UCLA
UCLA Electronic Theses and Dissertations

Title
Display of Listenership in Korean Conversation

Permalink
https://escholarship.org/uc/item/0hc323ph

Author
Won, Hyobin

Publication Date
2022

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Display of Listenership in Korean Conversation

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Asian Languages and Cultures

by

Hyobin Won

2022

© Copyright by

Hyobin Won

2022

ABSTRACT OF THE DISSERTATION

Display of Listenership in Korean Conversation

by

Hyobin Won

Doctor of Philosophy in Asian Languages and Cultures

University of California, Los Angeles, 2022

Professor Sung-Ock Shin Sohn, Chair

This dissertation investigates the display of listenership in Korean conversation that is achieved through manipulation of reactive responses. Particularly, it focuses listener's affective and epistemic stance towards the prior turn in the course of telling. The stance display is illustrated by analyzing the prosodic and multimodal features of reactive responses. The analysis is conducted on the two major types of Korean reactive responses; yes/nod type and oh/really type. The findings indicate that through yes/nod type responses, listeners display their interest towards the prior turn and show different levels of engagingness by managing the token type and the number of repeats. In terms of epistemic stance, listeners express that they are better informed than before while they are still on a path between K– and K+. Listeners manage the degree of informedness and readiness to obtain further information through the lexical choice, the number of repeats, and the co-occurring nods. With oh/really type responses, listeners not

only display their interest elicited by the process of realization but also express surprise by carefully designing prosody and facial expressions. Epistemically, they position themselves from K- to K+ with micro-management of pitch configuration. Through the close examination on reactive responses, this dissertation illustrates how listeners play their role by actively engaging in conversation.

The dissertation of Hyobin Won is approved.

Hongyin Tao

Shoichi Iwasaki

Sun-Ah Jun

Sung-Ock Shin Sohn, Committee Chair

University of California, Los Angeles

2022

To myself

TABLE OF CONTENTS

LIST OF TABLES viii
LIST OF FIGURES ix
ACKNOWLEDGMENTS xi
VITA xiii

CHAPTER 1: INTRODUCTION 1
1.1. Goals of the Study 1
1.2. Data 2
1.3. Methodology 3
1.4. Organization of the Study 4

CHAPTER 2: THEORETICAL BACKGROUND 5
2.1. Listening and Listenership 5
2.2. Reactive Response 6
2.2.1. Terminologies and Classifications 7
2.2.2. Form 9
2.2.3. Prosody 11
2.3. Stance 12
2.3.1. Affective Stance 12
2.3.2. Epistemic Stance 14
2.3.3. Affiliation and Alignment 15
2.4. Storytelling 16
2.5. K-ToBI and Boundary Tones 18

CHAPTER 3: THEORETICAL FRAMEWORK 21
3.1. Listenership to Speakership 21
3.2. Essential Aspects 23
3.3. Stances 26

CHAPTER 4: YES/NOD TYPE RESPONSES 28
4.1. Lexical token "*ung*" and "*e*" 30
4.2. Vocalization "*mm*" and nods 40
4.3. Multiple saying "*mm*" and "*e*" 50
4.4. Summary and Discussion 63

CHAPTER 5: OH/REALLY TYPE RESPONSES 67
5.1. Free-standing "*a*" 69
5.1.1. Rising-falling "*a*" 69
5.1.2. Falling "*a*" 76
5.2. Lexical token "*cincca*" and "*kulay*" 82
5.3. *a*-prefaced Structure "*a cincca*" and "*a kulay*" 97
5.4. Summary and Discussion 115

CHAPTER 6: CONCLUSION 120
6.1. Summary of Findings 120
6.2. Suggestions for Future Research 123
6.3. Implication of the Study 124

APPENDIX A..... 125
APPENDIX B 126
REFERENCES 127

LIST OF TABLES

Table 1. Participant Information 3

Table 2. Categorizations of reactive responses 8

Table 3. Classification of English reactive tokens by Tsuchiya (2013) 10

Table 4. Classification of Korean reactive tokens by Young and Lee (2004) 10

Table 5. Findings on "*mm*","*un*g", "e" tokens and nods 65

LIST OF FIGURES

Figure 1. Epistemic Stance represented in terms of Epistemic Gradient by Heritage (2012) 14

Figure 2. Intonational Structural of Seoul Korean 18

Figure 3. Praat information of the first "*ung*" token (HS) and its prior turn (HB) 34

Figure 4. Praat information of the second "*ung*" token (HS) and its prior turn (HB) 34

Figure 5. Praat information of the first "*e*" token (HS) and its prior turn (HB) 35

Figure 6. Praat information of the second "*e*" token (HS) and its prior turn (HB) 36

Figure 7. Praat information of the third "*e*" token (HS) and its prior turn (HB) 37

Figure 8. Praat information of the first "*mm*" token (HS) and its prior turn (HB) 44

Figure 9. Praat information of the second "*mm*" token (HS) and its prior turn (HB) 44

Figure 10. Praat information of the "*e*" token (HS) and its prior turn (HB) 45

Figure 11. Praat information of the "*mm mm mm mm*" token (GA) and its prior turn (HS) 55

Figure 12. Seating arrangement of the participants; GA (left) and HS (right) 57

Figure 13. Listener GA's head gesture and facial expression 57

Figure 14. Praat information of the "*e e e e e*" token (GA) and its prior turn (HS) 58

Figure 15. Praat information of the "*e e*" token (GA) and its prior turn (HS) 60

Figure 16. Praat information of the prior turn of the "*a*" token (CY) 72

Figure 17. Praat information of the "*a*" token (HS) 72

Figure 18. Seating arrangement of the participants; CY (left) and HS (right) 75

Figure 19. Listener HS's head and eye movement 75

Figure 20. Praat information of the prior turn of the the "*a*" token (HS) 78

Figure 21. Praat information of the first "*a*" token (CY) 78

Figure 22. Praat information of the prior turn of "*kuleh-keyss-ne*" (HS) 80

Figure 23. Praat information of "*kuleh-keyss-ne*" (CY) 81

Figure 24. Praat information of the prior turn of the first "*cincca*" (PT) 89

Figure 25. Praat information of the first "*cincca*" (HB) 89
Figure 26. Seating arrangement of the participants; HB (left) and PT (right) 90
Figure 27. Listener HB's nods and facial expressions 90
Figure 28. Praat information of the prior turn of the second "*cincca*" (PT) 91
Figure 29. Praat information of the second "*cincca*" (HB) 91
Figure 30. Praat information of the prior turn of the second "*kulay*" (PT) 93
Figure 31. Praat information of the second "*kulay*" (HB) 93
Figure 32. Listener HB's nods and facial expressions 94
Figure 33. Listener HB's nods and eye movement 96
Figure 34. Praat information of the first "*a cincca*" structure (GA) and its prior turn (HS) 102
Figure 35. Listener GA's facial expressions 104
Figure 36. Seating arrangement of the participants; GA (left) and HS (right) 105
Figure 37. Praat information of the second "*a cincca*" structure (GA) and its prior turn (HS) ..105
Figure 38. Praat info of the first "*a cincca*" (left) and second "*a cincca*" structure (right) 106
Figure 39. Listener GA's head movement 107
Figure 40. Praat information of the first "*a kulay*" structure (GA) and its prior turn (HS) 108
Figure 41. Listener GA's head movement 110
Figure 42. Praat information of the second "*a kulay*" structure (GA) and its prior turn (HS) ... 111
Figure 38". Praat info of the first "*a cincca*" (left) and second "*a cincca*" structure (right) 113

ACKNOWLEDGMENTS

This dissertation would not have been able to be completed without the tremendous assistance given to me. I first wish to express my deep appreciation for the endless support from my committee members.

Professor Sohn, I am truly grateful to you for giving me this opportunity to study at UCLA. If it were not for you, my life would have been very different now. My life in LA was full of experience and I want to thank you for that.

Professor Iwasaki, you were always there for me, not only as my academic mentor, but also as my emotional support. I was deeply moved by your kindness and wisdom every time I talked to you. You have taught me and shown me what a true teacher should be. I hope to become a teacher like you someday.

Professor Tao, I was always amazed by your profound knowledge and insight. You made me feel very lucky to be a part of ALC linguistics program. After learning from you and researching with you, I have become a better researcher than I have ever been. Thank you so much for being such a great academic role model.

Professor Jun, I cannot express how grateful I am to you. Without your detailed guidance, I would never have been able to go though such a difficult time. Thanks to your patient teachings, I was able to see everything clear and truly understand how things work. You always

made me feel confident about myself even when I was the lowest. I sincerely appreciate your thoughtful care.

I have to acknowledge my loving Korean SSNs, Jane Choi, Jae-eun Im, Seonkyung Jeon, Jae-eun Yoon, and Hee Ju. You all gave me so much support more than I deserved. I will never forget your unconditional love.

Also, I am grateful to my loving colleagues Jaehyun Jo, Dayoung Jeong, Ariel Chan, Yan Zhou, Liz Carter, Yi Ren, and Cheer Wu. Thank you guys for making my academic life more exciting.

Finally, I want to thank my family for their endless love and support.

VITA

2009 - 2014	B.A. in Korean Language
	Kyung Hee University, Korea
2014 - 2016	M.A. in Korean Education
	Kyung Hee University, Korea
2013	Visiting student instructor
	Brigham Young University, Provo, Utah, USA
2016 - 2021	Graduate Student Researcher
	Department of Asian Languages and Cultures
	University of California, Los Angeles, California, USA
2017 - 2021	Teaching Assistant/Associate/Fellow
	Department of Asian Languages and Cultures
	University of California, Los Angeles, California, USA
2018 - 2022	Instructor of the Intermediate Korean Language Course
	Korean Cultural Center, Los Angeles, California, USA

PUBLICATIONS AND PRESENTATIONS

Won, Hyobin (2019). *Prosodic Features of Backchannel Responses in Korean Conversation,* The 27th Japanese/Korean Linguistics Conference , Seoul, Korea.

Won, Hyobin (2019). *Who did that to you?: Conversational Tactics for Solicitation and Evasion in Referent Identification,* The 27th Japanese/Korean Linguistics Conference (poster presentation), Seoul, Korea.

Won, Hyobin and Zhou, Yan (2019). *Emerging negative stance in Korean and Mandarin Chinese interaction,* The 16th International Pragmatics Association (IPrA) Conference, The Hong Kong Polytechnic University, Hong Kong.

Won, Hyobin (2016). *Various Aspects of Case Marker Use in L2 Korean Writings*, The 38th Korean Semantics Conference, Seoul, Korea.

Won, Hyobin. (2015). *A study on the Case Marker Acquisition of Korean L2 Learners through the view of Linguistic Typology,* Korean Journal of Studies in Linguistics 38, 357-382.

CHAPTER 1: INTRODUCTION

1.1. Goals of the study

This dissertation explores the manifestation of listenership in Korean conversation. The term listenership refers to "the role of listener" and presupposes that listeners are not merely passive hearers, but active participants in conversation. The previous studies across languages have proven that listeners cooperate in developing sequences and their role is as influential as that of speakers. It has also been discovered that listeners display listenership through verbal and nonverbal signals called reactive response. The findings have shed new light on the listener's role in conversation and contributed to the research on human interactions. However, despite the massive volume of ongoing research, the existing studies are still missing some important points. Researchers have tended to simplify the functions of reactive responses and labeled them with a few specific types. The multiple aspects that a single response can manifest have not been thoroughly taken into account. Also, reactive responses in Korean conversation have been scarcely studied, especially from conversational analytic view.

This dissertation is motivated by such reasons and aims to demonstrate conversational meanings of Korean reactive responses. Korean listeners deliver their intentions and stances towards the prior turn or the entire telling by fine-tuning the prosodic and multimodal features of reactive responses. The study investigates the various aspects displayed through the responses, particularly focusing on listeners' affective and epistemic positioning in the course of telling. The research questions of the dissertation are presented below.

1. What are the most frequently used forms of reactive responses in Korean conversation?

2. How is listeners' affective stance displayed through reactive responses?

3. How is listeners' epistemic stance displayed through reactive responses?

4. How does the token type of verbal tokens affect the conversational meaning of the response?

5. How do the prosodic features of a verbal token affect the conversational meaning of the response?

6. How do the multimodal features affect the conversational meaning of the response?

1.2. Data

The conversation data for this dissertation were collected from 3 male and 3 female Korean native speakers who were UCLA graduate students. Their ages range from early 20s to early 30s and their birth places are Seoul (the capital city of Korea) and Kyungsang (southern part of Korea). One participant (SW) was born in Arizona, USA, but he grew up in Seoul and is able to speak Korean fluently. While all of the participant are considered to speak the standard accent, two participants (CY and HB) who were born in Kyungsang area have a mild southern accent. However both of the participants' accents were very minor and did not show a significant difference in their production of reactive responses compared to the other participants. The summarization of the detailed participant information is presented in Table 1.

<Table 1. Participant Information>

	CY	HB	GA	HS	PT	SW
Gender	M	F	F	F	M	M
Age	31	30	29	28	26	24
Birth place	Kyungsang	Kyungsang	Seoul	Seoul	Seoul	Seoul
Accent	Mild southern	Mild southern	Standard	Standard	Standard	Standard

Each participant was paired with another participant forming 15 dyads and had 2 sets of conversation with each person. The participants were not given any specific topic, but allowed to have free conversations. Each set of conversation lasts for 20 minutes. A total of 30 sets of face-to-face conversation (approximately 600 minutes) were collected and transcribed.

1.3. Methodology

This study explores the conversational meaning conveyed through reactive responses, particularly the stances displayed by the prosodic features of verbal tokens as well as the non-verbal behaviors. First, for the prosodic analysis, I used the software Praat and K-ToBI convention (Jun, 2000; Park, 2003). The prosodic features of each response token, such as pitch range, amplitude, and intonation contour, were measured with Praat. Then, the response tokens were transcribed with K-ToBI convention to identify their pragmatic meanings delivered through the boundary tones. Finally, the main analytic tool I used for the detailed analysis is Conversation Analysis (CA) which aims to study social interaction by understanding talk through turn-taking, preference organization, and sequence organization (Sidnell, 2011). Through CA

approach, I closely examined the fragments and demonstrated affective and epistemic stance displayed by the reactive responses.

