Toward a Corpus of Cantonese Verbal Comments and their Classification by Multi-dimensional Analysis

Oi Yee Kwong
Department of Translation
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
oykwong@arts.cuhk.edu.hk

Abstract

The information explosion in modern days across various media calls for effective opinion mining for timely digestion of public views and appropriate follow-up actions. Current studies on sentiment analysis have primarily focused on uncovering aspects like subjectivity, sentiment and credibility from written data, while spoken data are less addressed. This paper reports on our pilot work on constructing a corpus of Cantonese verbal comments and making use of multi-dimensional analysis to characterise different opinion types therein. Preliminary findings on the dimensions identified and their association with various communicative functions are presented, with an outlook on their potential application in subjectivity analysis and opinion classification.

1 Introduction

Nowadays there are numerous channels for expressing personal opinions. Views expressed in written forms are no longer confined to newspapers and magazines but are found everywhere in social media on the almost boundless internet. Meanwhile, the boom in all kinds of talk shows and phone-in programmes on radio and television have allowed both experts and non-experts to voice their views on many different subjects such as politics, finance, entertainment and leisure, just to name a few examples.

The challenge in such information explosion has to be met by effective opinion mining, which has so far focused on the subjectivity, sentiment, credibility, etc. from written data. Spoken data and the sub-types of opinions are relatively less addressed, which motivated our current work.

Opinionated utterances, or verbal comments, are likely to form a specific informal spoken genre as social media text has made a specific type of written language. They are distinct for utterance lengths, incompleteness, presence of speech errors, self-repairs, and speech planning evidence, amongst others. The comments may also be further categorised according to their communicative functions, such as presenting the speaker's stance, giving advice to someone, providing information, making prediction, and evaluating or judging something. The effective classification of these different functions will be essential. This paper thus reports on a pilot study on the construction of a corpus of Cantonese verbal comments and the use of multi-dimensional analysis for characterising different types of opinions expressed therein, and discusses the potential application of the results in subsequent opinion mining work.

Section 2 reviews related work. Section 3 introduces our corpus of Cantonese verbal comments. Section 4 discusses the linguistic features used in the preliminary multi-dimensional analysis done in the current study and the initial results, while Section 5 concludes with future directions.

29th Pacific Asia Conference on Language, Information and Computation: Posters, pages 10 - 18
Shanghai, China, October 30 - November 1, 2015

2 Related Work

Opinion mining often involves subjectivity and sentiment analysis. Subjectivity analysis aims at distinguishing opinionated sentences from factual statements, where the former is also known as private states, referring to one's mental and emotional states which may express one's attitude, feeling, beliefs, evaluation, speculation, etc. Sentiment analysis attempts to classify the polarity of subjective views as positive, neutral, or negative. A comprehensive survey can be found in Pang and Lee (2008), and Liu (2010).

Past studies have mostly been concerned with written data, typically first-hand opinions like movie reviews (e.g. Pang et al., 2002), product reviews (e.g. Hu and Liu, 2004), or debates on web forums (e.g. Somasundaran and Wiebe, 2009), and second-hand opinions reported or quoted in news articles (e.g. Wiebe and Wilson, 2002; Tsou et al., 2005; Ku et al., 2006).

Systems often leverage some sentiment lexicons (e.g. Wilson et al., 2005; Esuli and Sebastiani, 2006) and are thus primarily lexically based (e.g. Pang et al., 2002; Turney, 2002; Polanyi and Zaenen, 2006; Li et al., 2012), although tasks requiring more fine-grained information like opinion holders and targets would require more than simple lexical clues (e.g. Kim and Hovy, 2006; Lu et al., 2010; Zirn et al., 2011). Approaches using multi-lingual data are also gaining attention (e.g. Banea et al., 2010).

