every day conversation styles. Each monologue is about 3 minutes long and contains 370 words on an average. For the dialogue collection, the participants were asked to have a 10-minute conversation with another test person. The topic of the discussions was always the same: the education system in Germany. The average number of turns in the dialogues is 55. Afterwards, verbal intelligence of the candidates was measured using the Hamburg Wechsler Intelligence Test for Adults (Wechsler, 1982). Using this test, we obtained verbal intelligence scores of the test persons with a mean value of 113 and a standard deviation of 7.2. A more detailed description of the corpus can be found in (Zablotskaya et al., 2010; Zablotskaya et al., 2012).

2.2 Clustering

Using the k-means algorithm, the verbal intelligence scores of the test persons were partitioned into:

a) 2 clusters (Cluster L consisted of test persons with lower verbal intelligence, H contained candidates with higher verbal intelligence);

b) 3 clusters (L - lower verbal intelligence, M - average verbal intelligence, H - higher verbal intelligence).

Using the two clusters L and H, all the dialogues were partitioned into the following groups:

c) L-L is a group of dialogues where both partners had lower verbal intelligence scores;

d) H-H is a group of dialogues where both partners had higher verbal intelligence scores;

e) L-H is a group of all the other dialogues.

Using the information about the level of acquaintance of the dialogue partners, the following groups were created:

f) F-F is a group of dialogues with dialogue partners who were friends or relatives;

g) S-S is a group of dialogues with dialogue partners who had not met each other before the experiment (were strangers).

In the following sections the degree of adaptation will be compared between these groups.

3 Measuring Adaptation

There exist different approaches for measuring adaptation of dialogue partners. Reitter et al. (2006) used regression models to show that a speaker in human-human interactions aligns his syntactic structures with those of his dialogue partner. Ward and Litman (2007) modified the measures of convergence offered by Reitter. According to this modification, prime words of the first dialogue partner were determined. For measuring lexical convergence, the use of prime words by the second dialogue partner for each turn was calculated. In (Nenkova et al., 2008) the measurements of adaptation between dialogue partners were based on the usage of high-frequency words. Stoyanchev and Stent (2009) analysed adaptation calculating the number of reused verbs and prepositions by a speaker that occurred in his dialogue partner's turns.

In this work we measure adaptation as cross referencing, proportion of "I", "You" and "We" words, between-subject correlation and similarity between two texts. These approaches are described in the following sections.

3.1 Cross Referencing

Cross referencing is calculated as a number of repeated nouns and adjectives by a speaker P_1 from his dialogue partner P_2 divided by the total number of P_1's words (Sillars et al., 1997).

A one-way analysis of variance (ANOVA) showed significant difference between *Cross referencing* of speakers from the groups L, M and H ($AV_L = 0.08, AV_M = 0.047, AV_H = 0.042, F(2,97) = 8.43$, $p = 0.00062$). As we may see, speakers with lower verbal intelligence reused more nouns and adjectives of their dialogue partners than speakers with average and higher verbal intelligence.

3.2 "I", "You" and "We" words

The number of "I", "You" and "We" words in a discussion may reflect the degree of closeness of speakers. In Sillars et al. (1997) these measures were used for the analysis of language use in marital conversations and closeness of relationships between partners. It was found out that partners who had lived with each other for a long time and were happy together used "we" pronouns more often than separate pairs. In addition, the proportion of "I" and "You" words were higher for separates. In our investigation we also calculated the proportion of "I", "You" and "We" words for each groups and compared them using ANOVA. Interestingly, the proportion of "I"-words of friends was greater than that of strangers (averaged value of "I"-words for friends $AV_F = 0.0033$, for strangers $AV_S = 0.0017$, $F(1,109) = 5.33,\ p = 0.024$). This phenomena may be explained in the following way. Even discussing the German education, friends might talk about themselves. People who had not met each other before avoided talking too much about their own experience. On the other hand, the difference of "We"-words was not significant. This means that even friends were not able to linguistically show their closeness discussing such kind of topic.

3.3 Between-Subject Correlation

All the dialogue transcripts were compared with the LIWC dictionary for the German language (Wolf et al., 2008). The dictionary consists of different words sorted by 64 categories. The categories may be divided into the following groups:

- *Language composition*, for example number of words, number of unique words, pronouns, articles, etc.
- *Psychological processes*, for example positive and negative emotions, causal words, words expressing certainty, etc.
- *Relativity*, for example words related to space, motion and time.
- *Topic-related categories*, for example job, school, sleep, etc.