1.4. Organization of the Study

This dissertation is organized as follows. Chapter 2 will explain the theoretical background of the dissertation. It introduces the definition of listening and listenership and reviews reactive response in the previous studies. Then, the notion of stance, the main focus of this study, will be explained in detail. Also, storytelling and K-ToBI will be introduced as the specific type of conversation dealt with in this study and the model for the prosodic analysis, respectively. In Chapter 3, I will provide the theoretical framework of this dissertation by demonstrating how reactive responses should be analyzed considering the multiple layers they display. Chapter 4 will show the analysis of yes/nod type responses and oh/really type responses will be discussed in Chapter 5. Chapter 6 will summarize the findings and suggest some ideas for future research.

CHAPTER 2: THEORETICAL BACKGROUND

2.1. Listening and Listenership

In conversations, when one person speaks, the other person is - or people are - required to listen. The act of listening is a requirement of building a conversation. Then, what is listening? Is it the same activity as hearing? According to Merriam-Webster dictionary, the word “listen” is defined as “to pay attention to someone or something in order to hear what is being said, sung, played, etc.”, whereas “hear” merely means “to perceive or become aware of by the ear”. Through the definitions it can be inferred that when people listen, they do not just perceive the sound, but they make effort to understand the meaning being relayed. This distinction was also made by conversation analysts in their research on “listenership”, namely “the role of listener”. McCarthy (2003) describes listenership as “the active involvement with the floor-holding speaker that displays more than just hearership referring to a recipient’s neutral status of information reception”. O’Keeffe et al. (2007) define listenership in a similar way that it is “the active, responsive role that listeners have in conversation”. In sum, contrary to “hearership”, “listenership” requires “active engagement” in conversation. Then, how do people “actively engage” in conversation as listeners? Studies have been discovered that listeners cooperate in the construction of conversation by sending verbal and nonverbal signals to the floor-holding speaker. This study closely examines how Korean native speakers deploy these signals as an active listener rather than a passive hearer.

2.2. Reactive Response

"Good listenership" is generated by a listener's contribution to interpersonal relationships and this can be achieved by employment of "reactive tokens" (McCarthy, 2003; Tsuchiya, 2013). Reactive token refers to responses produced by a listener during the primary speaker's turn in conversation (Yngve, 1970; Duncan and Fiske, 1977). It has been introduced in the different terminologies, such as backchannel (Yngve, 1970), minimal response (Fishman, 1983), receipt token (Heritage, 1984), reactive token (Clancy et al., 1996) etc. Backchannel, as its name illustrates, was treated to play a secondary role by recipients merely following the primary speaker's lead. However as later studies started to shed more light on the importance of the collaborative role of recipients, their responses are now considered an essential component of steering conversation. Since then, how listeners manage reactive responses to cooperate with their co-participants has been an intriguing topic among conversation analysts. In this section, I will provide a comprehensive overview of the previous studies on listener's response. Note that, I chose to use the term "reactive response" adopting the definition of reactive token by Clancy et al. (1996), "a short utterance produced by an interlocutor who is playing a listener's role during the other interlocutor's speakership", while using the word "response" instead of "token" not to confine the concept to "a short utterance" but to also capture the multimodal aspects such as facial expressions and gestures that co-occur with the tokens.

2.2.1. Terminologies and Classifications

Scholars have coined new terminologies in efforts to identify the discourse functions of reactive responses. The most famous term “continuer” was suggested by Schegloff (1982). He defines continuer as non-lexical or lexical items by which the listener exhibits his understanding when passing the chance to take the floor. He focuses on the response’s specific trait of “not taking the floor” and thereby “encouraging the current floor-holder to continue his talk”. The concept of “acknowledgement” was first canvassed by Jefferson (1984). When a recipient of an ongoing talk neither shifts the topic nor talks on the topic in progress, he or she produces an acknowledgement token, recurrently. In this case, a talk is neither disattended nor taken up, but it is “acknowledged” by the recipient. The term “newsmark” was also introduced by Jefferson (1881) and it refers to the objects that treat a prior turn’s talk as “news”. This notion was furthered by Heritage (1984) who distinguishes “newsmark” from “information/news receipt” which merely marks the “receipt” of new information. Gardener (1997) coined the term “newsmarker”, including both “newsmark” and “news receipt”, referring to responses that mark the prior turn as “newsworthy” in some way. He argues that some lexical tokens such as “really” and “right” and the change-of-state token “oh” (Heritage, 1984) can display that the turn being responded to is new to the recipient.

Using the newly coined terminologies and the concepts of them, researchers have been trying to categorize reactive responses based on their discourse functions, forms and several other aspects. Maynard (1990) proposes six functions based on its discourse functions and sequential structures or functions such as minor addition, correction or request of information. Gardner (2001) introduces eight types of reactive responses classified by functions and forms

even including non-verbal vocalizations and kinesic actions. However in his study in 2002, the form-based types are excluded and two discourse functions are added (terminator and closing). O'Keeffe and Adolph (2008) distinguish discourse functions and conversation functions and suggest four types for each group. Although different terminologies are deployed, many of the functions overlap across the studies. Below is the summary of the different categorizations of reactive responses suggested by these researchers.

<Table 2. Categorizations of reactive responses>

Maynard (1990)	Gardner (2001)	Gardner (2002)	O'Keeffe and Adolph (2008)
(1) Continuer (2) Display of understanding (3) Support towards the speaker's judgement (4) Agreement (5) Strong emotional response (6) Minor addition, correction or request of information	(1) Continuers (2) Acknowledgements (3) Newsmarkers (4) Change of activity (5) Assessments (6) Brief questions (7) Collaborative completions (8) Non-verbal vocalizations and kinesic actions	(1) Continuer (2) Acknowledgement (3) Terminator (4) Newsmarker (5) Closing	<Discourse Functions> (1) Continuer (2) Convergence token (3) Engaged response token (4) Information receipt token <Conversation Functions> (1) Floor taking (2) Floor seeker (3) Listenership (4) Floor giving

The criteria used for the categorizations, however, lack consistency in the attempt to present the manifold aspects of the different conceptual layers within a single table. For example, in Gardner's (2001) list, while (1) "continuer" is a term denoting a discourse function, that is, "the listener's passing the opportunities to take the floor", (8) "non-verbal vocalizations and kinesic actions" only reveals the formation of response. The inconsistency in listing is not the only

problem, but also the terminologies themselves are vague and disputable. Why can't continuers "acknowledge" the prior turn at the same time? Can't newsmakers let the speaker continue? What do you "understand" when you "display" understanding. Haven't you also "understood" something when you let the speaker continue, acknowledge something, or mark something newsworthy? To solve this problems, I will suggest a new perspective on identifying reactive responses by considering the diverse aspects they display and treating them as different layers that can be activated simultaneously.

2.2.2. Form

What tokens are considered reactive response is directly related to the definition of reactive response. The classification based on the form of reactive response, therefore, has been an important argument. In English studies, Maynard (1990), Gardner (2001, 2002) and O'Keeffe and Adolf (2008) have discovered that reactive responses can be produced not only in verbal forms but also in physical forms, such as head nods. Drawing from the work of these researchers, Tsuchiya (2013) summarizes the classification of English response tokens based on their forms. The two broad categories are vocal tokens and visual tokens. There are five components under the vocal category comprising minimal response, non-minimal response, clustering of response tokens, laughter and pause. The visual category includes six gestures such as head nods, head shakes, head turnings, hand gestures, foot movements, and self comfort. He suggests example tokens belonging to each vocal token type and a description of each visual token type as tabulated in Table 3.

<Table 3. Classification of English reactive tokens by Tsuchiya (2013)>

Vocal response tokens	(1) Minimal response	*Mm, Uh-uh, Yeah, Right*
	(2) Non-minimal response	*Lovely, definitely, I see*
	(3) Clustering of response tokens	*Mm mm, yeah right*
	(4) Laughter	Chuckles and laughs
	(5) Pause	Silent pause
Visual response tokens	(1) Head nods	Any vertical head movement
	(2) Head shakes	Any horizontal head movement
	(3) Head turnings	Head moving towards speaker
	(4) Hand gestures	Any hand movement
	(5) Foot movements	Crossed legs
	(6) Self comfort	Crossed arms

Tsuchiya's (2013) classification is useful in that it shows most of the sub-categories suggested by other previous studies including the various physical response types. It is interesting that it treats a single gesture as a single "token". In my study, however, "a token" only refers to "a short utterance" and the multimodal aspects will be considered the other elements composing "a response".

Young and Lee (2004) explore the forms of Korean reactive response. They suggest a list of Korean reactive response tokens with English translation as shown in Table 4.

<Table 4. Classification of Korean reactive tokens by Young and Lee (2004)>

Korean Reactive token	**English gloss**
yey/ey/ney	"uh huh", "yeah", "okay"
a/e	"uh huh", "yeah", "okay"
mm/ung	"mm hm"

.hh	“wow"
eme	“oh”
kulay(yo)	“oh yeah?”
cengmal(yo)	“really?”
maca(yo)	“right”

The list includes monosyllabic responses such as “yey”, “e”, and, “mm” and short utterances such as “kulay(yo) (really?)” and “maca(yo) (right)”. With Tsuchiya’s (2013) terminologies, this list contains minimal responses and non-minimal responses. In my study, the target token forms also include some monosyllabic responses (minimal responses) and short utterances (non-minimal responses) and show a highly similar inventory as that of Young and Lee (2004). In the analysis chapter, I will first group them by similar meanings and then suggest the most frequently occurring tokens in my data.

2.2.3. Prosody

Prosody is defined as “the patterns of stress and intonation in a language”. In the research of reactive responses, prosody is a pivotal feature for identifying verbal tokens. Müller (1996) claims that prosody can change the impact of reactive tokens. While prosodically weak tokens only acknowledge a recognition of the emergent speech object and thus remaining limited to *de dicto*[1] recognition, prosodically salient tokens can signal *de re*[2] recognition of what is being said

[1] Relating to the form of an assertion or expression itself, rather than any property of a thing it refers to (Oxford English dictionary, 2020).

[2] Relating to the properties of things mentioned in an assertion or expression, rather than to the assertion or expression itself (Oxford English dictionary, 2020).

and assume a more differentiated stance. Gardner (1997) illustrates how intonation contours can can change what reactive tokens display in English. According to the study, a falling intonation signals that the producer has clearly understood the content and has nothing to add to the topic, while a fall-rising contour shows that it requires further work from the speaker. Thompson et al. (2015) demonstrate the role of intonation on responses in developing further sequences in English. According to them, a rising intonation tends to express doubt about the informing and makes a further turn from the informer relevant, whereas a falling intonation accepts the validity of the informing and declines to pursue further expansion. They also focus on "pitch upgrade/ downgrade" of response tokens in relation to their prior turn and claim that it can add an "affective lamination". In my study, prosodic features of verbal tokens will be deeply delved into along with other multimodal factors to investigate the affective and epistemic stance that reactive responses display.

2.3. Stance

The conceptual foundation of this study lies in the claim that listeners display their stance towards the immediate preceding turn or the telling provided up to that point. In this section, I will discuss the types of stance examined in this study; affective and epistemic stance. In addition to that, the related concepts "alignment and affiliation" will also be explained.

2.3.1. Affective Stance

According to Du Bois (2007), affective stance is displayed when the speakers position themselves along an affective scale by indexing specific aspects of their feelings with affective predicates such as glad or amazed. Also, when the predicates are transitive (like, don't like, love, hate and etc.), they regularly specify the object the speaker is affectively orienting to. Therefore affective stance often presents a particular emotional status with an orientation to a specific object of the speaker's stance.

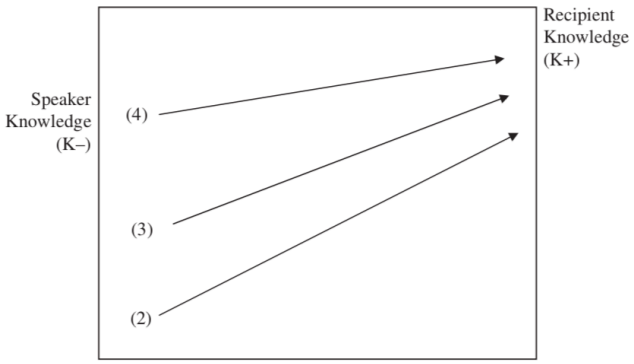
However, emotional status is not considered the same concept as affective stance, particularly in CA where the focus is on how emotions are displayed and treated in interactions rather than the actual emotional messages and intentions of the individuals (Peräkylä & Ruusuvuori, 2006). In his studies on facial expression, Fridlund (1991) argues that facial expressions are vehicles of either real or imaginary interaction designed for social uses. Goffman (1961) also says that the participants adjust their emotional states in accordance with the demands of the situation. Therefore, affective stance can be manipulated to appear to have certain emotions at the moment by the participants.

Affective stance is particularly important in this study that focuses on the participants' reactive responses through which emotions are expressed towards the preceding turn or the entire telling. Emotions in reactions were observed by Sacks (1972) in the discussion of expressions of sorrow and joy in conversation. He finds that these emotions seem to be situated in responses to news announcement rather than in first things in sequence. In this study, the delicate emotions in reactive responses will be investigated by analyzing their prosodic features and accompanying gestures.

