Automatic Speech Act Classification of Korean Dialogue based on the Hierarchical Structure of Speech Act Categories

Youngeun Koo
Dept. of German Linguistics & Literature, Sungkyunkwan University /
25-2, Sungkyunkwan-Ro, Jongno-Gu, Seoul, Korea
sarah8835@skku.edu

Jiyoun Kim
Dept. of German Linguistics & Literature, Sungkyunkwan University /
25-2, Sungkyunkwan-Ro, Jongno-Gu, Seoul, Korea
kite92@skku.edu

Munpyo Hong*
Dept. of German Linguistics & Literature, Sungkyunkwan University /
25-2, Sungkyunkwan-Ro, Jongno-Gu, Seoul, Korea
skkhmp@skku.edu

Abstract

Speech act is an utterance intention and should be understood correctly for a successful communication. However, in some cases, analyzing speech act of an utterance is not simple. This is because we can roughly grasp 'representative' speech act relatively easily through 'utterance-internal' features, but 'concrete' speech act varies depending on 'utterance-external' features. Therefore, this paper proposes a hierarchical structure of speech acts and a two-step classification method of speech acts, for a better understanding of the human conversation and the improvement of automatic speech act classification. The experiment, using Korean tutorial dialogues and telephone calls, showed 83% for the 1st step and 84% for the 2nd step, while using a flat structure showed 71% of accuracy.

1 Introduction

Speech act (SA) is an intention of an utterance. Austin (1962) argues that SA is "a functional unit in communication". For successful communication and correctly understanding the intention of an utterance, SA is very important. Understanding correct SA is crucial not only in a real-life, but also in the field of 'Intelligent tutoring systems(ITS)' as well as in various dialogue systems such as Apple's 'Siri' assistant and Amazon's 'Echo' speaker. Along with the increasing demand for such systems which can interact with human naturally, the system is expected to improve its performance through implementing a better SA classification method into it. However, it is often not easy to analyze it clearly. Sometimes, it is hard to describe SA because it varies according to in which situation the utterance is made. Besides, some utterances are difficult to be defined as one specific SA.

In this study, we examine various factors that disrupt SA analysis and claim the necessity of a hierarchical structure of SA categories. We propose the hierarchical structure of SA by comparatively analyzing two different corpora: Korean tutorial dialog and Korean telephone call.

This paper is organized as follows. In section 2 we look through related works of speech act theory. Section 3 shows some difficult cases of speech act classification and proposes the hierarchical structure of speech act categories to solve this problem. Then, in section 4, we present how we set the experiment of an automatic speech act classification to verify our methodology and then discuss the result of the experiment. Finally, we conclude this paper with future works.

* Corresponding author

Pacific Asia Conference on Language, Information and Computation (PACLIC 33), pages 432-441, Hakodate, Japan, September 13-15, 2019

2 Related Works

In pragmatics, the study of the use of language, speech act theory is one of the most important topics. Speech act theory was established by J. Austin and J. Searle. Austin (1962) insisted that a language of itself is an action and introduced the concept of a 'performative' sentence. Then, the works of Searle further developed the speech act theory. According to Austin (1962), depending on the 'force' that affects an utterance, there are three actions: locutionary act, illocutionary act, and perlocutionary act. Normally, we call illocutionary act as speech act.

Discovering and classifying SA categories are of great importance in that it allows us to understand our language use and real-life interaction. Nevertheless, to classify SA is a tough task, because an utterance can have diverse intentions depending on a situation.