Subjectivity may be associated with various communicative functions, such as presenting one's stance, giving advice, making prediction, evaluating and commenting, etc. Such functions are achieved with a combination of rhetorical devices including but not limited to lexical choices. Corpus-based discourse analysis has thus often relied on multiple linguistic patterns to characterise register variations (Biber, 1988; Kaufer and Ishizaki, 2006).

Multi-dimensional analysis, as explained and applied in Biber (1988) as well as Conrad and Biber (2001), makes use of multivariate statistical techniques like factor analysis to identify salient linguistic co-occurrence patterns (called "dimensions") from a wide range of linguistic features. The dimensions are functionally interpreted and then used to characterise various spoken and written registers. Biber (1993), for instance, identified five dimensions for the texts in the LOB corpus and London-Lund corpus based on 67 linguistic features. The first dimension has been labelled as "informational vs involved production", where the former is marked by features like word length, nominalizations, prepositions, etc. and the latter by present tense verbs, contractions, first and second person pronouns, etc.

The current work forms part of our project in which we investigate Cantonese verbal comments made in various domains and intend to use multi-dimensional analysis to characterise the comments and their respective communicative functions. Some preliminary results on corpus construction and the pilot study involving multi-dimensional analysis are reported and discussed in this paper. Our plan is to further employ the identified co-occurrence patterns of linguistic features for opinion mining in the future.

3 Corpus of Verbal Comments

3.1 Data Collection

The corpus compiled contains transcribed spoken Cantonese data from television and radio programmes broadcasted in Hong Kong during late 2013 to early 2014. They cover various domains (politics / current affairs, economics / finance, and food / entertainment) presented in different styles (such as interviews, phone-in programmes, singing contests, and food/film critics). Table 1 shows the data sources.

3.2 Pre-processing and Annotation

The transcription was done in verbatim with respect to individual speaker turns. The start time and end time for each turn were recorded. The role of a speaker within the programme (such as host, guest, reporter, and caller) was also noted. Self-repairs, hesitations, and pauses in the speech were indicated in the transcription accordingly. Table 2 shows an example, where the symbols //, ^^ and -- indicate intonational pause, self-repair and lengthening (by second) respectively. Transcription in Jyutping (for Cantonese) and an English translation for the content is given for reference. The talking speed for a given speech sample was calculated by the average number of syllables per minute.

Domain	Content and Programmes
Politics / Current Affairs	Interview programmes on TV/radio by host(s) with a guest, sometimes with phone-in sessions • 星期六主場 (Face to Face): *A one-to-one interview programme on TV, produced by RTHK* • 星期六問責 (Accountability): *An interview programme (with two hosts and one guest) broadcasted on radio, containing phone-in sessions, produced by RTHK* • 講清講楚 (On the Record): *A one-to-one interview programme on TV, produced by TVB*
Economics / Finance	TV programmes with discussions between host and financial analysts, sometimes with phone-in sessions • 理財博客 (Finance Blog): *A financial analysis programme on TV, usually with one host and one guest analyst, plus phone-in sessions, produced by ATV* • 華爾街速遞 (Wall Street Express): *A financial analysis programme on TV, with one host and one guest analyst, containing phone-in sessions, produced by Cable TV* • 樓盤傳真 (Property): *A real estate commentary programme on TV, with two or more hosts, reporters and interviewees, produced by Cable TV*
Food / Entertainment	TV/radio programmes with critics on food/film, and singing contests on TV with judge comments • 一粒鐘真人蘇 (One Hour So): *An entertainment programme with a main host introducing food and restaurants with critics, sometimes with cooking demonstration, may have co-host in some episodes* • 超級巨聲/星夢傳奇 (The Voice): *A series of singing contests on TV with instant comments from adjudicators, produced by TVB* • 亞洲星光大道 (Asian Million Star): *A series of singing contests on TV with instant comments from adjudicators, produced by ATV* • 電影兩面睇 (Movie World): *A film critics programme broadcasted on radio, usually with three hosts, produced by RTHK*