2.3.2. Epistemic Stance

The term “epistemic” is related to the knowledge territories of the participants in the conversation. According to Kamio (1997), the participants have their own territories of information, and they can share the same element of knowledge to different degrees. Heritage (2012) draws a distinction between epistemic status and epistemic stance. While epistemic status means the relative access to some domain of knowledge that the participants have, epistemic stance refers to how speakers position themselves to appear to be knowledgeable to a certain extent through the design of turns at talk. Therefore, the realized epistemic stance is not necessarily congruent with the actual epistemic status. The participants can dissemble their stance to look more, or less, knowledgeable than they are.

He also considers that the participants occupy different positions on an epistemic gradient and the gradient itself may vary in slope from shallow to deep.



<Figure 1. Epistemic Stance represented in terms of Epistemic Gradient by Heritage (2012)>

Epistemic gradient can answer two different questions: (1) who is more knowledgeable between A and B?; (2) how knowledgeable is each participant positioning him/herself to be? The first

question can be simply answered by the comparison between the two ends of the slope and represented as [K+] (more knowledgeable) and [K-] (less knowledgeable). The second question is, however, more complicated to answer. In figure 1, three actions from [-K] speaker are stratified by their grammatical realizations; (2) Are you married?; (3) You're married, aren't you?; (4) You're married. While (2) proposes that the speaker has a very low level of knowledge of the information, (3) and (4) show that he or she already has some knowledge and seeks to confirm it. Through managing the question from (2) to (4), the speaker can express his or her increasing certainty or knowledge of the information as shown in the different gradients. As seen in this diagram, epistemic stance can be considered on a linear range of values. Each participant's positioning can vary along the range from less knowledgeable to more knowledgeable and a stance always falls onto a certain point on the spectrum. In this study, epistemic stances displayed in reactive responses will be closely examined as one of the aspects to identify each response.

2.3.3. Alignment and Affiliation

The notion of alignment and affiliation has often been used to express the participants' emotional support to each other managed through various conversational tactics in pursuit of building solidarity. While the two terms had been interchangeably used, Stivers (2008) distinguishes these two and considered each of them to display a distinct function. She differentiates nods from vocal continuers in that nods are produced when the recipient is offered access to the teller's stance, whereas continuers can only acknowledge that the telling is in

progress without any access towards the event being reported. Here, what she means by "access" is that the recipients are provided with the means to understand "what it was like to experience the event" being reported through the eyes of the teller. The tellers can provide access to their stance through story prefaces, information packaging through grammatical and lexical devices, prosody, and the context of the telling. When the recipients are given access from the teller, according to her, they are more likely to nod. While the recipients are only able to structurally align (or disalign) with the telling activity via vocal continuers, they can socially affiliate (or disaffiliate) with the tellers by endorsing their perspectives through nodding.

This distinction gave this study an insight for the basic functions of reactive responses. In the aspects of the functional categorization in this study, Stivers's alignment is equivalent to "listener's understanding of the floor situation" and her affiliation is related to "listener's affective stance". The detailed explanation will be provided in the next chapter.

2.4. Storytelling

The interest of this study is in the listener's role in conversation activated through reactive responses. To concentrate on this, I chose a specific type of conversation to investigate; storytelling. Storytelling is entitled to a distinctive type of activity in Conversation Analytic research. Unlike turn-by-turn talks where each participant produces one turn-constructional unit (TCU) at a time at its possible completion point, in storytellings the main teller leads the floor with extended turns-at-talk while the recipient passes multiple opportunities to take turns. It is achieved by the recipient who recognizes it as a story at the beginning of the telling, thereby the

teller can continue to talk (Sacks, 1974; Mandelbaum, 2013). As the telling being the main speaker and the recipient being the main listener, the role of listener can be magnified, and therefore, reactive responses can be captured better in storytellings.

The distinctive characteristics of storytelling can be more noticeable when compared with other types of conversation. Stivers (2012) compares the organization of adjacency pairs and storytellings. While adjacency pairs are composed of first and second-pair part actions, storytellings are established around conveying stances. In storytellings, the teller usually begins the story with a preface such as "did you hear what happened to Debbie?" in order to secure the interlocutor's willingness to be a story recipient. The teller also designs the lexical, phonetical, and grammatical features of turns to indicate when the story will be complete and what kind of stance is preferred. Storytelling is also differentiated in terms of the recipient's incremental uptake. While uptake is required at completion points in turn-by-turn talks, the recipient of storytelling provides response tokens throughout the telling to express their evaluative stance. In storytellings, the separate roles of the speaker and the listener are more clearly presented.

For these reasons, I take storytelling as the optimal type of conversation for investigating reactive responses and displayed stances in the course of telling. How the speaker and the listener negotiate their stance to affiliate and, sometimes disaffiliate, with each other will be discussed in the next chapter.

2.5. K-ToBI and Boundary Tones

In this dissertation, prosody is used as one of the major tools to determine the discourse functions and stances of reactive responses. Particularly, it investigates the intonational tones occurring at the end of a phrase, called boundary tones that are known to convey various pragmatic meanings (Jun, 1990; Park, 2003). As a transcription tool of Korean intonation, I will adopt K-ToBI (Korean-TOnes and Break Indices) conventions (Jun, 2000).

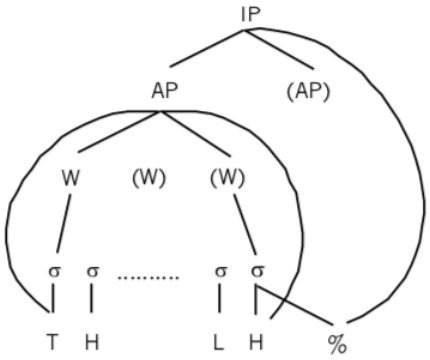


Figure 1. Intonational Structure of Seoul Korean
IP: Intonation Phrase AP: Accentual Phrase
w: phonological word σ: syllable
T= H, when the syllable initial segment is aspirated/tense, otherwise, T= L
%: Intonation phrase boundary tone

<Figure 2. Intonational Structural of Seoul Korean>

K-ToBI, developed by Jun (2000, 2005) based on the phonological model of Korean intonation (Jun 1993, 1996, 1998), is a prosodic transcription convention for standard (Seoul) Korean. As shown in Figure 2, the intonational structure of Korean has two prosodic units, AP (Accentual Phrase) and IP (Intonational Phrase). An AP is smaller than an IP but larger than a phonological word (w). An AP has a tonal pattern of /HH... LH/ or /LH...LH/. That is its left edge is marked by a high tone (HH) or a rising tone (LH) and its right edge is marked by a rising tone (LH). When

an AP has four or more syllables, all four tones are realized, but when an AP has fewer than four syllables, only the AP-initial and AP-final tone are realized. That is, the AP-medial tones, H or L or both, can be deleted. An IP can have one or more APs and is marked by a boundary tone (%). When an AP is ithe last AP of an IP, the AP-final boundary tone is overridden by the IP-final boundary tone (%). The IP boundary tones can be H%, L%, LH%, HL%, HLH%, LHL% and convey various pragmatic meanings and information about the sentence type.

Park (2003) investigates the meaning of Korean boundary tones. According to her, each element of a boundary tone has a distinct meaning; a low tone signifies high degree of awareness or certainty toward the information, whereas a high tone denotes low degree of awareness or certainty and surprise based on the new information. In a combination of such elements, each boundary tone denotes a pragmatic meaning. The meaning of the boundary tones suggested by Park (2003) are listed below.

(i) H% tone means speaker's immediate surprise based on the new information,

(ii) LH% tone highlights the shift from the speaker's previous belief to the current surprise,

(iii) HLH% tone is the same as LH% but an emphasis is on surprise (H),

(iv) L% tone has the focus on the realization alone,

(v) HL% tone reveals surprise followed by realization,

(vi) LHL% tone is the same as HL% tone but the emphasis is more on realization (L).

Park's (2003) study is meaningful in that it has discovered the fact that each of the multi-tonal units independently displays its own meaning and what the meanings are. These findings have

given an inspiration and become the foundation of my dissertation. In my study, the conversational functions and stances displayed by the reactive responses are closely examined in relation with their intonation contour and boundary tones. In Chapter 3, the theoretical frameworks and technical details of the analysis will be introduced.

CHAPTER 3: THEORETICAL FRAMEWORK

3.1. Listenership to Speakership

The term “listenership” in conversation has been considered the antithesis of “speakership” along with other terms such as hearership, recipientship and recipiency. Jefferson (1984) discusses recipiency and speakership in her research on deployment of the acknowledge tokens. She claims that there are distinctions among the various tokens, for example, while “*mm hm*” shows “passive recipiency” by which she means “its user is proposing that his co-participant is still in the midst of some course of talk, and shall go on talking”, “*yeah*” can exhibit a preparedness to shift from recipiency to speakership. “Speakership” refers to, based on her description, “its user is proposing that he becomes the main speaker by taking the floor”. However, the recipient’s intention is oftentimes not clear to identify. For instance, if the recipient produces a “*yeah*” token in order to express his excitement and the highly engaging attitude towards the telling, is this person considered to be assuming speakership or merely showing an intense version of listenership?

Xu (2014) shows how Mandarin speakers employ reactive tokens in order to manifest recipiency. In her study, the different levels of recipiency were presented as a cline in following order; silent - passive - neutral - active - affiliative. According to her, while the absence of reactive tokens is interpreted as disengaged and indifferent, semantically empty vocalizations such as “*mh*", “*uh huh*”, and “*mhm*" show passive recipiency avoiding interference and encouraging the floor-holding speaker. Recipiency can be neutrally expressed through lexical

tokens such as ‘‘*yes*’’, ‘’*right*’’, and ‘’*okay*’’ and it is more actively conveyed when these tokens are repeatedly used. Finally, laughter tokens can show affiliation or empathy in terms of the sense of “togetherness” which, he claims, is the highest level of recipiency. Xu (2014)’s work is meaningful in that she mobilizes the concept of “continuum” to show the different types of displaying recipiency. However, the different levels are still labeled separately, as if there are clear cuts between them. As she briefly mentions in the conclusion, there is no one-to-one relationship between a type of reactive token and the status of recipiency, particularly when other facts such as prosody can add extra layers of meaning or function to it.

It is the fundamental claim in this study that displaying listenership is a highly sophisticated interactional work manipulated by a number of different devices. For this reason, it is not readily possible to label a contingency as a single function or a single type. Since any response can express multiple aspects through the choice of lexical, prosodic, and gestural information, it is important to take into account multiple dimensions. In the past research, the various aspects of reactive responses have been lumped into rather simple functional categories lacking clear distinctions.

To reorganize this, I first attempt to establish the basic qualifications for reactive response that are required for them to be entitled to “reactive responses” by my definition. I claim that a “reactive response” through which listenership is assumed must display three essential aspects; listener’s (1) receipt of the preceding turn; (2) understanding of the floor situation; and (3) unwillingness to take the floor. Then, I will investigate the qualified “reactive responses” with respect to two types of stances; affective stance and epistemic stance. I argue that the

conversational functions of reactive responses will be discovered through examining the stances they display in storytellings. The detailed explanation will be continued in the next section.

Let me clarify the terms I use in this dissertation. For the role of the participants, I use the term "listener" to refer to the participant who is assuming listenership (the role of listener) by producing reactive responses and "speaker" to refer to his or her interlocutor who is assuming speakership (the role of speaker) by holding the floor of current telling. For the definition of "floor", I borrow the concept of "unit-floor", introduced by Iwasaki (1997), which is defined as "a unit of conversation which has a coherent speech activity such as topic development". Therefore, when a participant is creating a story consisting of several turns and the story is considered one united speech activity developing the same topic, the participant is considered a "speaker" who is holding a "floor".

The assignment of the roles as the speaker and the listener keeps changing as a conversation evolves. The participants dynamically take turns and sometimes fight for the floor. This dissertation however deals with the fragments where the roles of the speaker and the listener are set in stone and the listener plays his or her role enthusiastically without attempting to become the speaker.

3.2. Essential Aspects

In this section, I suggest three essential aspects displayed by reactive responses; l listener's (1) receipt of the preceding turn; (2) understanding of the floor situation; and (3)

unwillingness to take the floor. The three aspects are the qualifications for responses to be entitled to “reactive responses” in this study.

Aspect 1. Listener’s receipt of the preceding turn

The first aspect displayed by reactive responses informs the speaker that the preceding turn was clearly heard and has been registered as a valid turn. Therefore nothing is problematic and no repair is needed in terms of auditory perception. If the listener indicates something is off or initiates a repair, this activity is not considered a reactive response.

Aspect 2. Listener’s understanding of the floor situation

The second aspect expresses the listener’s awareness that his or her interlocutor has the current floor and it is in progress. He or she also understands that the newly registered turn is not just a single turn, but a unit that constitutes the expanded talk. This concept is equivalent to Stivers’s (2008) *structural alignment* by which the listener can support the progress of the current speaker’s telling. The misunderstanding of the floor situation will lead to a collapse of the current role allocation between the listener and the speaker, and the listener’s activity derived form such a misunderstanding is not considered a reactive response.

Aspect 3. Listener's unwillingness to take the floor

> The third aspect indicates the listener's unwillingness to take the current floor and remain as a listener. This aspect is should be differentiated from the listener's understanding of the floor situation. For instance, even if the listener is aware of the person who currently holds the floor, he or she can still attempt to take it away from the speaker. If the listener is trying to become a speaker by taking the floor, this activity is not considered a reactive response[3].

The famous terminology "continuer" may seem to be related to all the three aspects listed here. The literal meaning of the term signifies that the speaker is allowed or encouraged to "continue" to talk. The speaker's right to continue is granted by the listener. In other words, the speaker's continuation can happen as a consequence of the listener's opting out of taking the floor. For the listener to make this choice, he or she has to (1) successfully register the prior turn without any hearing problems, (2) understanding who has the floor and that it is in progress, and (3) be unwilling to become a speaker. If the definition of the term "continuer" implies all the steps listed here, all the reactive responses are fundamentally continuers by my definition.

