Linguistic variability in unimodal and multimodal dialogues

6 G － 7　　　　　K.H. Loken-Kim, F. Yato, T. Morimoto

ATR Interpreting Telecommunications Research Labs.

1. INTRODUCTION

The research team at ATR Interpreting Telecommunications Research Laboratories has been exploring the possibility of introducing multimodality to our on-going interpreting telecommunications research. Multimodal (speaking, marking, writing, typing ..etc.) human-machine-human communication has a great potential to enhance communicative quality. However, the very nature of the multimodality induces interlocutors to use multiple modalities concurrently or to shift from one modality to another. Such behavior often result in unfinished and fragmented utterances. Therefore, this study set out to investigate linguistic variability in multimodal (MM) dialogues in comparison with linguistic variability in unimodal (Tel) (telephone-only) dialogues. For this study, linguistic variability was measured as 1) Interjection rates, usage of interjections, and interjection environments, 2) False starts, repairs and their patterns, 3) Number of words and turns, 4) Unique lexical items. A detailed description of the experimental procedure appears in [1, 2].

2. RESULTS

2.1) Interjection rate, Usage of interjections, Interjection patterns

There were a total of 823 interjections in MM and 617 in Tel respectively in the 16 dialogues. The interjection rates (interjections/total communication time in seconds, MM: 3202 seconds, Tel: 2118 seconds) for MM and Tel were about the same: 0.26 interjections/second in MM vs. 0.27 interjections/second. The most frequently used interjections in both MM and Tel were ano (あの) and ano- (あの－)and the overall interjection distributions were very similar (Figure 1).

Interjection environments were observed by counting the number of interjection patterns. When the interjection environments were analyzed, 3 main patterns were found (Table 1).
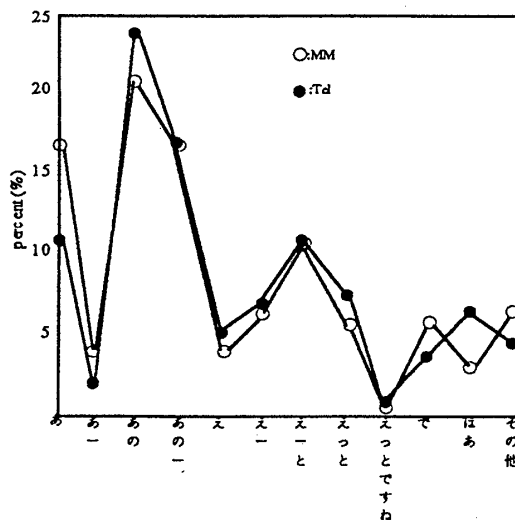

Figure 1. Usage of Interjections

Table 1. Interjection Environments (FSR)

Pattern types	MM			Tel		
Speaker	A&C	A	C	A&C	A	C
IE1	296	98	713	206	83	123
IE2	241	196	48	188	140	46
IE3	21	11	9	15	13	4
IE1+IE2+IE3+other	689	421	268	538	318	220

IE1: interjection at sentence onset, IE2: interjection follow a particle
IE3: interjection follow a conjunction
A: agent, C: client

ユニモーダルとマルチモーダル対話の特徴
ローケン・キム　キュンホ、谷戸　文廣、森元　逞
ATR音声翻訳通信研究所

2.2) False starts and repairs (FSR) and their patterns

There were a total of 142 FSRs (Table 2). Since the MM dialogues were generally longer than the Tel dialogues, there were slightly higher number of FSRs in MM than in Tel. FSR rate (repairs/second) in MM, however, was 30% lower than in Tel (MM: 0.0234 repairs per second = 75 repairs /3202 seconds, Tel: 0.0316 repairs per second = 67 repairs/2118 seconds).

Two FSR patterns [3]; <RP><R> and <RP><ET><R>, were defined and their occurrences were counted (Table 3). Out of the 142 FSRs, 105 cases were in the <RP><R>pattern, and 26 cases were in the <RP><ET><R> pattern, and both clients and agent retraced slightly more syllables in <RP><R> pattern in MM than in Tel. However, when using the <RP><ET><R> pattern, they retraced more syllables in Tel than in MM.

2.3) Total number of words and turns

The total number of words uttered by the agent in MM and Tel were 6854 words and 4700 words respectively. The speaking rates for the agents were slightly slower (2.14 words per second = 6854 words/3202 seconds) in MM than in Tel (2.21 words per second = 4700 words/2118 seconds).

The total number of turns in MM and Tel were 979 and 764 respectively. The presence of the graphic information (map, hotel reservation form, etc.) prompted more turn-taking in MM.

2.4 Unique lexical items

Figure 2 illustrates the lexical items used by the agent in MM and Tel (Total: 791 lexical items). The agent clearly used more lexical items in MM than in Tel, while 44% of the total lexical items was shared.

3. Conclusion

The difference in interjection rate between MM and Tel was negligible. The rate of FSR was less in MM, but slightly more retracing of syllables. MM users took more turns, spoke more, and spoke more slowly than Tel users. They also had a longer lexical range in MM. In the future, a study of the syntactic ambiguity of the sentences used in MM and Tel will be measured by computing the average number of parse trees per sentence.

Table 2. False starts and Repairs

Environment	MM		Tel	
Role	Agent	Client	Agent	Client
Total	34	41	27	40
ave./dialogue	4.25	5.13	3.38	5
s.d.	3.60	4.83	3.16	8.0

Table 3. Repair Patterns and Retracing Span in Number of Syllables

Pattern		MM		Tel	
		Agent	Client	Agent	Client
<RP><R>	average	2.2	2.6	1.7	2.5
	s.d.	3.16	1.89	0.82	2.26
<RP><ET><R>	average	1.3	3.1	2	3.2
	s.d.	0.43	1.83	0	1.81

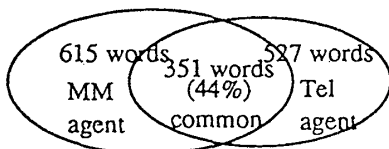<RP>: Reparandum5 <R>: Repair <ET>: Editing Term

Figure 2. Percentage of Common Words in MM and Tel

REFERENCES

[1] Loken-Kim, Yato, Morimoto, A Simulation Environment for Multi-Modal Interpreting Telecommunications, 情報処理学会AV研究会, March, 1994

[2] Loken-Kim, Furukawa, Yato, Morimoto, 情 報 処 理 学 会 第 4 8 全 国 大 会 、 March, 1994.

[3] Levelt, Monitoring and self-repair in speech, Cognition, 14 (1983) 41-104