Each word from the dictionary may refer to several categories. For example, the word *traurig* (sad) refers to the categories *Affective Processes*, *Negative emotions* and *Sadness*.

For analysing the degree of adaptation of dialogue participants, Pearson's correlation coefficients between $F(A_i)$ and $F(B_i)$ for each feature F were calculated ($F(A_i)$ is the value of a feature F extracted from the utterances of the first dialogue partner A from a dialogue i, $F(B_i)$ is the value of a feature F extracted from the utterances of the second dialogue partner B from a dialogue i). For participants from the group L-L, 30% of the features showed a significant correlation, for participants from the group H-L this value was 23%, for H-H this value was 12%. Table 1 shows the percentage of features with significant correlation for each LIWC group.

LIWC group	H-H	L-L	H-L
Language composition	28%	37%	9%
Psychological precesses	10%	19%	23%
Relativity	10%	35%	30%
Topic-related categories	11%	37%	27%

Table 1: Percentage of LIWC categories with significant correlation coefficients.

As we can see from the results, for almost all LIWC groups lower verbal intelligence speakers engaged in a conversation showed a higher degree of adaptation.

3.4 Similarity between two Texts

If two dialogue partners adapt to each other during a conversation, the similarity between their utterances should be high. For measuring the similarity between two texts, we calculated the degree of alignment between frequency distributions of certain features (tokens) extracted from the dialogues. For comparing the frequency distributions, the chi-square test was chosen because it does not require the normality of distributions and is easy to implement. A detailed explanation of this method may be found in (Vogel and Lynch, 2007) and (Straker, 2012).

Let F_i and F_j be two text files containing n_i and n_j tokens correspondingly. If F_i and F_j have the same language style, we consider the texts to be taken from the same population and the distributions of tokens from the two files should not be significantly different (null hypothesis). The chi-square statistic

is calculated based on the observed and expected values of tokens in both text-files. If the chi-value χ_i^2 is less than a certain significance threshold c_i^2 (based on the degrees of freedom and significance level), the null hypothesis is accepted and the two files may be considered as having a similar language style (making an assumption that the language style is reflected by tokens of this type). For estimating the degree to which the two texts are similar, we calculate the distance between these two values:

$$Similarity_i = S_i = \chi_i^2 - c_i^2.$$

If $-c_i^2 <= S_i <= 0$, the similarity between the texts is significant. If $S_i > 0$, the null hypothesis is rejected: the analysed texts have different language styles.

In this investigation four different types of tokens were extracted: *Letter n-gram distributions*, *Word n-gram distributions*, *Lemma n-gram distributions* and *Part-of-speech n-gram distributions*.

The mean values of S_i for each group were compared to each other using ANOVA. Features with significant ANOVA results for the groups F-F and S-S are shown in Table 2:

Feature	S_i for F-F	S_i for S-S	F(1,54)
Word 3-g	-48.7	-29.8	10.6
Lemma 3-g	-41.8	-23.5	10.1
P.-of-speech 4-g	-38.9	10.0	8.1
P.-of-speech 5-g	-59.4	-34.3	8.6

Table 2: Significant features for F-F and S-S ($p < 0.05$).

The results show that the similarities of language between friends or relatives were greater than between participants who had not met each other before.

Our next purpose was to check whether verbal intelligence plays a certain role if we analyse dialogues between friends and strangers separately. ANOVA was applied to the mean values of the similarity measure S_i calculated for the following groups:

a) L-L, H-H and L-H only for dialogues between friends;

b) L-L, H-H and L-H only for dialogues between strangers.

ANOVA significant feature are shown in Tables 3 and 4.

Feature	S_i(L-L)	S_i(H-H)	S_i(L-H)	F
Word 4-g	-77.8	-62.4	-53.81	3.9
P.-of-sp. 6-g	-83.5	-63.7	-53.9	4.7

Table 3: Significant features for L-L, H-H and L-H only for dialogues between friends ($p < 0.05$, $F(1,53)$).

Feature	S_i(L-L)	S_i(H-H)	S_i(L-H)	F
Word 4-g.	-59.9	-90.1	-45.2	2.2

Table 4: Significant features for L-L, H-H and L-H only for dialogues between strangers ($p < 0.05$, $F(1,53)$).