[3] As for "*yeah*" from Jefferson's (1984) study, although it exhibits a "preparedness" to shift from recipiency to speakership because the listener attempts to take the floor in the following turn, the token itself expresses all of the three aspects until he or she actually produces the next turn. Therefore, I treat it as a reactive response which tends to be followed by the listener's attempt of floor taking.

3.3. Stances

While the three aspects discussed above are the essential qualifications of reactive responses, the stances are unique features that all responses display differently. Therefore, the stances are determinants of the distinctive qualities of each reactive response. The study mainly focuses on affective stance and epistemic stance to identify the discourse functions of reactive responses.

(1) Affective stance

The listener can exhibit their feelings towards the newly received information. Stivers (2008), in her explanation of affiliation, defines stance as "the teller's affective treatment of the events he or she is describing". Affective stance can be managed to appear to have certain emotions in accordance with the demands of the situation. These emotions are conveyed through prosodic features and non-verbal behaviors such as facial expressions and gestures. In this study, the delicate emotions of reactive responses will be investigated.

(2) Epistemic stance

The listener can express his or her understanding of the newly introduced information or knowledge. According to Heritage (2012), epistemic stance refers to how speakers position themselves to appear to be knowledgeable to a certain degree through the design of turns at talk and it may or may not be congruent with

the actual epistemic status, the relative access to some domain of knowledge that the participants have. I will examine how the listener's epistemic stance is presented through the management of reactive responses. Therefore, the focus is not on their actual access to the knowledge territories, but on their positioning as a listener which might have been masqueraded to appear more, or less, knowledgeable than they actually are.

In this chapter, I have explained the three essential aspects displayed by reactive responses; (1) receipt of the preceding turn; (2) understanding of the floor situation; and (3) unwillingness to take the floor and the two types of stances that reactive responses display; affective stance and epistemic stance. In the following chapters, the analysis of "yes/nod" type and "oh/really" type reactive responses will be investigated with respect to the two stances. For "yes/nod type" responses, "*ung*", "*e*", "*mm*", and multiple saying "*e*" and "*mm*" tokens, and for "oh/really" type responses, "*a*", "*cincca*", and "*kulay*" tokens and *a*-prefaced structures "*a cincca*" and "*a kulay*" will be discussed. The response tokens are selected based on the frequency and classified by their conversational functions. Although these tokens and token types do not include all the Korean reactive responses, this analysis will give a general idea on how Korean listeners display listenership.

CHAPTER 4: YES / NOD TYPE RESPONSES

In this chapter, "yes" and nod type reactive responses will be discussed. The word "yes" has been studied in various languages with respect to its discourse functions (Jefferson 1984; Drummond and Hopper 1993; Zimmerman 1993; Angles et al. 2000; Aoki, 2008). "*Yes*" in Korean can be presented in several different forms including the plain or intimate forms, "*ung*" and "*e*", the polite forms,"*ney*" and "*yey*", and even the non-lexical vocalization "*mm*" in a broad sense. The previous studies on Korean "*yes*" tokens have discovered their interactional functions in conversations. Kim (1993) found that when "*yes*" tokens are produced with a rising intonation, they function as other-initiated repair signaling a trouble processing the prior turn. In their study on confirmation sequences in language proficiency of non-natives speakers, Kim and Suh (1998) use the term "receipt marker" referring to "*yes*" tokens marking the listener's receipt of the preceding turn. Kim (1999) argues that Korean "*yes*" tokens function similar to English continuers, however they tend to appear at intra-turn unit boundaries, before the turn is complete, while English continuers occur at "transition relevance place (TPR)". Pyun and Yoon (2022) discovered fifteen discourse-pragmatic functions of Korean "*yes*" tokens; (1) affirmative answer, (2) confirmation, (3) acceptance, (4) agreement, (5) answer to summons, (6) acknowledgement, (7) change-of-state, (8) change-of-activity, (9) response solicitation, (10) reinforcement, (11) other initiation of repair, (12) closing of phone call, (13) continuer, (14) proposal to discontinue the on-going action for the sake of a larger course of action, and (15) arguably hesitation marker. They provide a highly specified classification of the discourse functions of "*yes*" tokens that can

occur in possibly all kinds of conversation. However, the explanation of each function provided in the study is rather too simple to generalize.

In this chapter, the most commonly occurring "yes" type reactive response tokens in Korean conversation will be discussed; "*ung*", "*e*", "*mm*", and multiple saying "*e*" and "*mm*". The goal of this chapter is to demonstrate how the listeners manage "yes" type responses to produce sequentially and socially preferred responses while displaying their own stances in storytellings. Meanwhile, since the focus of my study is on "reactive responses", responses in "question-answer" sequences, such as information-seeking or action-requesting sequences, will be excluded from the discussion.

Although it is not a verbal form, "nod" will also be examined in this chapter. Nod has been known as one of the listener's reactions indexing the speaker's turn is still in progress (Goodwin, 1980; Schegloff, 1982; Jefferson, 1984; Goodwin, 1986). The studies on nod have explored its unique discourse functions particularly compared to verbal tokens. Stivers (2008), in her famous study on nod, discovered the difference between nods and verbal continuers. According to her, through nods, a recipient can claim to have access to the event being reported, thereby affiliates with the teller's stance, while with vocal continuers, he or she can only structurally align with the telling. Voutilainena et al. (2019) also found the difference between vocal continuers and nods. In their findings, vocal continuers indicate the perception of empathy, whereas the impact of nods remains weak. Berger and Rae (2012) claim that "people may choose to respond non-vocally" in order to "fulfill the obligation to respond in a particular time and place but can evade some of the constraints that impinge on vocal responses".

In my data, while nods frequently co-occur with verbal tokens, or follow them, they also appear independently. I consider the independent nod(s) a separate type of response and will discuss it in this chapter. Also I divide the independent nod(s) into two types depending on their placements; during the ongoing turn and at the turn completion points. The former, mid-turn nods, and the latter, turn-completion nods, will be illustrated, respectively. The analysis of nods will be presented in section 4.2. along with the vocalization "*mm*".

4.1. Lexical token "*ung*" and "*e*"

The lexical tokens "*ung*" and "*e*" both of whose meanings are non-polite "yes" are the two most frequently used response tokens. A distinction between "*ung*" and "*e*" was once made by Oh and Park (2017) in their study of Korean acknowledge tokens "*ung*" and "*e*". They claim that even though the two tokens seem to be used interchangeably in conversation, "*ung*" functions as a typical continuer treating the preceding turn as an in-progress multi-unit, whereas "*e*" is used to provide the listener's appreciation which has been made relevant to the turn-so-far. Particularly, "*e*" appears after the speaker's assessment or noteworthy and emphasized telling.

Stivers' (2008) findings on the different functions of nods and vocal continuer should also be revisited here although the token types are different. According to her, when the recipients are offered access to the teller's stance they tend to nod showing their endorsement to the teller's perspective. On the other hand, the recipients produce vocal continuers when they are not given such access and these simply support the structural asymmetry of the storytelling in progress.

The findings of Oh and Park (2017) and Stivers (2008) seem relevant to the two tokens, "*ung*" and "*e*", found in my data.

In fragment 1, the speaker HB gradually builds a foundation for a question that she eventually wants to ask the listener HS. Through the seven consecutive turns, HB explains why she is curious about HS's earlier statement that there was a husband in her future plan.

<Fragment 1: Husband>

```
1 HB:      어 아까 너가
           e     akka  ne-ka
           yes earlier you-NOM
           "Earlier you said"

2 HS: →    응
           ung

3 HB:      꿈이
           kkwum-i
           dream-NOM
           "Your dream was"

4 HS: →    응
           ung

5 HB:      뭐 (0.2) 집에 피시방 차리고
           mwe    cipey     phisipang chali-ko
           what home-LOC    PC room   make-CNN
           "Well, to make a computer room at your place and"

6 HS: →    어
           e

7 HB:      남편이랑          게임하는                    거랬잖[아
           namphyen-ilang keyim-ha-nun  ke-la-yss-canh [a
           husband-with       game-do-RL thing-QT-PST-you:know
           "Play video games with your husband, you said."

8 HS: →                                                  [어(0.5)
                                                         [e(0.5)
```

```
9          (nod)(nod)(nod)(nod)

10 HB:     그래서 남편이 있어서 난 놀랬어 사실 그 꿈에
           kulayse namphyen-i  iss-ese        na-n  nolla-yss-e
           sasil     ku   kkwumey
           so       husband-NOM exist-because I-TOP surprised-PSS-INT
           actually that dream-LOC
           "I was actually surprised because there was a husband
           in your dream."

11 HS:     아  진짜
           a  cincca
           oh really
           "Oh, really."

12 HB:     어  [>피시방을 차리고< 혼자 게임할 수도 있잖아
           e  [>phisipangul  chaliko< honca keyim-ha-l swu-to
           iss-canha
           yes  PC room-ACC  make-CNN alone game-do-RL way-also
           exist-you:know
           "You could make a computer room and play video games
           by yourself."
  (HS)        [(nod) (nod)

13 HS: →   어 (0.2) (nod) [(nod)(nod)(nod)
           e

14 HB:                    [근데 왜 남편이 있어?
                          [kuntey way namphyeni   isse?
                          [then   why husband-NOM exist-INT
           "Then why is there a husband (in your picture)?"

15 HS:     어?
           e?
```

At the end of each speaking turn, the listener deploys a reactive response. After the speaker's first turn "earlier you said" and the second turn "your dream was", HS produces a lexical word "*ung*" with a falling intonation as shown in line 2 and 4. In line 5 and 7, the speaker gives more specific details about what HS mentioned earlier in the two separate turns, "to make a computer room at your place" and "play video games with your husband, you said". A notable thing is that the speaker's fourth turn "play video games with your husband, you said" ends in the ending "-

canha (you know)" which functions to solicit agreement (Kawanish and Sohn, 1993). Here, the speaker uses "-*canha*" to confirm whether HS's statement was correctly quoted by her. These two turns are, unlike the first two speaking turns, responded to with a lexical word "*e*" carrying a falling intonation. After the speaker's fourth turn, the listener also nods four times while producing "*e*". When the speaker shares her feelings that she was actually shocked by HS's comment, HS takes HB's reaction as quite surprising and displays this through the response "*a cincca* (really)" (line 10-11). In line 12, the speaker briefly responds to the listener's reaction with "*e* (yes)" and goes on to her next turn "you could make a computer room and play video games by yourself" again with the ending "-*canha* (isn't it)". This time "-*canha*" may function to seek an agreement on the legitimacy of the alternative option suggested by the speaker. This turn is responded to with "*e*" with a falling intonation in line 13. In line 14, HB finally brings up the question that she has built up the foundation for, "why is there a husband in your picture?", and hands over the floor to the current listener HS.

(1) Affective stance

In this fragment, the listener responds to the speaker's turn with two "*ung*" tokens and three "*e*" tokens. The two "*ung*" tokens appear in line 2 and 4. The affective stance of the responses is shown in their prosodic features. Both of the two "*ung*" tokens carry a falling intonation as presented in Figure 3 and 4[4].

[4] [4] The values used for HB and HS voice range were: 120 - 380 Hz.

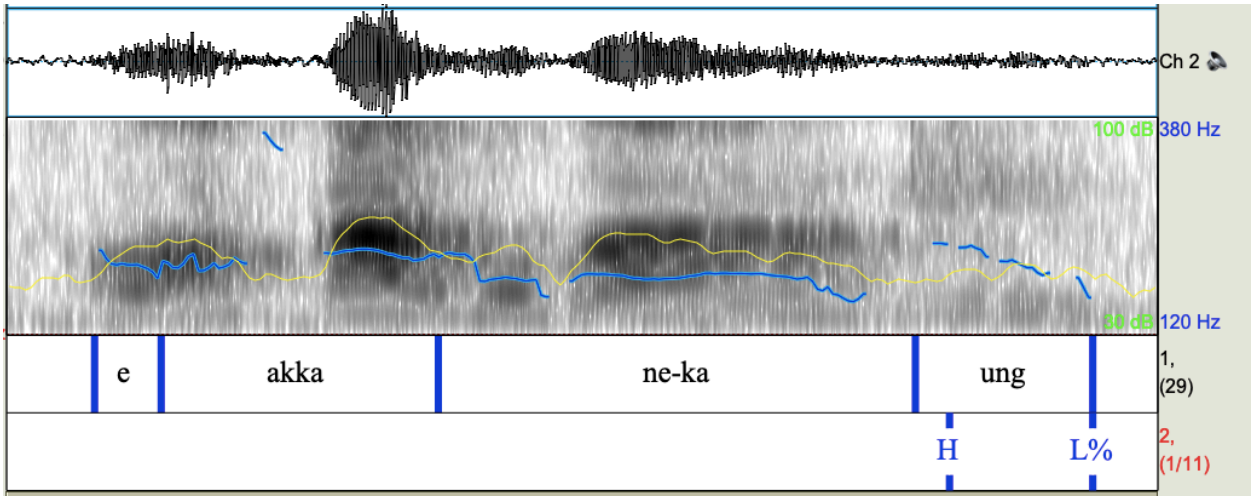


Figure 3. Praat information of the first "*ung*" token (HS) and its prior turn (HB)