Table 1: Data Sources

Programme	星期六主場 (Face to Face)
Date	2013-10-12
Start time	00:02:09
End time	00:02:16
Role	Guest
Content	呀 唔係 // 佢哋 倡議 嘅 嘢 冇 問題 // 但係 aa3 m4hai6 keoi5dei6 coeng3ji5 ge3 je5 mo5 man6tai4 daan6hai6 ah no they propose 's thing have-not problem but Well, no, there is no problem with what they proposed, but 佢哋 嘅 做法 呢 就 有 問題 // 咁 所以 呢 keoi5dei6 ge3 zou6faat3 ne1 zau6 jau5 man6tai4 gam2 so2ji5 ne1 they 's method PAR ADV have problem so therefore PAR the way they did it was problematic, and so 即係 我 就 覺得 -- 即係 要 ^^ 要 出嚟 講 zik1hai6 ngo5 zau6 gok3dak1 zik1hai6 jiu3 jiu3 ceot1lai4 gong2 that is I ADV feel that is need to need to come out speak I mean, I feel that … well … I have to … have to speak out.

Table 2: Example of a Speaker Turn

Sentiment annotation mainly follows the way explained in Wiebe et al. (2005). Annotators were asked to identify any opinions and other private states (e.g. beliefs, sentiment, speculation, etc.) expressed in the transcribed speech. First, the speech content may contain objective factual information as well as subjective opinions, as in the following examples respectively:

今集星期六主場嘅嘉賓呢//亦有參與其中
(Our guest today on Face to Face also participated [in the event].)

即係你冇理由呢 eh 我叫所謂越級偷步嘅
(That means you have no reason to skip the steps and jump the gun.)

Second, opinions and other private states may be expressed explicitly with private state verbs or specific polar elements or implicitly with different styles of language, as in the following examples respectively:

呢個情況令人擔憂
(This situation is worrying.)

咁終於都出咗呢個諮詢文件喇喺啱啱呢個禮拜啦
(The consultation documents eventually come out this week. [implies the documents come out late])

Third, opinions may be from sources other than the speaker, especially when the speaker quotes someone else who is the source of the opinion. Fourth, opinions may be expressed with different strengths, which may have to be judged in context. Fifth, different attitudes may be conveyed in opinionated speech, which may typically be neutral, positive or negative. Sixth, the opinions or attitudes may be expressed with respect to certain things, people or events, which we call the target. Finally, opinionated speech may serve various communicative purposes, as demonstrated by the examples in Table 3.

For the annotation, the transcribed speech was first split into speech segments. In general the intonational pauses (marked with //) were taken as segment boundaries. Hence each speaker turn may contain one or more speech segments, and these segments may make up one or more speech events. A speech event is considered to correspond roughly to a full sentence in written text.

For each speech event, as well as any private state expressed in a speech event, the following fields are to be filled: From (the starting segment), To (the ending segment), Word span (for events with explicit speech verbs), Subjective (whether it is an opinionated segment), Source (speaker by default or otherwise indicated), Target (the object of the opinion), Strength (how strong the opinion is: low, medium, high), Polarity (positive, neutral, negative), and Function (purpose of the speech event).

Polar elements, or expressions in the speech conveying positive or negative sentiments, are also identified. For each polar element, the following information is to be provided: From (the starting segment), To (the ending segment, usually the same as From), Word span (the expression conveying polarity), Source (speaker by default, otherwise indicate nested sources), Strength (how strong the expression is: low, medium, high), and Polarity (positive, neutral, negative).