Category	Explanation
Representatives	to commit the speaker to something's being the case, to the truth of the expressed proposition
Directives	attempts by the speaker to get the hearer to do something
Commissives	commit the speaker to some future course of action
Expressives	express the psychological state specified in the sincerity condition about a state of affairs specified in the propositional content
Declaratives	brings about the correspondence between the propositional content and reality

Table 1. Speech Act Categories of Searle (1976)

Table 1 is the SA categories of Searle (1976). Later, this had a great influence on further works. Fraser (1974), Kats (1977), Bach and Harnish (1979) and Leech (1983) attempted to improve categories of Searle (1976). In European countries, researchers concentrated on 'sub-classifying' SAs of Searle (1976) (Lee, 2015).[1]

Among SAs of Searle (1976), Hindelang (1978, 1981) concentrated on directives: demanding and questioning directive. Hindelang (1978) subdivided demanding directive into 18 SAs with criteria, such as 'obligation of counterpart to carry out the request' and 'relationship between the speaker and the counterpart'. Similarly, Hindelang (1981) sub-classified questioning directive into 10 SAs. Next, Rolf (1983) subdivided representatives into 36 SAs with two criteria: 'existence of its preceding speech act' and 'speaker's attitude toward the information'. Furthermore, Graffe (1990) sub-classified commissives into 'sp1-preferred type', 'sp2-preferred type' and 'complex type', depending on who has an interest in realizing the commissive. Lastly, Marten-Cleef (1991) dealt with sub-classifying expressives with respect to 'speaker and counterpart's attitude and judgment on the uttered situation'.

These works enabled not only better understanding of SAs of Searle (1976), but also suggesting various SAs that can appear in our real life. Yet, an empirical investigation on whether each SA actually appears in real communication is not completed.

Speech act theory is also studied in the field of computer science. Lampert et al. (2006) and Qadir et al. (2011) adopted 5 SA categories of Searle (1976). Kim (2006) and Buckley et al. (2008) utilized the DAMSL(Dialog Act Markup in Several Layers) (Core and Allen, 1997) tag-set. Some studies like Lee et al. (1997) and Bayat et al. (2016) proposed their own SA categories.

3 Multi-level Speech Act Categories

SA categories proposed in previous researches are mostly in a flat structure. In other words, all categories are on the same level. In fact, some studies show an approach to deviate from a flat structure. Core and Allen (1997) suggested each utterance having multiple SA labels in the DAMSL tag-set. They analyzed SA of an utterance in 3 layers: forward communicative function, backward communicative function and utterance feature.[2]

[1] Kohl and Kranz (1992) explain why sub-classifying 'the global speech act type' of Searle's taxonomy is necessary. First, some speech acts need to be divided in much detail. When we focus on the speech act types of Searle (1976), the types are distinguishable each other. But, in fact, in many cases, the boundary between the types is ambiguous. Second, 'global speech act type' is insufficient to explain the link between the utterance intention and its uttered expression. Third, through sub-classifying 'global speech act type', we can ascertain whether the overall taxonomy is well-founded and plausible.

[2] Some studies describe the top label of DAMSL in 4 dimensions including 'communicative status', 'the

However, they focus on the 'role of an utterance' in a dialogue flow, rather than the 'intention of an utterance'.

Kang et al. (2013) also took an approach to use hierarchical structure in the SA classification. 'Question type', 'response type' and 'other type' are suggested as 3 SA types of the first layer. By structuralizing SAs into the hierarchical structure for SA classification, the accuracy reached 85% in hotel, airline, tour reservation corpus and 91% in schedule management corpus. But the corpora of Kang et al. (2013) are mostly composed of 'question-answer' pair. Thus, it remained yet as a limitation, that the hierarchical structure is highly restricted to the 'question-answer' paired domain corpus.

This paper proposes the multi-level hierarchical structure of SA for a better understanding of human conversation and the improvement of automatic SA classification. In the following section, we look at some difficult cases of SA analysis and discuss the reasons why the hierarchical structure of SA is necessary.

3.1 Importance of the Hierarchical Structure of Speech Act Categories

In some cases, it is difficult to classify an utterance into a specific SA. First, each person can understand SA differently. To be specific, not all people understand the utterance intention 'same', as 'one' specific SA.

(1) A: He is our new teacher!
B: It's not him.
(*'disagree', 'inform'*)[3]

In example (1), speaker B's utterance can be understood as either disagreeing with the speaker A's assertion or informing new information to speaker A. It depends on how people read and perceive this utterance. Even if people perceive the utterance similarly, not all people would denote its SA with the same SA category. People can denote (1B) as 'inform', 'assert', 'disagree', 'dispute', 'react' or 'response'. To solve this, these similar SAs should be grouped into the same type.