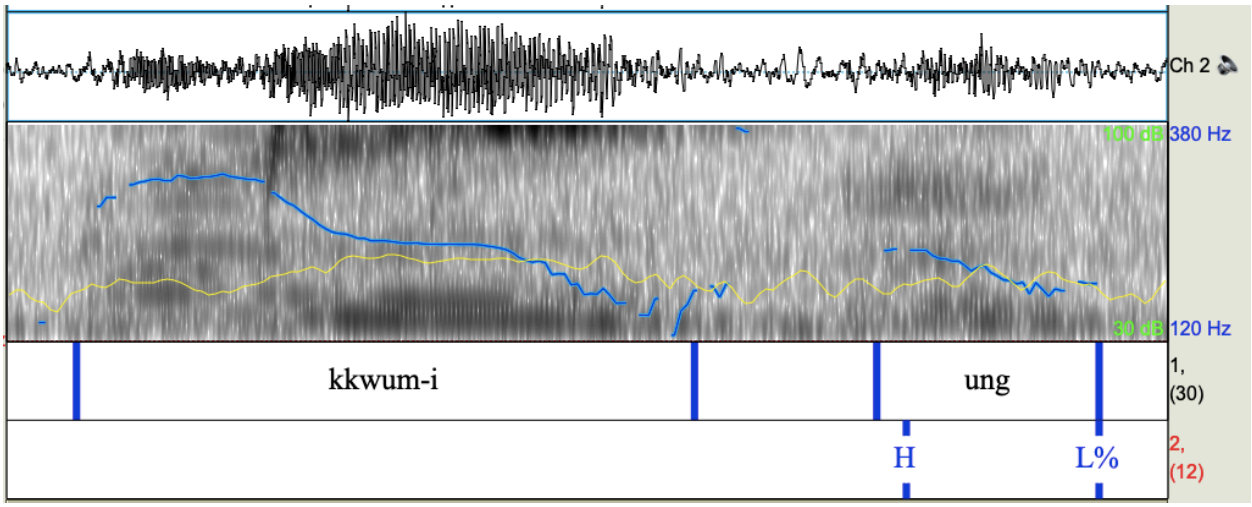


Figure 4. Praat information of the second "*ung*" token (HS) and its prior turn (HB)

In line 1 and 3, the speaker begins the floor with the opening comments, "earlier you said" and "your dream was", by which it is projected that the upcoming story will be regarding something the listener HS mentioned about her dream. While the first two turns only offer limited information of the story to be unfolded, they do not insinuate the speaker's actual intention of the telling. Therefore, at this point, the listener has no other choice but to wait until she is given more specifics. However, those information provided thus far still gives some hints about the telling

that is not so trivial to be passed on or simply nodded on. Thereby the listener chooses the lexical token "*ung*" carrying a falling intonation. On yes-type tokens, the falling intonation occurs the most frequently in my data expressing neither excessive nor trivial but mild interest in the ongoing telling. Also, in relation to their prior turns, the tokens show a similar pitch height and pitch range (10 - 350 Hz), but a lower intensity. I claim that because of the low intensity, the tokens are considered to be displaying a minor pitch downgrade (Curl, 2005; Ogden, 2006; Thompson et al., 2015). This minor pitch downgrade supports the mildness of the listener's interest towards the prior turn.

The first two "*e*" tokens appear in line 6 and 8. The affective stance of the responses is presented in their prosodic features. Both of the two "*e*" tokens also carry a falling intonation as shown in Figure 5 and 6[5].

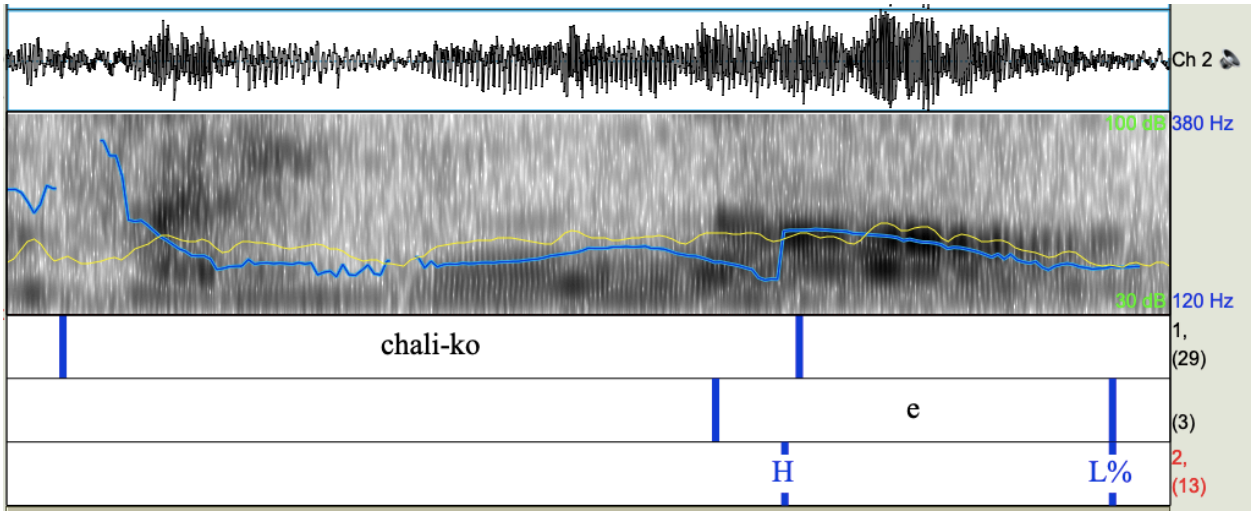


Figure 5. Praat information of the first "*e*" token (HS) and its prior turn (HB)

[5] The values used for HB and HS voice range were: 120 - 380 Hz.

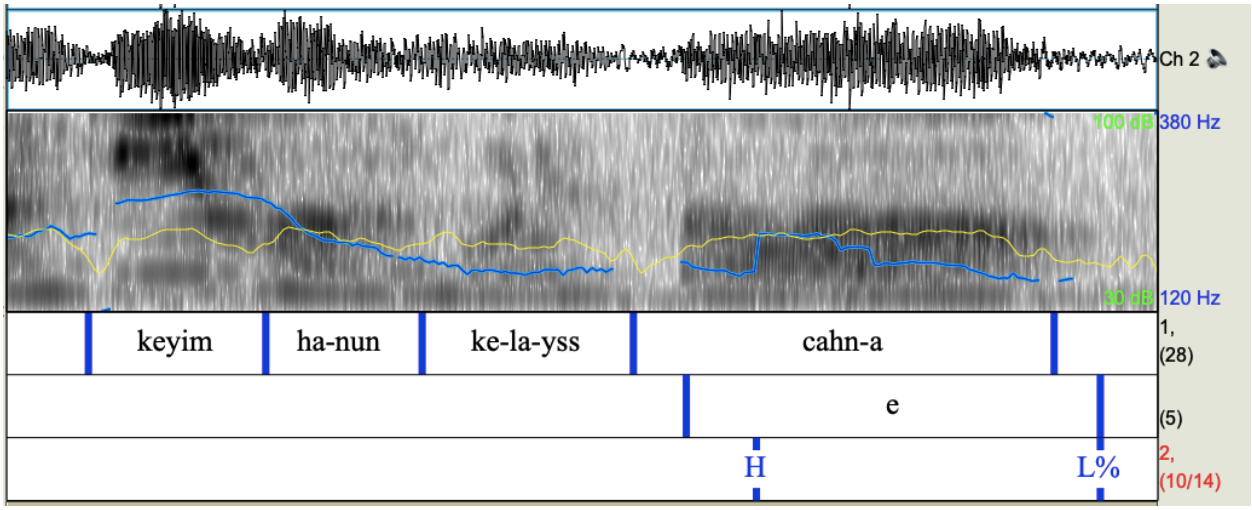


Figure 6. Praat information of the second "*e*" token (HS) and its prior turn (HB)

In line 5 and 7, the speaker specifies the telling by adding details. Through the two turns, "to make a computer room at your place" and "play video games with your husband, you said", the speaker quotes a particular point of the listener HS's statement. In contrast to the first two turns in line 1 and 3, these turns provide more crucial information that seems to be closely related to the speaker's intention of telling. Therefore, the listener responds to these with the lexical word "*e*" which is more engaging than "*ung*". Also, compared to their prior turn, the tokens show a higher intensity, while their pitch height and pitch range are similar to them. I argue that the higher intensity denotes a minor pitch upgrade (Curl, 2005; Ogden, 2006; Thompson et al., 2015). Contrary to the "*ung*" tokens above, the pitch upgrade of the two "*e*" tokens reinforces the listener's relatively stronger interest towards the prior turn.

One notable thing is that the second "*e*" token co-occurs with a few nods. As briefly mentioned above, the speaker's fourth turn "play video games with your husband, you said" ends in the ending "-*canha* (you know)". "-*canha*" is known to function as an agreement seeking

method (Lee, 1999; Kawanish and Sohn, 1993). By "-*canha*", the speaker tries to get confirmation of whether her quotation of HS's remark is correct. Even though "-*canha*" does not shape the sentence into an interrogative form and therefore it does not obligate the listener to provide an answer as in a question-answer sequence, it strongly invites the listener's confirmation. Here, the strong function of this device seems to have motivated the listener to respond with more than a single lexical token. Therefore, the listener produces several nods along with the lexical token to express her confirmation in addition to her interest.

The "*e*" token is shown one more time in line 13[6]. The "*e*" token also draws a falling intonation.

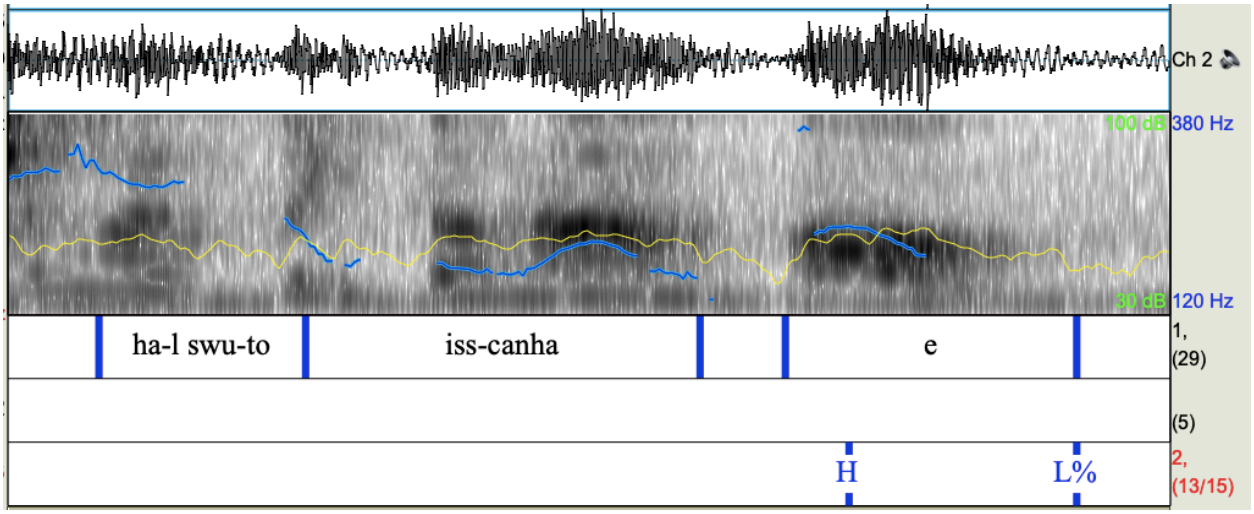


Figure 7. Praat information of the third "*e*" token (HS) and its prior turn (HB)

After expressing that she was shocked by HS's future plan in line 10, the speaker proposes what she thinks can be an alternative option in line 12 and this turn also ends in "-*canha*". Similar to "-*canha*" in line 7, the listener is highly encouraged to display a response with extra decoration to

[6] The values used for HB and HS voice range were: 120 - 380 Hz.

express her confirmation as well as her interest. Here, the listener produces a lexical word "*e*" and three subsequent nods. However, unlike the previous response in line 8, this time the nods do not co-occur with the lexical token simultaneously. There is a pause of 0.5 seconds before the nods begin. While "-*canha*" in line 7 simply requests confirmation of correctness of the quotation, "-*canha*" in line 13 seeks for an actual agreement on the legitimacy of the newly suggested option that could possibly threaten the listener's negative face. Therefore, the pause and the delayed nods reveal the listener's hesitancy to readily expressing a full agreement on the idea that is presumably opposing to her own. Also, in relation to the prior turn, the token shows a similar intensity, pitch height, and pitch range (10 - 450 Hz). The "*e*" token is not displaying a pitch upgrade or a pitch downgrade but stays in the neutral status. This, what I call, "pitch retention" displays the listener's mild interest and neutral stance towards the prior turn.

In sum, while both the lexical token "*ung*" and "*e*" with a falling intonation display mild interest towards the prior turn, "*e*" token is deployed when the information given is relatively more significant. Also the nods that co-occur with or follow the verbal token can be used to display the listener's confirmation of the validity of the information besides her interest.

(2) Epistemic stance

Since "yes" type tokens have been considered to serve as "continuers", they have not been investigated with respect to the epistemic stance they can display. However, the listeners do epistemically position themselves through this simple type of response.

First, the listener displays her epistemic authority through the lexical choice of the verbal token. Epistemic authority is a "right" to claim something that follows from "the knowledge gained through epistemic access" (Heritage and Raymond 2005; Enfield, 2011; Thompson et al., 2015). Through line 1 to line 7, the speaker HB retrieves what the listener HS said, using it as the background of the question that she eventually wants to ask in line 14. Therefore it is evident that the listener has direct access to the telling. However, after the speaker's first two turns, since the information given by this point is relatively limited to figure out what it is about, the listener produces the "*ung*" tokens. On the other hand, after line 5 and 7 where the more significant information of the telling has been delivered, the listener deploys the "*e*" tokens. Unlike the first two turns, these turns quote a very specific part of the listener's statement which is sufficient for her to grasp the story being unfolded. After the listener has comprehended that the exact part of the telling that she has direct access to, she changes the lexical form from "*ung*" to "*e*". I argue that the listener claims the higher epistemic authority than the speaker for the telling by employing the "*e*" tokens, while she does not do so with the "*ung*" tokens.

The distinction between "*ung*" and "*e*" can be compared to that between vocal continuers and nods in the study of Stivers (2008). According to her, the listeners produce nods when they have access to the teller's stance, whereas they deploy vocal continuers otherwise. The access depends on how specific the information given to the listener is. The more specific it is, the more accessible it can be for the listener "to understand what it was like to experience the event" (Stivers, 2008). In terms of accessibility, this differentiation is similar to the one shown in my data. However it is different in that, in this fragment, the listener has the higher authority than

the speaker with respect to the telling because it was originally said by the listener. Therefore, by using "*e*" token, the listener can claim her higher epistemic authority for the current telling.