As we may see from the results, a lower verbal intelligence speaker may adapt to his dialogue partner if they both are relatives or friends. On the other hand, if dialogue partners have not met each other before, higher verbal intelligence speakers are better able to adapt to their dialogue partner than lower verbal intelligence speakers.

4 Discussions

As we may see from the results, it was difficult for the candidates to linguistically show their closeness discussing the education system in Germany. However, similarity of utterances in dialogues between friends was greater than similarity in dialogues between strangers. Lower verbal intelligence speakers repeated nouns and adjectives from their dialogue partners and used words from the same linguistic dimensions more often than higher verbal intelligence speakers. The first hypothesis is just partly proven because we did not find features that reflect adaptation of higher verbal intelligence speakers. In our future work we are going to further investigate how higher verbal intelligence speakers linguistically show their closeness to the other. The results also showed that speakers with lower verbal intelligence are better able to adapt to the other if they both are relatives or friends. As we suggested in our second hypothesis, if dialogue partners are strangers, higher verbal intelligence speakers show a higher degree of adaptation. In our future work we are going to use this information for improving the classification of speakers into two and three groups according to their verbal intelligence coefficients.

Acknowledgments

This work is partly supported by the DAAD (German Academic Exchange Service).

Parts of the research described in this article are supported by the Transregional Collaborative Research Centre SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG).

References

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G. and Newlands, A. 2000. Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42, p.1-22.

Brennan, S. E. and Clark, H. H. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, p. 1482-1493.

Buzan, T. 2002. The power of verbal intelligence. HarperCollins Publishers, Inc.

Cleland, A. A. and Pickering, M. J. 2003. The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, 49, p.214-230.

Garrod, S. and Anderson, A. 1987. Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27, p. 181-218.

Levelt, W. J. M. and Kelter, S. 1982. Surface form and memory in question answering. *Cognitive Psychology*, 14, p.78-106.

Logsdon, A. 2012. Learning Disabilities. http://www.learningdisabilities.about.com/

Nenkova, A., Gravano, A. and Hirschberg, J. 2008. High frequency word entrainment in spoken dialogue. *Proceedings of ACL/HLT 2008*, p.169-172.

Pickering, M. and Garrod, S. 2004. Toward a mechanistic psychology of dialog. *Behavioral and Brain Sciences*, 27, p.169-190.

Reitter, D., Keller, F. and Moore, J. 2006. Computational modelling of structural priming in dialogue. In Proceedings of the Human Language Technology Conference of the NAACL, companion volume, 2006, p. 121-124.

Sillars A.L., Shellen W., McIntosh A. and Pomegranate M.A. Relational characteristics of language: Elaboration and differentiation in marital conversations. *Western Journal of Communication*,61, p.403-422.

Stoyanchev, S. and Stent, A. 2009. Lexical and Syntactic Priming and Their Impact in Deployed Spoken Dialog Systems. *Proceedings of NAACL HLT 2009*, Boulder, Colorado, p.189-192.

Straker D. Changing Minds. http://changingminds.org/explanations/research/analysis/chi-square.htm.

Vogel C. and Lynch G. Computational Stylometry: Who's in a Play? In *Proceedings of COST 2102 Workshop (Patras)'2007*, p.169-186.

Ward, A. and Litman, D. 2007. Automatically measuring lexical and acoustic/prosodic convergence in tutorial dialog corpora. In Proceedings of the SLaTE Workshop on Speech and Language Technology in Education.

Wechsler D. 1982. *Handanweisung zum Hamburg-Wechsler-Intelligenztest für Erwachsene (HAWIE)*. Separatdr., Bern; Stuttgart; Wien; Huber.

Wolf M., Horn A.B., Mehl M.R. and Haug S. Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count *Diagnostica*, 54, Heft 2, p.85-98.

Zablotskaya K., Fernández-Martínez F. and Minker W. 2012. Investigating Verbal Intelligence using the TF-IDF Approach. *Proceedings of International Conference on Language Resources and Evaluation (LREC) 2012*, European Language Resources Association (ELRA).

Zablotskaya K., Walter S. and Minker W. 2010. Speech data corpus for verbal intelligence estimation. In *Proceedings of International Conference on Language Resources and Evaluation (LREC) 2010*, European Language Resources Association (ELRA), Valetta, Malta.