Function	Example
Stance	業主唔應該再加租囉 (The owner should not further raise the rent.)
Evaluation/ Comment	唱就唱得唔錯//但台風差啲 (The singing is not bad, but the poise is not good enough.)
Speculation	我好懷疑呢單嘢係咪真 (I really doubt the truth of this case.)
Prediction	樓價應該會跌番啲 (Property prices will probably fall a little.)
Elaboration/ Justification	因為可能就嚟加息所以… (The interest rate will probably go up, therefore …)

Table 3: Communicative Functions of Opinions

3.3 Materials for Pilot Study

The current pilot study made use of a subset of the corpus from two domains: current affairs and finance. Speaker turns by host or guest were selected. Only those turns which last at least 10

seconds were included. Each speaker turn was considered one speech sample. Drawn from about 150 minutes of transcribed speech from current affairs programmes and about 230 minutes of transcribed speech from finance programmes, a total of about 340 minutes of speech containing 495 samples with over 117K syllables were used in the analysis. The breakdown for individual domains and roles is shown in Table 4.

Domain	**Time (mins)**		**Syllables**		**Samples**	
	Host	Guest	Host	Guest	Host	Guest
Current Affairs	31.01	101.64	10,434	32,593	86	154
Finance	72.80	134.10	25,367	49,164	129	126

Table 4: Data Size for Current Study

4 Multi-dimensional Analysis

4.1 Linguistic Features Used

The following linguistic features were extracted and counted from the annotated materials described above. The quantitative data were then used in the current pilot analysis:

- Pronouns, including first person pronouns (我 *I*, 我哋 *we*), second person pronouns (你 *you*, 你哋 *you*), and third person pronouns (佢 *he/she*, 佢哋 *they*).
- Modals, including modal verbs likes 可能 *may*, 應該 *should*, 可以 *can*, 會 *will*, etc.
- Private verbs, including verbs indicating private states such as 諗 *think*, 覺得 *feel*, 認為 *think*, 相信 *believe*, etc.
- Yes/No question words, including 有冇 have or have not, 係咪 did or did not, etc. and other A-not-A patterns.
- Wh-question words, including *what* (乜, 咩…), *why* (點解, 為乜…), *who* (邊個…), *when* (幾時, 如何…), and *how* (點樣…).
- Discourse connectives, including words indicating concession (雖然 *although*, 但係 *but*…), causal relation (因為 *because*, 所以 *therefore*…), conditions (除非 *unless*, 不論 *whether*…), and hypothetical situations (如果 *if*, 就算 *even if*…).
- Speech planning features, including common fillers like 即係 *that is*, 其實 *actually*, etc., as well as speed and number of self-repairs, hesitation, lengthening and pauses in a speech sample.

4.2 Procedures

The frequency data obtained for the above linguistic features were tabulated and subject to Factor Analysis, following the process of multi-dimensional analysis discussed in Biber (1988, 1993). Factor Analysis is a kind of multivariate analysis which reduces a large number of features to a smaller set of factors based on their co-occurring patterns. In this study, SPSS was used as the tool to do this. With reference to the factors identified and the loadings associated with their component features, dimensional scores were computed for each type of text (in this case, speech produced by a certain role in a certain type of programme) with respect to each factor/dimension. These scores were based on the average of the sum of normalized frequencies for positively loaded features less that for negatively loaded features under a particular dimension for a given text type.

4.3 Preliminary Results

Four factors, corresponding to the dimensions in multi-dimensional analysis, were identified in the process. As demonstrated by Biber (1993), each dimension could be functionally interpreted according to the positive features and negative features associated with it. For example, the abundance of personal pronouns, especially first and second person pronouns, might indicate a high degree of interaction and involvement. Individual text types, or genres, could be characterised by not just one but many features which often co-occur or are simultaneously absent. Given the relatively small set of features of limited variety used in this pilot study, not all dimensions were found to associate with negative features. The possible functional interpretations of the identified dimensions with the corresponding positive and negative features are shown in Table 5.