Second, the listener also displays her epistemic stance through the reactive responses. Although the listener claims the higher epistemic authority for the telling, it is the speaker who is more knowledgeable about the progress of the telling. Through each turn, the speaker provides with new information and the listener becomes more informed of the telling in progress. In this fragment, through the "*ung*" and "*e*" tokens, the listener positions herself as more informed than before. This does not mean that the listener is a complete K+ and the epistemic gap is fully filled. However, it is evident that she becomes more and more knowledgeable about the telling going through each turn. Here, I claim that, borrowing the expression from Thompson et al. (2015), the listener is on the "intermediary stage on a path between K– and K+". Also, in terms of epistemic stance, I argue that the "*e*" tokens display a higher K+ status than do the "*ung*" tokens, given the fact that the "*e*" tokens are employed after the more significant information, while the "*ung*" tokens are used after the background information.

In sum, the listener claims the higher epistemic authority for the telling than the speaker with the "*e*" tokens, but she does not do so with the "*ung*" tokens. Also, in terms of epistemic stance, the listener displays that she is more informed than before through both the "*ung*" and "*e*" tokens, while she shows a higher K+ status through the "*e*" tokens.

4.2. Vocalization "*mm*" and nods

In this section, vocalization "*mm*" and nods will be investigated. First, I will discuss vocalization "*mm*". Although "*ung*" and "*mm*" are phonetically very similar, they are differentiated semantically and phonetically. First, the former is a lexical word meaning non-polite "yes", whereas the latter is merely a non-lexical vocalization. Therefore "*mm*" is semantically much weaker than "*ung*". Second, the lexical word "*ung*" consists of an unrounded back vowel /ɯ/ and the following velar nasal consonant /ŋ/, therefore the lips are relaxed but slightly open during the articulation. On the contrary, the vocalization "*mm*" is a syllabic nasal sound [m̩] and the lips remain closed throughout the whole articulation. Since the syllabic nasal [m̩] can be pronounced without lip movement, producing "*mm*" takes far less effort than producing "*ung*". Therefore the vocalization "*mm*" can be taken to be less engaging than the lexical word "*ung*" because it is semantically weaker and phonetically effortless.

In fragment 2, the speaker HB tells about her experience in adapting cats and how her dermatological condition triggered by them was cured.

<Fragment 2: Cats>

```
1 HB:      나도 처음에 고양이들 데려왔을 때 집에
           na-to cheum-ey   koyangi-tul teylyewa-ss-ul ttay cip-ey
           I-also first-TMP cat-PL      bring-PST-RL   when home-LOC
           "When I first brought cats to my house"

2 HS:      (gaze only)

3 HB:      아토피가                  [내가 원래 있어서
           athophi-ka                [nay-ka wenlay    iss-ese
           Atopic:dermatitis-NOM     [I-Nom  initially exist-because
           "Because I used to have atopic dermatitis"
 (HS) →                              [(nod)(nod)(nod)
```

```
4 HS: →    음
           mm

5 HB:      한 3년을            [(0.2) 되게 안  좋았거든
           han   3nyen-ul      [(0.2) toykey an  coh-ass-ketun
           about 3 years-ACC   [      very   not good-PST-CORL
           “It was not good for about 3 years.”
(HS)  →                        [(nod) (nod) (nod) (nod)

6 HS: →    [음
           [mm
           [(nod)(nod)

7 HB:      집에      [가면은 >나는          [같이 살진 않았는데<
           cip-ey    [ka-myen-un >na-nun    [kathi  salci-n anh-ass-nuntey<
           home-LOC  [go-if-TOP  >I-TOP     [together live-TOP not-PST-but<
           “When I go home, I did not live together but,”
 (HS) →              [(nod)(nod)            [(nod)(nod)

8 HS:      (gaze only)

9 HB:      근데 3년       [지나니까
           kuntey 3nyen   [cina-nikka
           but 3 years    [pass-because
           “But after three years”
 (HS) →                   [(nod)(nod)

10 HB:     [뭔가 익숙해져 가지고                  [없어졌어
           [mwenka     ikswukha-ycye kaciko       [eps-ecy-ess-e
           [something:is accustomed-because       [disappear-become-PST-INT
           “I got used to it and it was gone.”
 (HS)  →   [(nod)(nod) (nod)(nod)

11 HS: →                                          [어
                                                  [e

12 HB:     이제 더이상 상관없어
           icey teisang sangkwaneps-e
           now  anymore do:not:matter-INT
           “It does not matter any more.”

13 HS: →   (nod) (nod)

14 HB:     (0.3) [같이 살면 괜찮아지나 봐 또
           (0.3) [kathi    sal-myen kwaynchanh-aci-na pwa tto
                 [together live-if  fine-become-CONJ      again
           “I think it gets better when you live together (with cats).”
 (HS)  →         [(nod)(nod) (nod)(nod)
```

15 HS: → (nod)(nod) [(nod)(nod)

16 HB: [(nod)(nod)

In line 1, 3 and 5, the speaker HB gives the background information that she had a hard time due to her atopic dermatitis for the first three years after she adapted cats. In line 4 and 6, HS responds to it with a single vocalization "*mm*" with a falling contour at the end of each turn. The vocalization is treated as an appropriate response to the speaker. In line 6, the speaker quickly appends a compressed utterance and this turn is not responded to vocally or physically while the listener's gaze is fixed at the speaker. However, it is not considered problematic by the speaker because this is an additional fact that belongs to the background information stated in the preceding turn. In line 8, HB finally provides the new information that after spending three years with the cats, she got accustomed to being around them and her dermatitis was cured. This turn contains the key point of the story and the listener reacts to it with a lexical word "*e*" meaning "yes" in line 11. A notable thing is that HS produces the response before the turn is completed, immediately after the connective "*ikswukha-y-cye kaciko* (I got used to it)" and it overlaps with the first syllable of the remaining part of the same speaking turn "*epse-cy-ess-e* (it was gone)". In line 12, HB's turn "it does not matter any more" is a rephrasing of the previous turn and this is simply responded to by HS's nods. In line 14, the speaker summarizes the whole story into a sentence "I think it gets better when you live together (with cats)". The listener reacts to this utterance with a few nods and her reaction is again reciprocated to by the speaker's synchronizing nods.

(1) Affective stance

In line 4 and 6, the listener produces the "*mm*" tokens with a falling intonation as presented in Figure 8 and 9[7].

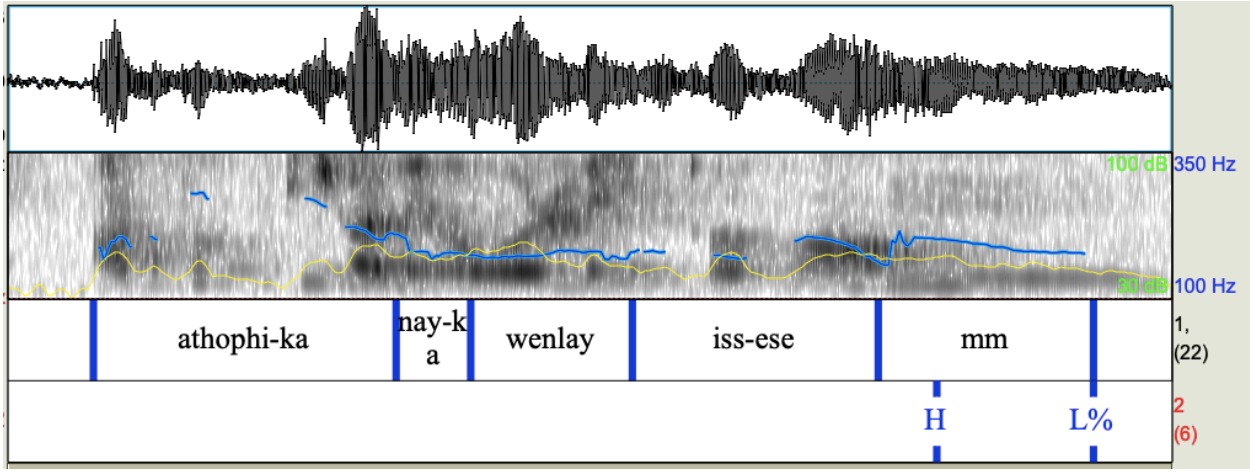


Figure 8. Praat information of the first "*mm*" token (HS) and its prior turn (HB)

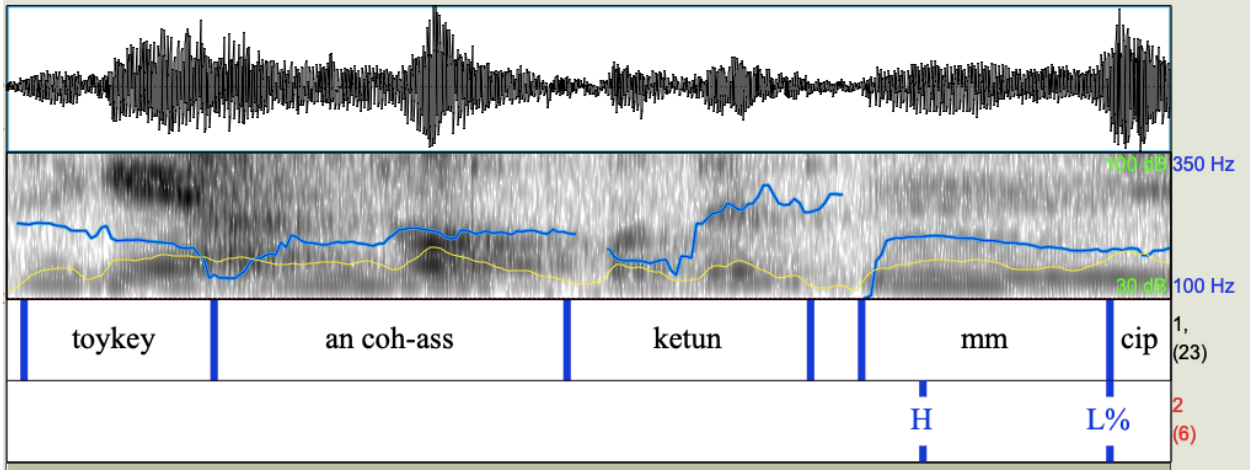


Figure 9. Praat information of the second "*mm*" token (HS) and its prior turn (HB)

As discussed earlier, a falling intonation on "yes" type tokens expresses mild interest towards the telling. On the semantically weakest form "*mm*", the mildness is even enhanced. The listener is

[7] The values used for HB and HS voice range were: 100 - 350 Hz.

displaying a lowered level of interest through the single vocalization "*mm*" with a falling intonation at the end of the first and the second speaking turn through which the speaker delivers the background of the telling. The "*mm*" tokens show a minor pitch downgrade because of the lower intensity in relation to their prior turns. The pitch downgrade supports that the listener's interest is restricted to a mild degree.

On the other hand, the lexical word "*e*" is produced in line 11 (Figure 10[8]).

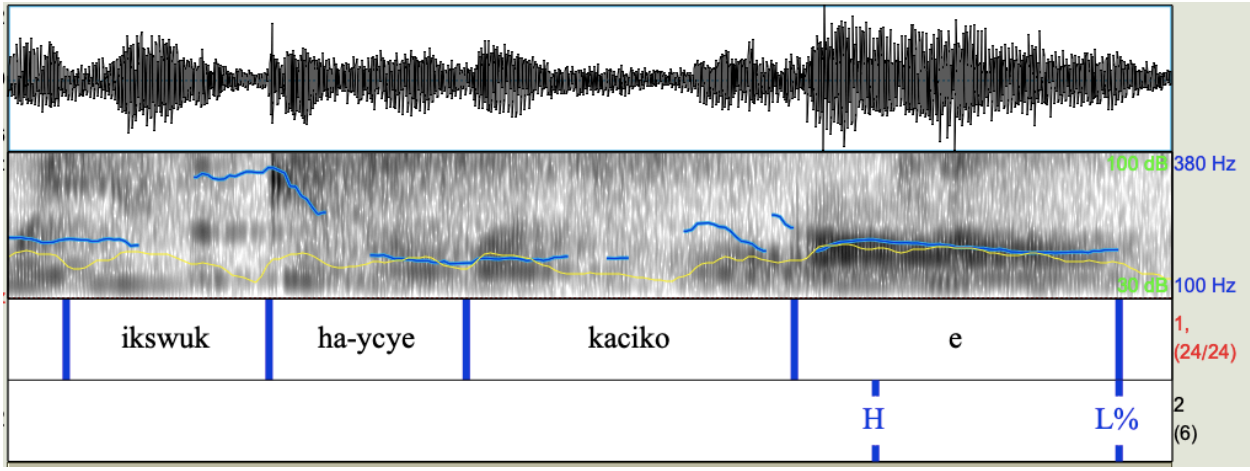


Figure 10. Praat information of the "*e*" token (HS) and its prior turn (HB)

Not only is the token type different, but also this token shows a minor pitch upgrade through the higher intensity. The upgraded lexical "*e*" token is deployed after the speaker has provided the core information that her dermatitis was cured after spending three years with the cats. This contrasts to the two "*mm*" tokens that are produced after the relatively less remarkable background information. The environmental difference between "*mm*" and "*e*" are analogous to the difference between "*ung*" and "*e*" discussed in the previous section. However, through the

[8] The values used for HB and HS voice range were: 100 - 380 Hz.

vocalization "*mm*", the listener can take it down a notch with regard to the level of interest. I argue that the vocalization "*mm*" expresses the mildest interest among all of the verbal tokens.