Second, we understand SAs differently depending on a communication situation. Examples (2) ~ (5) explain this in more detail.

(2) A: This experiment is due tomorrow!
(*'command', 'request'*)
B1: Yes, Mrs. Jones.
B2: Okay, no need to worry.

In example (2), SA of speaker A's utterance differs depending on the relationship between two speakers. If speaker A is at a higher position or can impose a sanction against, or if speaker B is bound to perform what speaker A demands, the utterance of speaker A is a 'command' (Hindelang, 1978). However, if not, the utterance would be a 'request'.

(3) A: Can you pass me that?
(*'question', 'request'*)

In example (3), SA of speaker A varies according to the situation where the utterance is made. If it happens during a doctor's appointment, it is a 'question'. Instead, if it happens in daily life and speaker A is pointing something close to another speaker, it is likely to be a 'request'.

(4) A: You know I hate messing up the house, don't you?
(*'criticism', 'warning'*)
B1: I'm sorry, mom.
B2: Yes, I'll keep in mind.

In example (4), speaker A's intention differs by to whom the utterance is made. If it is toward a kid who messed up the house, speaker A intends to 'criticize' the kid. However, if speaker A utters toward the other kid who did not mess up the house, speaker A intends to 'warn' this kid.

(5) A: He gave a Christmas present to his boss again this year!
(*'compliment', 'criticism'*)

In example (5), the utterance of the speaker A can be interpreted differently depending on the speaker A's attitude. If the speaker A is favorable to something/someone, which he/she is talking about, his/her utterance is a 'compliment'. Instead,

information level', 'forward-looking function', and 'backward-looking function' (Fisel, 2007). For this paper, whether the number of dimensions in top label is 3 or 4 does not make much difference. Here, we focus on the fact that DAMSL attempted multiple-labeled annotation. Therefore, this paper follows the description in Core and Allen (1997).

[3] Speech acts are marked in italic font.

if the speaker A is hostile, the utterance is a 'criticism'.

Lastly, SA analysis is often difficult since SAs proposed so far inevitably overlap somehow each other. This is because it is almost impossible to establish SA categories fully complimentarily.

For these several reasons, we insist that SAs should be understood and classified automatically based on the hierarchical structure of SAs, rather than on the flat structure.

3.2 Proposing the Hierarchical Structure of Speech Act Categories

The proposed hierarchical structure is designed through empirical analysis of two different domains of conversation corpora: Korean tutorial dialogues and Korean telephone calls, both built by the National Institute of the Korean Language (NIKL). They are comprised of two separate one-to-one conversations. The tutorial dialogue corpus consists of 1,833 utterances between a teacher and a student. Most of the utterances in this tutorial dialogue were collected in a math class. The telephone call corpus consists of 2,005 utterances between a graduate student and an undergraduate student. Table 2 shows further information about corpora.

Speech Act	Utterance		Speech Act	Utterance	
	Tut	Tel		Tut	Tel
Accept	3	1	Exclamation	56	153
Acknowledge	221	254	Greeting	0	4
Agree	45	44	Guess	42	65
Answer	288	185	Induce	110	28
Apologize	2	1	Inform	360	496
Ask-answer	338	269	Praise	1	8
Ask-confirm	44	28	Reject	3	7
Assert	161	293	Request	33	18
Avoid	13	9	Suggest	28	23
Command	14	2	Thank	0	4
Correct	10	9	Will	11	8
Criticism	32	36	Wish	2	11
Disagree	16	49			

Table 2. Number of Speech Acts in each Corpus (Tut: tutorial dialogue, Tel: telephone call)

This paper aims at organizing utterance intentions overall, rather than sub-classifying only one particular SA. Also, instead of designing completely new SA categories, we utilize SA categories proposed in Koo (2018) to build the hierarchical structure of SAs.