Dimension	Positive features	Negative features
D1: Interaction, Involvement, Stance	Private verbs 1st person pronouns Wh-questions 2nd person pronouns	
D2: Uncertainty, Prediction	Speed Speech fillers Modals Yes/No questions	
D3: Elaboration, Explanation	Speech planning features Causal connectives 1st person pronouns	2nd person pronouns Yes/No questions Concession words
D4: Argumentative	3rd person pronouns Concession connectives Hypothetical connectives Causal connectives	

Table 5: Dimensions Identified

The speech samples were divided into four categories (or registers) by the two domains (current affairs and finance) and two roles (host and guest). The dimensional scores for each category along each dimension were computed. Figure 1 shows the comparison of the categories along the first dimension (D1) and Figure 2 shows their comparison along the second dimension (D2). In the figures, "Cur" stands for current affairs and "Fin" stands for finance.

It can be seen that guests and hosts in both domains are quite clearly distinguished by D1, which is characterised by private verbs, first and second person pronouns, and wh-questions. These are indicative of interaction, involvement and stance. Guests in interview programmes are often asked for their views on certain subjects, while hosts are expected to remain as neutral as possible.

D2, which is characterised by faster speed and abundance of speech fillers, modals and yes/no questions, reflects the uncertainty of the speakers and is likely to be associated with predictions rather than factual statements. This dimension thus singles out guests in financial programmes, who usually make predictions on financial matters and give investment advice. Table 6 shows a guest speaker turn from each domain of similar duration for a quick visualisation of the features found for D1 and D2. The relevant features are bolded and underlined.