Next, let me look into nods that occur independently, not the ones accompanied with verbal tokens. As the story unfolds, the listener also continuously nods throughout the telling. However the nods during the utterance and the ones at the end of the turn should be treated differently. Compared to other responses, nods are more likely to be produced when the turn is on the way, because they do not acoustically disrupt speaker's turn. This mid-turn type of nods are commonly found in Korean and Japanese in which the overall occurrences of reactive response are relatively higher than in English and Mandarin (Mayland, 1990; Clancy, 1996; Iwasaki, 1997). This type of nodding is the mildest reactive response that minimizes the interruption while showing the listener's consistent attention to the telling. I will call this "mid-turn nods".

In fragment 2, the speaker HB begins a turn projecting a story consisting of the several consecutive turns. However she is not responded by the listener at the end of the first turn. The speaker continues her story with her second turn but the listener still does not react to it. There, the speaker shifts her gaze towards the listener to solicit a reaction in the midst of the second turn. The listener finally starts nodding at the moment she is gazed at. From this moment, the listener continuously produces nods during the speaker's turn throughout the telling (line 3, 5, 7, 9, 10 and 14). Through these mid-turn nods, the listener signals that she is not going to interrupt but will stay attentive to the story. The affective stance it displays can be seen as very mild interest or attentiveness. However, compared to the turn completion points, the listener is much less obligated to provide responses in the midst of the turns. The listener's gaze alone can display

the listener's attention although it is weaker than nods (line 2 and 8). Therefore, offering mid-turn nods should be considered more than just attentiveness; it shows extra effort from the listener. I claim that mid-turn nods display a high level of interest.

The second type of nods are the ones situated at the completion points. Nods at the completion points are differentiated from the mid-turn nods. While the mid-turn nods are produced to minimize interruption when the speaker's turn is on the way, the listener is given more options to choose from in terms of types of response at transition relevant places. By choosing nods over other types of response at these places, the listener intentionally lowers the level of engagingness treating the preceding turn less notable.

In this fragment, it is first shown in line 13. In line 10, the speaker provides the main information "after three years, I got used to it (being around the cats) and it (atopic dermatitis) was gone". The listener responds to this turn with the "*e*" token (line 11). On the contrary, after the speaker has appended a redundant remark "it does not matter any more" in line 12, the listener only nods at the turn completion point. By managing the type of responses, the listener differentiates the level of engagingness towards the speaker's two consecutive turns. From the view of affective stance, I consider that nods at turn completion points display "attentiveness" that is even weaker than the mildest interest displayed through the vocalization "*mm*".

A similar type of nods is also found in line 15 where it is used in response to the speaker's last turn of the sequence that encapsulates the telling in one sentence "I think it gets better when you live together (with cats)". Since the relatively more important parts of the telling have been already provided, the summary does not offer any new information, but merely winds

up the sequence. Therefore, the listener reacts to this utterance with a few nods signaling that her interest towards the prior turn is not the extent to which she would produce a vocalization or a lexical token, but is something to simply nod to. This reactive response is accepted properly by the speaker as shown in her reciprocating nods in line 16.

To sum up, while mid-turn nods display a more intense degree of interest because of the less obligation imposed on the listener in the midst of the turns, the nods at turn completion points express mere attentiveness since they are even weaker than the mildest form, the vocalization "*mm*".

(2) Epistemic stance

The vocalization "*mm*" tokens and nods also display the listener's epistemic stance. First, similar to the "*ung*" and "*e*" tokens, the "*mm*" token expresses that the listener is now informed of a piece of information delivered through the prior turn. Throughout the telling, the speaker HB shares her own experience to which the listener HS does not have any access. The listener produces "*mm*" in line 4 and 6, after she has been given the background information of the telling. After each turn, "because I used to have atopic dermatitis" in line 3 and "it was not good for about 3 years" in line 5, the listener has obtained new information and positions herself as more informed than before. In other words, through the "*mm*" tokens, the listener shows that she has processed the new information and has become "better" informed although she is not fully K+ yet because she is aware of the more crucial telling is coming up.

I will also briefly discuss the "*e*" token in line 11. In relation to the two "*mm*" tokens employed after the limited background information, the "*e*" token is used after the main information which seems to be more "unexpected" to the listener. In my data, at these locations, "oh/really" type tokens[9] are commonly deployed expressing the listener's upgraded epistemic stance and "surprise" in terms of affective stance. Here, however, the "*e*" token is chosen here instead. The "*e*" token with a falling intonation also displays the listener's upgraded epistemic stance, however, it does not express "surprised" feeling. I claim that the only difference between the "*e*" token and "oh/really" type tokens at this place is the affective stance they display. By choosing the "*e*" token over one of the "oh/really" type tokens, the listener simply shows her becoming more informed with the telling while displaying a high level of interest instead of astonishment.

Nods, on the other hand, show the two different types as I explained above; (1) mid-turn nods; (2) turn-completion nods. Unlike the turn-completion nods that are placed where a response is due, the mid-turn nods occur when the turn is on the way. Therefore, when producing such nods, the listener has not been informed with the information being conveyed in the moment. I consider the listener deploying the mid-turn nods to position herself as K-.

On the other hand, the nods placed at turn completion points should be compared to other type of responses used at the same places. In this fragment, the turn-completion nods appear in line 13 and 15. Compared to the core information, "I got used to it and it was gone", given in line 10, the turns in line 12 and 14 contain rather "redundant" or "predictable" information; "It does not matter any more", "I think it gets better when you live together (with cats)". Therefore, the

[9] "Oh/really" type tokens will be discussed in Chapter 5.

listener's nods after these two turns claim her prior knowledge and display her already earned K+ status treating the information not as "informative" as the one in line 10 to which she responded with the "*e*" token.

In sum, the "*mm*" tokens indicate that the listener is "better" informed than before while she is on a path between K– and K+. The "*e*" token shows the listener's upgraded epistemic stance, but it does not display "surprise". Also, through the mid-turn nods, the listener positions herself as K-, because she has not been informed with the information being conveyed. However, the listener displays her already earned K+ status with the turn-completion nods treating the prior turn as not informative, because the information given is either redundant or predictable to her.

4.3. Multiple saying "*mm*" and "*e*"

In this section, multiple saying type of responses will be discussed. The term "multiple saying" is defined by Stivers (2004) as a unit of talk (e.g., a word, phrase, or sentence) repeated multiple times in the same turn under a single intonation contour such as "*no no no*" or "*wait wait wait*". The sound pattern of a multiple saying shows a slightly higher pitch in the beginning and draws a gradual fall until the end without a break in phonation. The repeats are commonly latched together and delivered quickly. The whole repeats are treated as a single action although each token comprises a syntactically complete unit. The finding shows that multiple sayings are a resource to address an in progress course of action rather than the prior turn alone. The user of multiple sayings display a stance that his or her interlocutor should halt their unnecessarily persisted actions. In their study of the discourse-pragmatic functions of Korean "*yes*" tokens,

Pyun and Yoon (2022) claim that Korean "yes" tokens can also function as "proposal to discontinue the on-going action for the sake of a larger course of action" when it is produced as a multiple saying. Nagata (2004) studied the difference between *aizuchi* in the middle of utterance and during pauses and found the formal tends to be presented in the repetitive type (e.g., *a:a:* and *nn nn).* He argues that the production of *aizuchi* in the middle of utterance displays that its producer has received sufficient information to comment on the matter and therefore he or she is ready to take a turn.

In her thorough study of "*nn*" token in Japanese conversation, Aoki (2008) suggests a different point of view on multiple saying. She uses the term "complex tokens", as opposed to "plain tokens", referring to "*nn*" and head nods produced successively in a variety of combination. She found that complex tokens can (1) display affiliation with its producer's affective/evaluative stance and (2) highlight a mutual epistemic stance by claiming prior knowledge of information provided. Yang (2013), who explored multiple saying "*dui dui dui*" in Mandarin conversation, discovered that the token is commonly used in response to the speaker's assertion displaying the listener's affiliation, in contrast with "*en*" token which simply denotes the listener's (structural) alignment. According to the study, the phonetic properties of "*dui dui dui*" show a stress on the first *dui* indicated by largest amplitude, longest duration and noticeable falling intonation and level pitch movement throughout the token. The token is also often accompanied with three head nods during its production. In this section, I will look into the "yes" type reactive responses in the form of multiple saying in Korean conversation and examine their stances and discourse functions based on the prosodic features and non-verbal behaviors.

In fragment 3, the speaker HS tells about an unexpected benefit of wearing a wedding ring at her workplace.

<Fragment 3: Wedding ring>

```
1 HS:       아  그리고 예상 외로 또 좋은 점이
            a  kuliko yeysang      oylo     tto    coh-un   cem-i
            oh and    expectation  outside  again  good-RL  point-NOM
            “Oh, and also the unexpected benefit is”

2 GA:       (nod)(nod)

3 HS:       내가 일할 때는 막
            nay-ka ilha-l    ttay-nun mak
            I-NOM  work-RL   when-TOP just
            “When I work”

4 GA:       (nod) (nod)

5 HS:       부모님들 (0.2) [상대하는 일이 많은데,
            pwumonim-tul [sangtayha-nun  il-i      manh-untey,
            parent-PL    [deal:with-RL   work-NOM  many-CIRCUM
            “I deal with parents parents most of the time”
 (GA)                    [(nod)(nod)(nod)

6 GA:       [음
            [mm
            [(nod) (nod)

7 HS:       [특히 한국 부모님들이랑 있을 때는 °내가°
            [thukhi    hankwuk pwumonimtul-ilang iss-ul    ttay-nun
             °nay-ka°
            [especially Korean  parents-with    exist-RL  when-TOP
              I-NOM
            “Especially when I am with Korean parents, I,”

8 GA:       (gaze only)

9 HS:       어려보이는 거에 대해서:
            elyepoi-nun     ke-ey tayha-yse:
            look:young-RL   thing-about
            “About (me) looking young”

10 GA: →    [음:음 음 음
```

```
            [mm:mm mm mm
            [(big nod)(nod)(nod)(nod)

11 HS:      쫌 그걸 내가 안 좋아하거든
            ccom   kuke-l          nay-ka an  cohaha-ketun
            little that:thing-ACC  I-NOM  not like-CORL
            "I don't like it."

12 GA: →    [어 어 어 어 어
            [e e e e e
            [(nod)(nod)(nod)

13 HS:      나는 되게 전문적이고                  [그렇게 보이고 싶은데 막
            na-nun toykey cenmwunceki-ko   [kulehkey poi-ko
            siph-untey  mak
            I-TOP  very   professional-CNN [like:that be:seen-CNN
            want-but    just
            "I want to look very professional and such but like"
  (GA)                                     [(nod)(nod)(nod)(nod)

14 GA: →    [어 어
            [e e
            [(nod)(nod)

15 HS:      어른들이        막 (0.5)    [그렇게 보니까
            eluntul-i     mak (0.5)  [kulehkey  po-nikka
            adult-PL-NOM  just       [like:that see-because
            "Because people see me that way."

16 GA:                                 [어
                                       [e
                                (nod) [(nod)(nod)
```

In line 1, the speaker HS begins the telling with the topic utterance with a preface "*a*" token equivalent to the "*oh*" token in Ensglish. In the next turn, she gives a detail of the background of the telling, "when I work". These first two turns are responded to by the listener with two small nods at the end of the each turn. The speaker offers more information in line 5 that she has to deal with parents most of the time at her workplace. This specified piece of information is reacted to by the listener GA with a vocalization "*mm*" with co-occurring nods at the end. A

notable thing is that the listener also produces short mid-turn nods when the speaker has said "parents" (line 6). In line 7, by saying "especially when I am with Korean parents, I", the speaker adds a small detail that the "parents" she is talking about are Korean. Here, the speaker finishes the turn with "*naeka* (I)" insinuating that the turn is to be continued soon. In addition to that, the speaker takes her gaze away from GA in the middle of the turn when saying "*iss-ul ttay-nun* (I am with)" and also lowers the intensity of the sound of "*naeka* (I)". Through this behavior, the speaker is trying to reassure the listener that there is no need to display any response this time because there is not so much of new information. The speaker's behavior minimizes the possible unnecessary effort from the listener's end because she considers the current turn is not worth the imposition. The listener, who seems to have understood the speaker's intention, does not provide any verbal or physical reaction throughout and after the speaker's turn although she keeps her eyes on the speaker's face. In line 9, the speaker HS finally says something noteworthy that the telling is regarding the fact that she looks young. While saying this turn, HS gives her gaze back to GA and lengthens the final syllable of the turn "-*se*". Here, the listener GA produces a multiple vocalization "*mm mm mm mm*" with a co-occurring big nod and following small nods. When she gives the first big nod with the first "*mm*" token, she also opens her eyes widely for a second. The speaker then goes on and says "I don't like that (fact)". As soon as this new and seemingly important information has been given, the listener articulates the lexical token "*e*" five times consecutively, "*e e e e e*", with accompanying nods. In line 13, HS appends extra information about her personal preference of the image of herself saying "I want to look very professional and such but like". The listener starts to nod with a relatively fast pace after the speaker has said "professional" and as the turn reaches the end, she produces another multiple token "*e e*" with

co-occurring nods. In line 15, the speaker rephrases her turn in line 9 saying that "because the middle aged people see me that way". This time the listener GA deploys a single "*e*" token with nods after HS has said "the middle aged people" and it overlaps with the speaker's remaining utterance "see me that way".

(1) Affective stance

The listener produces multiple saying tokens in different shapes during the telling. The first token, "*mm mm mm mm*" is shown in line 10 and its affective stance is displayed through the prosodic features as presented in Figure 11[10].