At first, we inspected which SAs are a 'representative speech act' that represents SAs with similar features. A representative SA can be understood easily without a complex analysis of the utterance and the conversation. We did not adopt 5 deductively derived SAs of Searle (1976) as our representative SA.[4] Instead, we investigated the corpora to determine the representative SAs.

We analyzed the frequency of each SA to judge whether it is a domain-independent representative SA. This enabled us to presume the status or position of each SA. The one, which is biased in one domain or does not appear often in both domains, cannot be considered as a representative category. In this case, we searched for other similar SAs and put them together in the same upper class, the representative SA.

For example, in the tutorial dialogues, there are many 'avoid' utterances compared to the telephone calls. Considering that 'avoid' is one of the negative reactions to the counterpart's demand, 'avoid' shares many features with 'reject' or 'disagree'. Also, there are many utterances that a teacher 'induces' a student to respond in the tutorial dialogue. Since, the speaker intends to get information through inducing, 'induce' can be combined into an upper SA category with 'ask-answer' and 'ask-confirm'.

On the other hand, 'assert' appears relatively a lot in the telephone call corpus. This is quite easy to infer, because telephone calls mostly occur when one of the speakers has something to the other to talk about. Since 'assert' and 'inform' are similar, in that they both deliver information to others, we can combine them together.

As a result, we propose a two-levelled hierarchical structure of SAs as Figure 1. [5] Representative SAs, the upper level SAs, are marked with all letters capitalized. Concrete SAs, the lower level SAs, are marked with the capitalized first letter.

[4] Pöring and Schmitz (1999) brought out a hierarchical structure of speech act by borrowing 5 speech act types of Searle (1976) and categorizing them into 3 types of speech act in the first class: representatives as 'information-searching', directives and commissives as 'obligative', and expressives and declaratives as 'constitutive'.

[5] Indeed, there might be some speech acts, omitted in this hierarchical structure. However, this is left as a further study. In this paper, we attempt to explore possibility of improved automatic speech act classification by using a hierarchical structure of speech act categories rather than a flat structure.

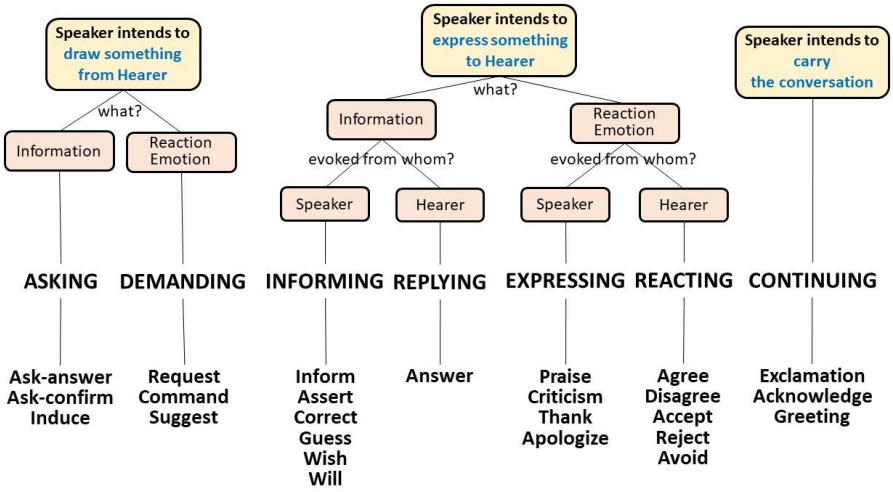

Figure 1. Proposed Hierarchical Tree Structure of SAs

There are a few important issues in the proposed hierarchical structure. This structure was made not only by an inductive approach, by empirically analyzing corpus, but also by a deductive approach, by theoretically analyzing our utterance intentions in terms of linguistic interactions. When making a hierarchical structure, it is important to understand an utterance as a part of an interaction between two or more speakers (Kang, 2004; Franke, 1990). Also, 'what' is involved in the interaction is important. So we considered it as one of the criteria in structuring SAs.[7]

Representative Speech Act	Concrete Speech Act
ASKING	Ask-answer, Ask-confirm, Induce
DEMANDING	Request, Command, Suggest
INFORMING	Inform, Assert, Correct, Guess, Wish, Will
REPLYING	Answer
EXPRESSING	Praise, Criticism, Thank, Apologize
REACTING	Agree, Disagree, Accept, Reject, Avoid
CONTINUING	Exclamation, Acknowledge, Greeting

Table 3. Proposed Hierarchical Structure of SAs

Once representative SA categories are determined, concrete SA categories can be mapped into their corresponding representative SA category. Table 3 shows how each SA of Koo et al. (2018) is linked to 7 representative SAs. Most of these subgroups are self-explanatory, except for the following cases.