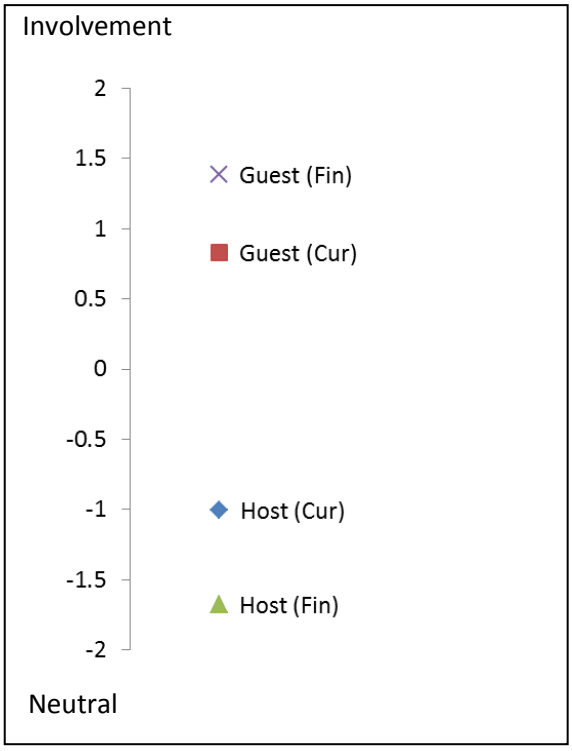

Figure 1: Comparison along D1

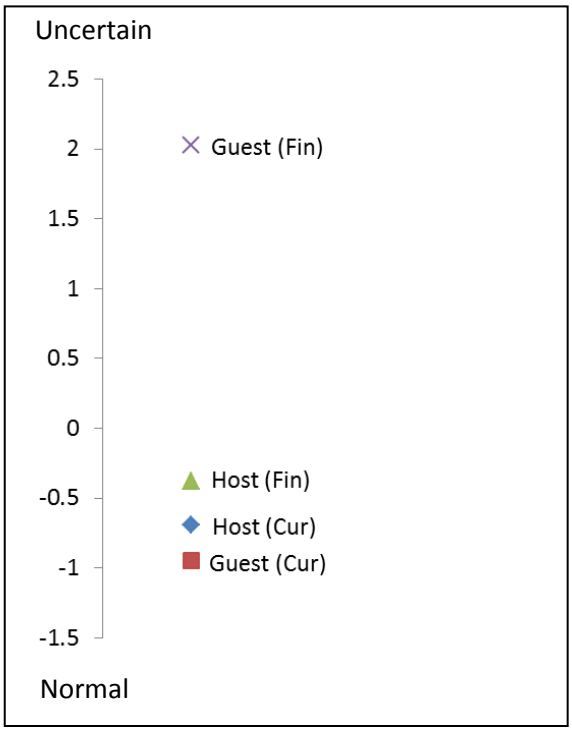

Figure 2: Comparison along D2

Dimension	Guest (Current affairs)	Guest (Finance)
D1	土地呢件事呢講親都好複雜嘅//因為有好多種--種類類嘅利益各個方面吓//一講親土地呢香港呢大家就即刻去^^向咗個錢字諗嘅//但係呢我哋其實講到土地嗰陣時一定係講全香港所有人嘅總體利益//嗱咁譬如話其實呢我哋人多嗰陣時點樣呢//就應該喺已經有嘅城鎮一路周邊俾佢擴充出去呢//呢個係最好嘅方法嚟嘅 aha//咁就 eh 因此呢我哋冇理由走去山卡啦度起啲樓㗎吓嘛	嗱汽車股嚟講嘅話呢我 諗其實係暫時嚟講仍然會係比嗰個大市呢困擾住//咁但係整體嚟講嘅我 覺得尤其是係長城啦或者係嗰個華晨呢//eh 長城食糊就靠 SUV 啦而華晨方面仍然係以佢嗰個比較高檔次嘅一啲 eh 豪華客車嚟講嘅話呢係受惠嘅//咁所以其實呢一類股份 eh..我 覺得就係逢低可以吸納啦//咁我 諗其實如果係嗰個跌定或者係內地嗰個銀根開始係寬^^放寬番嘅話呢//其實係可以吸納呢一類咁嘅股份囉
D2	土地呢件事呢講親都好複雜嘅//因為有好多種--種類類嘅利益各個方面吓//一講親土地呢香港呢大家就即刻去^^向咗個錢字諗嘅//但係呢我哋其實講到土地嗰陣時一定係講全香港所有人嘅總體利益//嗱咁譬如話其實呢我哋人多嗰陣時點樣呢//就應該喺已經有嘅城鎮一路周邊俾佢擴充出去呢//呢個係最好嘅方法嚟嘅 aha//咁就 eh 因此呢我哋冇理由走去山卡啦度起啲樓㗎吓嘛	嗱汽車股嚟講嘅話呢我諗其實係暫時嚟講仍然會係比嗰個大市呢困擾住//咁但係整體嚟講嘅我覺得尤其是係長城啦或者係嗰個華晨呢//eh 長城食糊就靠 SUV 啦而華晨方面仍然係以佢嗰個比較高檔次嘅一啲 eh 豪華客車嚟講嘅話呢係受惠嘅//咁所以其實呢一類股份 eh .. 我覺得就係逢低可以吸納啦//咁我諗其實如果係嗰個跌定或者係內地嗰個銀根開始係寬^^放寬番嘅話呢//其實係可以吸納呢一類咁嘅股份囉

Table 6: Comparison on D1 and D2 Features

Figure 3 and Figure 4 show the comparison of the categories along the third and fourth dimensions (D3 and D4) respectively. D3, with the abundance of causal connectives and speech planning features, is characteristic of guests in both domains who often need to elaborate and explain the views, especially in current affairs discussions. Hosts in current affairs programmes often pose concise questions and let the guest respond, whereas those in financial programmes may pose more elaborated questions, or may even express some of their personal views with considerably more interaction with the guest. The differentiation of the categories along D4, for argumentation, suggests that guests tend to speak with more logical reasoning than hosts, and this is more evident for guests in financial programmes than those in current affairs programmes.