First, 'Wish' and 'Will' are a concrete SA in the 'INFORMING' representative SA. Whereas, Rolf (1983) considered them as the expressives. Compared to other SAs, 'Will' and 'Wish' are the multi-faceted category. This causes researchers to analyze them differently. For this study, we focused on their essential intent, conveying some information, rather than their slightly different nuance, 'desiring to achieve' and 'hoping to achieve'.

Second, we mapped 'Answer' to the 'REPLYING' representative SA. In many studies, 'Answer' is regarded as 'INFORMING'. Nevertheless, 'REPLYING' and 'INFORMING' are different, especially with respect to which SA they pairs with. Therefore, we mapped 'Answer' to 'REPLYING', since we propose SA categories not only for understanding the human conversation, but also for automatically classifying SAs.

Third, 'Acknowledge', 'Exclamation' and 'Greeting' fall into the 'CONTINUING' SA. This paper emphasizes the role of a SA in a linguistic interaction. In this respect, we consider 'Acknowledge', 'Exclamation' as a neutral and an emotional reception signal, and 'Greeting' as a socially conventional expression in a conversation.

[7] Bunt (2013) classified 'general-purpose functions' into 'information-transfer functions' and 'action-discussion functions'.

4 Experiments

We conducted an experiment to verify our classification method and the feasibility of the proposed hierarchical structure. We also compared the accuracy of the automatic SA classification between using the flat structure and the hierarchical structure.

We utilized WEKA version 3.8 for machine learning and 'Support Vector Machine (SVM)' as a machine learning algorithm. As an evaluation measure, we employed '10-fold cross validation'. We used two corpora mentioned earlier, Korean tutorial dialogues and Korean telephone calls, as a training corpus. Each corpus consists of 1,833 utterances and 2,005 utterances, respectively.

The experiment used unigram, bigram, which are extracted from each utterance, and linguistic features proposed in Koo et al. (2018) for the automatic classification: 9 sentence features and 4 context features.

Feature Type	Feature Name
Sentence feature	sent_type, tense, sub_person, negation, interrogative, verb_num, sent_length, first two words, last two words
Context feature	prev SA, prev SA_oppo, SA pair, turn chng

Table 4. Linguistic Features for Speech Act Classification

First, we conducted experiments on two corpora together to verify the methodology of this paper. We compared the accuracy of SA classification with the flat structure and with the hierarchical structure. Table 5 shows the result of the experiment.

	Tutorial dialogue + Telephone call	
	1st level	2nd level
Baseline (flat)	71.04	
Proposed (hierarchical)	83.44	84.47

Table 5. Accuracy of the Experiment on Combined Corpus (%)

The accuracy of the '1st level' of the hierarchical structure indicates the accuracy of classifying an utterance into a representative SA. Similarly, the accuracy of the '2nd level' indicates the accuracy of classifying an utterance of a specific representative SA into a concrete SA. For now, we designed two levels of classification separately, aiming for a preliminary examination on our hierarchical structure.

As a result, the accuracy of the 1st level was 83.44% and the accuracy of the 2nd level was 84.47%. Since we classify SAs in two steps, the first step is very important. If the first step is incorrectly analyzed, then the next step is bound to fail.

To evaluate our method, we conducted an additional experiment in a deep learning approach. We used 'Convolutional Neural Networks (CNN)' (Kim, 2014) to classify utterances into 7 representative SAs by setting the model as following: filter windows of 2, 3, 4, batch size of 32, epoch of 50 and learning rate of 0.01. Also, same with the evaluation measure of the earlier experiment using SVM, randomly selected 10% of the training data is used as the test data.

SVM	CNN
83.44	81.75

Table 6. Accuracy of the Experiment on the 1st Level (%)

The accuracy of the CNN model for classifying utterances into the representative SA is 81.75%. Of course, this result will improve, if we elaborate the model. Nonetheless, even with this preliminary experiment, it is still enough to figure out the feasibility of the proposed representative SAs. Moreover, through comparing the accuracy of the feature-based machine learning approach and the deep learning approach, we could conclude that linguistic features of Koo et al. (2018) perform nearly as good as a deep learning model.

In Table 5, the performance of the 2nd level is comparatively lower than the 1st level, though an utterance is classified into 7 classes on the 1st level and 4 classes in average on the 2nd level. This happens to be attributed to the features for the SA classification. Specifically, the features that we used are not suitable or sufficient to classify a concrete SA of an utterance. Following examples are the incorrectly classified utterances.

Num	Utterance	Predicted SA	Actual SA
1 (Tel)	A: 냉면 좀 사줘봐. Buy me some cold noodles. (nayng-myen com sa-cwe-pwa.)	Suggest	Request
2 (Tut)	A: 숙제니까 알아서 적어. It's the homework, so write it down by yourself. (swuk-cey-ni-kka al-a-se cek-e.)	Request	Command
3 (Tel)	A: 니 좀 유종의 미 좀 바꿔요! You should change that 'end well'! (ni com yu-cong-uy mi com pa-kkwe-yo!)	Suggest	Command

Table 7. Incorrectly Classified Utterances of 'DEMANDING' type

In example (1), the Korean expression '~(해)봐(~(hay)pwa)' carries a meaning of suggesting something to someone for his/her sake. However, as this expression is made in the utterance, which is in the conversation between the two close friends and uttered toward the speaker who has more ability to perform something, it is more likely to be a 'Request' rather than a 'Suggest'. Linguistic features derived from the utterance are usually enough to understand the intention of an utterance quite roughly. Nevertheless, they are not enough to understand the intention of an utterance in precise. In other words, 'utterance-internal' features are apt for classifying representative SA, but not for concrete SA. To catch precisely the utterance intention, we need more detailed, 'utterance-external' features.

Similar to example (1), examples in (2) and (3) also require detailed information about the participants of the conversation, in order to be correctly classified. Example (2) is classified automatically as a 'Request'. However, when there is detailed information about two speakers, a teacher and a student, it can give us a clue to classify the SA correctly as a 'Command'.

In addition, considering only the information given from the utterance, the utterance that ends with Korean honorific expression '~요(~yo)', like example (3), could be a 'Suggest'. However, when the information is given that this conversation is between two close friends, the example (3) might be classified correctly.

Num	Utterance	Predicted SA	Actual SA
4 (Tel)	A: 그래서 피곤한가 보지. Guess that's why (he's) tired. (ku-lay-se phi-kon-han-ka po-ci.)	Inform	Guess
5 (Tel)	A: 니는 비디오 언제 산 거야? When did you buy videos? (ni-nun pi-ti-o en-cey san-ke-ya?)	Ask-answer	Ask-answer
	B: 중학교 때는 없었던 거 같은데? I guess it did not exist during middle school. (cwung-hak-kyo ttay-nun eps-ess-ten ke kath-un-tey)	Guess	Inform

Table 8. Incorrectly Classified Utterances of 'INFORMING' type

Example (4) and (5B) show the difficulty of distinguishing 'Inform' from 'Guess'. Whether the speaker has knowledge or information about the subject of the utterance or not, determines the concrete SA. Especially, the 'utterance-external' features could be employed usefully, when the 'utterance-internal' features, such as Korean expression '~거 같다(ke kath-ta)' that indicates an uncertain attitude, in example (5B), might affect a SA classification wrongly.

Next, we conducted an additional experiment on two corpora separately. This experiment is to test whether the performance of a SA classification based on the hierarchical structure can be affected by the domain of the training corpus or not. Table 9 shows the results of the experiment.