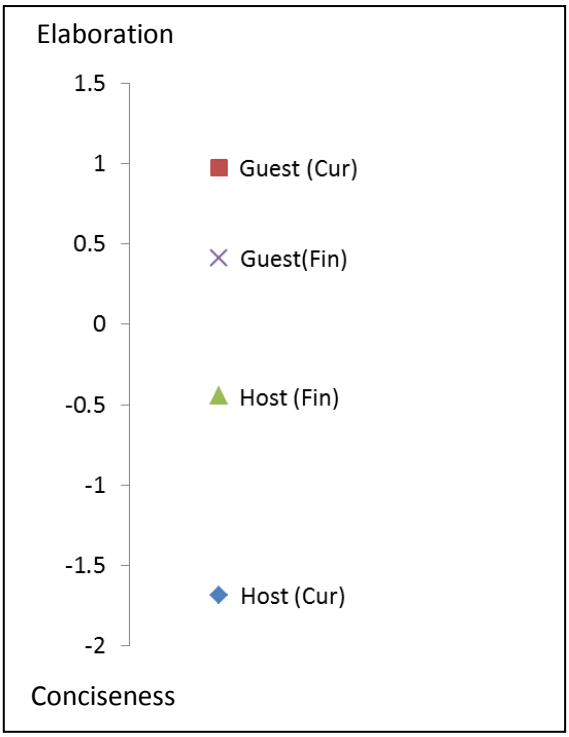

Figure 3: Comparison along D3

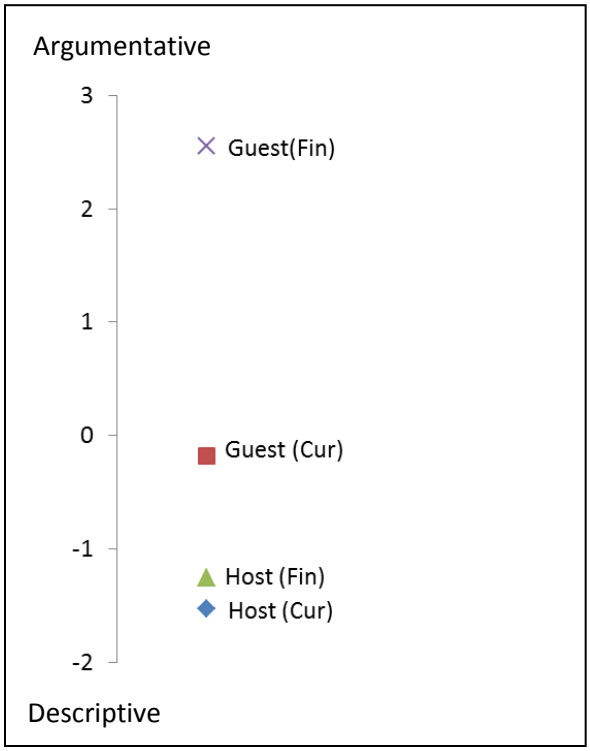

Figure 4: Comparison along D4

4.4 Potential Application

Given the feature co-occurrence patterns identified and their association with specific communicative functions, they can potentially be used as the features in a variety of approaches, including those based on rules or machine learning, for automatic subjectivity recognition and opinion type classification. Larger-scale data annotation is in progress, for opinions and other private states. The resulting annotated corpus is expected to provide more data as well as more variety of linguistic features for a more comprehensive multi-dimensional analysis. The relevant features for characterizing different categories of verbal comments will be applied in experiments on opinion mining.

5 Future Work and Conclusion

The preliminary study reported here suggested that multi-dimensional analysis is a promising approach to characterise opinionated speech samples and differentiate their sub-types based on communicative functions.

The immediate next step will expand the analysis to include more speech samples, possibly with a greater variety of domains and roles, to obtain more reliably distinguished dimensions, and to account for a wider range of opinion types and functions. So far we have relied mostly on lexical features, and more types of features will be necessary for a fuller picture of the genre characteristics of verbal comments. In particular, for lexical features we plan to add parts of speech, aspect markers, and sentence-final particles (which is very characteristic of Cantonese), and more importantly, for lexico-grammatical features we plan to include nominalisations, assertions, negation, and as far as possible, some discourse level features would be favoured. More tests on grouping and de-grouping the various features will be conducted and a more comprehensive analysis of the dimensions (with expanded datasets) will be done, for a descriptive account of Cantonese verbal comments as a specific spoken genre.

Another important direction will certainly be the application of the dimensions (and the features therein) and dimension scores for opinion mining. We have outlined their potential uses and experiments will be done when more annotated data for training and testing are ready. This future work is expected to showcase the synergy between corpus-based discourse analysis and opinion mining applications.

Reference

Banea, C., R. Mihalcea and J. Wiebe. 2010. Multilingual Subjectivity: Are More Languages Better? In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, Beijing, pp.28-36.

Biber, D. 1988. *Variation across Speech and Writing*. Cambridge University Press.

Biber, D. 1993. Using Register-Diversified Corpora for General Language Studies. *Computational Linguistics, 19(2)*: 219-241.

Conrad, S. and D. Biber. 2001. *Variation in English: Multi-Dimensional Studies*. Longman.

Esuli, A. and F. Sebastiani. 2006. SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. In *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC'06)*, pp.417-422.

Hu, M. and Liu, B. 2004. Mining and Summarizing Customer Reviews. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 168-177.

Kaufer, D. and S. Ishizaki. 2006. A corpus study of canned letters: mining the latent rhetorical proficiencies marketed to writers in a hurry and non-writers. *IEEE Transactions on Professional Communication, 49(3)*: 254-266.

Kim, S-M. and E. Hovy. 2006. Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, Sydney, pp.1-8.

Ku, L-W., Y-T. Liang and H-H. Chen. 2006. Opinion extraction, summarization and tracking in news and blog corpora. In *Proceedings of AAAI*.

Li, S., S. Ju, G. Zhou and X. Li. 2012. Active Learning for Imbalanced Sentiment Classification. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Jeju Island, Korea, pp.139-148.

Liu, B. 2010. Sentiment Analysis and Subjectivity. In N. Indurkhya and F. J. Damerau (Eds.), *Handbook of Natural Language Processing*. Boca Raton, FL: Chapman & Hall.

Lu, B., B.K. Tsou and T. Jiang. 2010. Supervised Approaches and Dependency Parsing for Chinese Opinion Analysis at NTCIR-8. In *Proceedings of the 8th NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*, Tokyo, Japan, pp.234-240.

Pang, B. and L. Lee. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval, 2(1-2)*: 1-135.

Pang, B., L. Lee and S. Vaithyanathan. 2002. Thumbs up? Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 79-86.

Polanyi, L. and A. Zaenen. 2006. Contextual Valence Shifters. In J.G. Shanahan, Y. Qu and J. Wiebe (Eds.) *Computing Attitude and Affect in Text: Theory and Applications*. Springer.

Somasundaran, S. and J. Wiebe. 2009. Recognizing Stances in Online Debates. In *Proceedings of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language of the AFNLP*, Singapore, pp.226-234.

Tsou, B.K., O.Y. Kwong, W.L. Wong and T.B.Y. Lai. 2005. Sentiment and Content Analysis of Chinese News Coverage. *International Journal of Computer Processing of Oriental Languages, 18(2)*: 171-183.

Turney, P. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In *Proceedings of ACL 2002*, Philadelphia, pp.417-424.

Wiebe, J. and T. Wilson. 2002. Learning to disambiguate potentially subjective expressions. In *Proceedings of the Conference on Natural Language Learning (CoNLL)*, pp. 112-118.

Wiebe, J., T. Wilson and C. Cardie. 2005. Annotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation, 39(2)*: 165-210.

Wilson, T., J. Wiebe and P. Hoffmann. 2005. Recognizing Contextual Polarity in Phrase-level Sentiment Analysis. In *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, pp. 347-354.

Zirn, C., M. Niepert, H. Stuckenschmidt and M. Strube. 2011. Fine-Grained Sentiment Analysis with Structural Features. In *Proceedings of the 5th International Joint Conference on Natural Language Processing*, Chiang Mai, Thailand, pp.336-344.