	Tutorial dialogue		Telephone call	
	1st level	2nd level	1st level	2nd level
Baseline (flat)	70.03		71.40	
Proposed (hierarchical)	82.11	83.26	86.25	87.18

Table 9. Accuracy of Experiments on each Corpus (%)

As to the accuracy of the baseline, the two corpora show a similar result. However, when we look at the accuracy of the proposed method, the results are different. Telephone call corpus has a higher accuracy on both levels, compared to a tutorial dialogue. This seems due to the complexity of an utterance in each corpus. In the tutorial dialogue, a considerable amount of utterances are not directly related to the conversation between a teacher and a student (Koo, 2018). For example, such utterances are the readings of a textbook by the speaker. Likewise, utterances in the tutorial dialogue are relatively complex, which makes it hard to train input sentences. Whereas, telephone calls are simpler and relatively restricted types of conversation patterns appear in telephone calls.

On top of that, the accuracy of the 2nd level of the tutorial dialogue is low, compared to that of the telephone call. This is presumed to be caused by the lack of data. In fact, for an accurate experiment, fair and balanced amount data for each representative and concrete SA is required. However, as Table 2 shows, the tutorial dialogue particularly lacks data.

5 Conclusion

In this paper, we proposed the hierarchical structure of SA categories by comparatively analyzing the corpus of two different domains: Korean tutorial dialogues and Korean telephone calls. With these corpora, we conducted an experiment of a SA classification using the hierarchical structure. On the 1st level where an utterance is classified into a representative SA, the accuracy of the classification was 83%. On the 2nd level where an utterance of a specific representative SA is mapped into a concrete SA, the accuracy showed 84%. We also discussed the results of the experiment with some examples of incorrectly classified utterances.

Through the experiment, we can infer that the hierarchical structure is adequate for an automatic SA classification. However, a more elaborated analysis of SA categories is necessary. In future works, we plan to verify the SA categories of this study in more detail. The linguistic motivation of the proposed SAs must be investigated further.

In addition, as mentioned earlier in this paper, to classify SA on a concrete level, more features are needed. 'Utterance-external' features like the information about participants of the conversation is necessary for a more sophisticated method for a SA analysis. Not only that, various conversational analysis based features are presumed to be useful.

Above all, for the completeness of this methodology, we plan to connect two steps of the SA classification and identify the performance of the hierarchical structure for automatic SA classification more accurately.

Acknowledgments

This work was partly supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (R7119-16-1001, Core technology development of the real-time simultaneous speech translation based on knowledge enhancement) and National Research Foundation of Korea (NRF) grant funded by the Korean government(MOE) (2018H-1A2A-1062645, Global Ph.D. Fellowship Program).

References

Andrew Lampert, Dale Robert, and Paris Cécile. 2006 Classifying speech acts using verbal response modes. In Proceedings of the Australasian Language Technology Workshop 2006. 34-41.

Ashequl Qadir, Ellen Riloff. 2011. Classifying sentences as speech acts in message board posts. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing(EMNLP), Association for Computational Linguistics, 748-758.

Berken Bayat, Christopher Krauss, Agathe Merceron, Stefan Arbanowski. 2016. Supervised Speech Act Classifica-tion of Messages in German Online Discussions. In: Proceedings of the International Florida Artificial Intelligence Research Society Conference(FLAIRS Conference), 204-209.

Bruce Fraser. 1974. An examination of the performative analysis. Papers in Linguistics, 7, 1-40.

Changwoo Kang. 2004. Possibility of Subclassification of Speech Act Types and Its Problem. Journal of Germanic Linguistics, 9(0):195-215. (in Korean)

Eckard Rolf. 1983. Sprachliche Informationshandlungen. Göppingen. (in German)

Geoffrey N. Leech. 2016. Principles of pragmatics. New York: Longman Inc.

Götz Hindelang. 1978. Auffordern. Die Untertypen des Aufforderns und ihre sprachlichen Realisierungs-formen. Göppingen. (in German)

Götz Hindelang. 1981. Zur Klassifikation der Frage-handlungen. In: Hindelang, Götz/Zillig, Werner (Hrsg.): Sprache: Verstehen und Handeln. Bd. 2. Tübingen, 215-226. (in German)

Harry Bunt. 2005. A Framework for Dialogue Act Specification. In: Paper presented at the 4th Joint ISO-SIGSEM Workshop on the Representation of Multimodal Semantic Information, Tilburg.

Hyeyong Lee. 2015. Korean Expressive Speech Acts. Youkrack, Seoul. (in Korean)

Hyunjung Lee and Jungyun Seo. 1997. Analysis of Speech-acts for Korean Dialog Sentences. Human and Language Technology, 24(2 Ⅱ):259-262. (in Korean)

Jerrold J. Katz. 1977. Propositional Structure and Illocutionary Force: a study of the contribution of sentence meaning to speech acts. Hassocks: Harvester Press.

John Austin. 1962. How to Do Things with Words. The William James Lectures Delivered at Harvard University in 1955, Clarendon Press.

John Searle. 1976. A classification of illocutionary acts. Language in society, 5(1), 1-23.

Jürgen Graffe. 1990. SICH FESTLEGEN UND VER-PFLICHTEN: Die Untertypen kommissiver Sprechakte und ihre sprachlichen Realisierungs-formen. Münster/New York. (in German)

Kent Bach and Robert M. Harnish. 1979. Linguistic Communication and Speech Acts. Cambridge: The Mit Press.

Mark Buckley, Magdalena Wolska. 2008. A classification of dialogue actions in tutorial dialogue. In Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1, 73-80.

Mark Fisel. 2007. Machine learning techniques in dialogue act recognition. Estonian Papers in Applied Linguistics, 3, 117-134.

Mark G. Core and James Allen. 1997. Coding Dialogs with the DAMSL Annotation Scheme, Association for Advancement of Artificial Intelligence(AAAI), 56:28-35.

Mathias Kohl und Bettina Kranz. 1992. Untermuster globaler Typen illokutionärer Akte. Zur Untergliederung von Sprechaktklassen und ihrer Beschreibung. In: König, Peter-Paul und Helmut Wiegers (Hrsg.): Sprechakttheorie. Münster: LIT. S. 1-44. (in German)

Min-Jeong Kim, Kyoung-Soo Han, Jae-Hyun Park, Young-In Song and Hae-Chang Rim. 2006. Dialogue Act Classification for Non-Task-Oriented Korean Dialogues, Human and Cognitive Language Technology, 2006(10):246-253. (in Korean)

Ralf Pörings and Ulrich Schmitz. 1999. Sprache und Sprachwissenschaft: eine kognitiv orientierte Einführung. Gunter Narr Verlag. (in German)

Sangwoo Kang Youngjoong Ko and Jungyun Seo. 2013. Hierarchical Speech-act Classification for Discourse Analysis, Pattern Recognition Letters, 34(11):1119-1124.

Susanne Marten-Cleef. (1991). GEFÜHLE AUS-DRÜCKEN. Die expressiven Sprechakte. Göppingen. (in German)

Walther Kindt. 2009. Pragmatik: die handlungs-theoretische Begründung der Linguistik. In H. M. Müller (Ed.), Arbeitsbuch Linguistik. UTB: Vol. 2169. Paderborn: Schöningh. 289-305. (in German)

Wilhelm Franke. 1990. Elementare Dialogstrukturen. Darstellung, Analyse, Diskussion. Tübingen. (in German)

Yoon Kim. 2014. Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1746–1751.

Youngeun Koo. 2018. A Study on Speech Act and Automatic Speech Act Classification of German Dialog (Master's Thesis). Sungkyunkwan University, Seoul. (in Korean)

Youngeun Koo, Jiyoun Kim, Munpyo Hong, and Youngkil Kim. 2018. A Linguistic Study of Speech Act and Automatic Speech Act Classification for Korean Tutorial Dialog. Journal of KIISE, 45(8), 807-815. (in Korean)