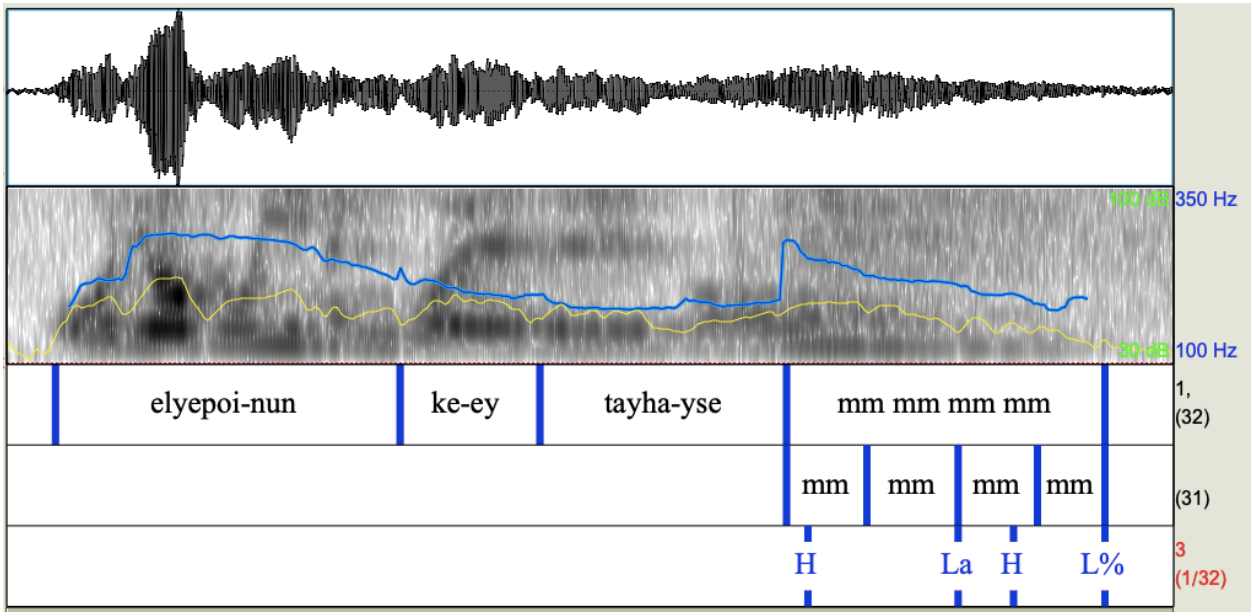


Figure 11. Praat information of the "*mm mm mm mm*" token (GA) and its prior turn (HS)

[10] Although the "*mm mm mm mm*" token consists of multiple syllables, not every syllable forms an AP because some syllables are pronounced faster and do not show any significant dips. In this token, the first two syllables and the last two syllables form an AP, respectively, cued by the intensity contour. Each AP is marked as "H La", if the pitch of the AP-initial tone is higher than that of the AP-final tone, even if the difference is slight.

The token is produced after the speaker has mentioned the first topic-related information that she looks young. The turn is topicalized by the word "*tayha-yse* (about)". The speaker strongly elicits a response by relocating her gaze back on the listener and indicating the turn boundary with the final lengthening on "-*se*". Here the listener produces a multiple saying token consisting of four vocalizations "*mm mm mm mm*" drawing a falling intonation contour throughout the production. Although the token is produced in succession, each vocalization is clearly pronounced separately carrying its own intonation contour. Similar to the prosodic features of "*dui dui dui*" (Yang, 2013), the first "*mm*" shows the most recognizable falling pitch while the following vocalizations carry a relatively level intonation. Also the first "*mm*" is most prominent and it is marked by the highest intensity and pitch among the four vocalizations. Compared to the single vocalization "*mm*" displaying the mildest level of interest, the multiple saying token "*mm mm mm mm*" expresses much higher engagingness. At this point, the listener seems to have inferred from the previous context that the upcoming telling will be somewhat negative (that the speaker does not like the fact that she looks young). Therefore, she expresses her enhanced empathy through the multiplied vocalization with a falling intonation. However, since the amount of the information given by far is still relatively little compared to what comes next, the weakest form, the vocalization "*mm*", is chosen for the token type.

In relation to their prior turn, the token shows a minor pitch downgrade with a lower intensity. Along with the choice of the weak vocalization form, the listener exhibits rather reserved engagingness through the reduced intensity. The impact of the decrease of intensity in this token will be more clear when compared to the next multiple saying token that is intentionally produced with a higher intensity.

Her affective stance is more highlighted by the co-occurring nods (a big nod and a few small nods) and her facial expression. The listener produces a big nod and raises her eyebrows in the moment when producing the first syllable of the repetition token with emphasis.

Figure 12. Seating arrangement of the participants; GA (left) and HS (right)

(Before the 1st nod) (1st nod & raising eyebrows) (Small nods)

Figure 13. Listener GA's head gesture and facial expression

The listener nods four times along with each vocalization and a few more after. The intensity of the nods gradually decreases as the prosodic features of each vocalization fade out. This finding is also congruent with the head nods accompanied with "*dui dui dui*" in Yang's (2013) study. I

argue that the co-occurring nods reinforce the listener's affective stance and here, it is interest and attentiveness. Also, while producing the first vocalization and the first big nod, the listener instantly raises her eyebrows. Eyebrow raising indicates surprise, fear, wonder or enthusiasm and genuine interest (Smart, 1986). In this case, the act of raising eyebrows emphasizes her interest towards the topic.

To sum up, the co-occurrence of the nods and the raised eyebrows with the first vocalization intensifies the interest displayed by the listener.

The second token, "*e e e e e*" is employed in line 12 and its prosodic features is presented in Figure 14[11].

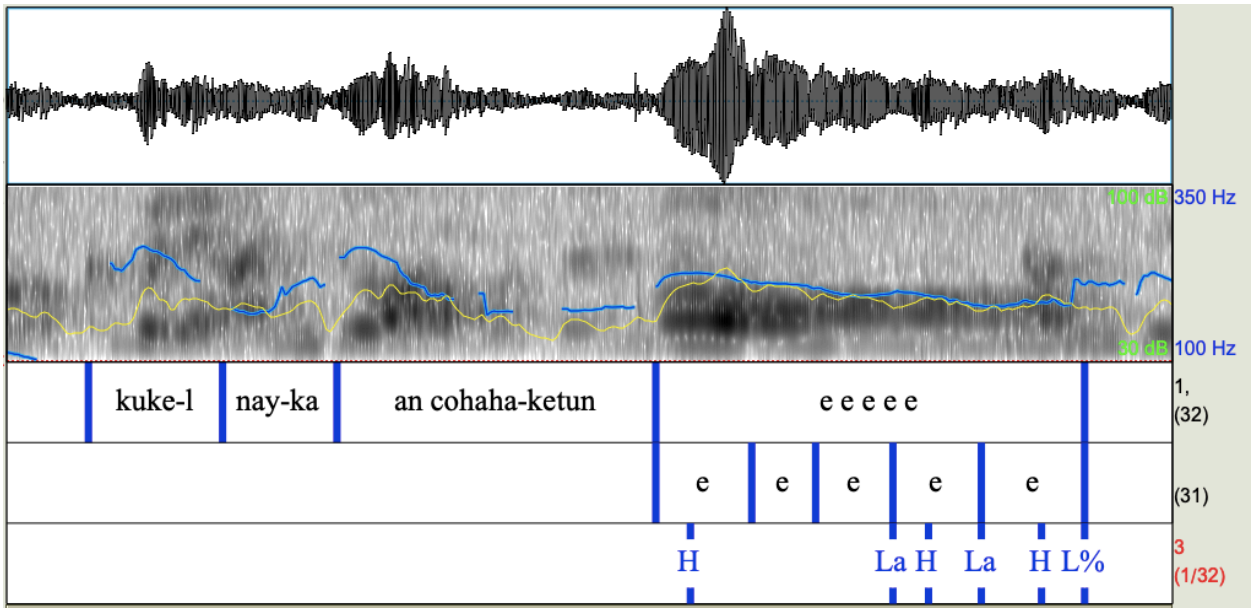


Figure 14. Praat information of the "*e e e e e*" token (GA) and its prior turn (HS)

Similar to Figure 11., not every syllable in the "*e e e e e*" token forms an AP because some syllables are pronounced faster and do not show any significant dips. I consider that this token has three APs cued by the intensity contour. Each AP is marked as "H La", if the pitch of the AP-initial tone is higher than that of the AP-final tone, even if the difference is slight.

This token is employed after the speaker has finally given the new information about her own preference that she does not like the fact she looks young. As a response, the listener articulates the lexical token "*e*" five times consecutively. Similar to the previous response, the multiple saying token "*e e e e e*" also draws a gradual falling intonation contour until it reaches the end with a stress on the first "*e*". By choosing the lexical token "*e*" and the higher number of repeats (five times), the listener noticeably upgrades the level of engagingness and thereby treats the immediate prior turn as more noteworthy information than the previous turn. Also, compared to the prior turn, the token shows a minor pitch upgrade with the higher intensity while pitch height and pitch range remain similar. The increase in the loudness of the token intensifies the listener's affective stance. Particularly, when compared to the previous token that shows a pitch downgrade, the upgrade is evident. This token is also accompanied with five nods, each of which is conducted with each vocalization. Therefore, with the higher intensity and co-occurring nods, the multiple saying token "*e e e e e*" shows enhanced interest towards the prior turn in accordance with the upgrade of the noteworthiness of the information given.

The prosodic features of the third token "*e e*" is shown in Figure 15. The third one is shown after the speaker has mentioned about her personal preference of the image of herself that she wants to look professional. The turn ends with the connective "-*untey* (but)" that strongly invites a response. The listener begins to nod at a rapid pace after the word "*cenmwunceki-ko* (professional)" and produces the multiple token "*e e*" with a falling intonation at the end of the turn.

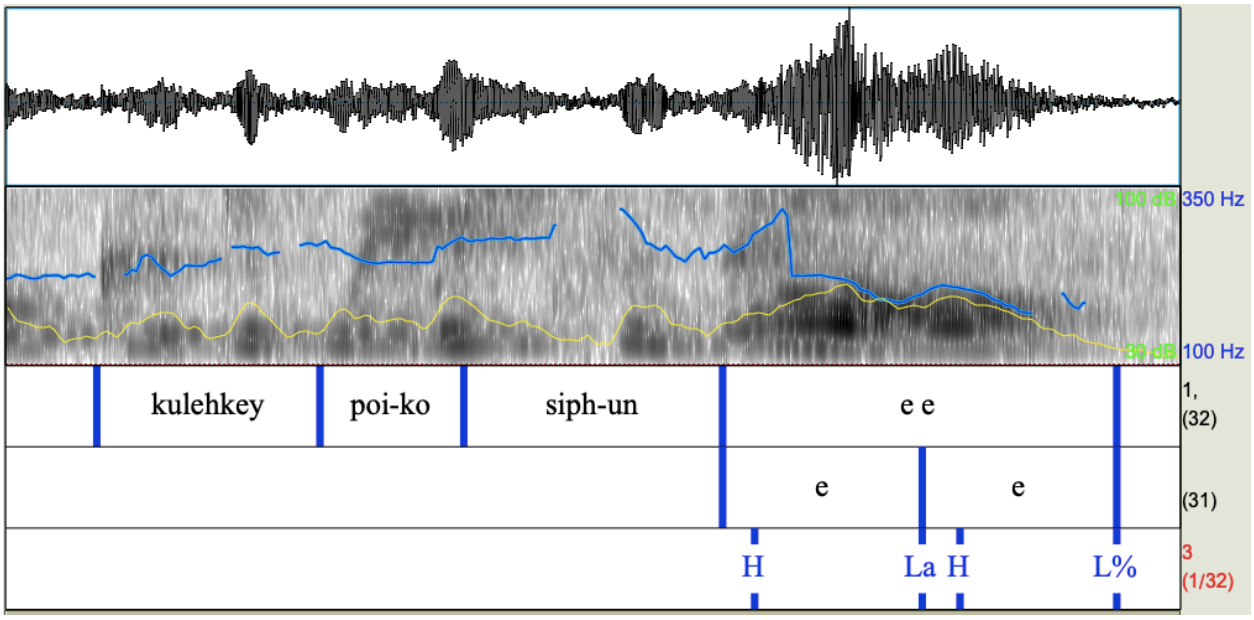


Figure 15. Praat information of the "*e e*" token (GA) and its prior turn (HS)

While the speaker's previous turn in line 11 gives the main idea of her disliking of the way people see her, the current turn only adds a detail without any significant new information. Therefore, the listener reduces the number of repeats of the token ("*e e*") differentiating it from the previous response ("*e e e e e*"). As the significance of the information decreases, the listener also manages her affective stance through the shorter token. In relation to the prior turn, the multiple saying token "*e e*" also shows a higher intensity. Similar to the previous token, the listener's interest is emphasized by the louder volume. This token as well co-occurs with two nods that also highlight the affective stance. Overall the multiple saying token "*e e*" shows similar prosodic qualities as "*e e e e e*" token, except for the number of repeats.

In sum, the listener shows different levels of engagingness to the speaker's turns by managing the token type and the number of repeats. This finding is congruent with that of Aoki (2008) and Yang (2013) that the multiple saying tokens display the listener's affiliation.

(2) Epistemic stance

The major functions of multiple saying tokens found in other languages are (1) to halt unnecessarily persisted actions (Stivers, 2004; Pyun and Yoon, 2022); (2) to display that he or she has received enough information and is ready to provide feedback (Nagata, 2004); (3) to display affiliation (Aoki, 2008; Yang, 2013); and (4) to display a mutual epistemic stance by claiming prior knowledge (Aoki, 2008). Among these, (2) and (4) are related to epistemic stance. (2) shows that the listener is now K+ and (4) is the case where the listener claims his or her epistemic independency. In this fragment, I found that the listener claims her upgraded epistemic stance as in (2). However, I argue that, unlike (2), the listener does not expresses that she is ready to "provide" feedback, rather, she is ready to "receive" more information from the speaker.

In this fragment, the speaker HS shares her own thoughts and feelings to which she has direct access. As demonstrated above, when the first token "*mm mm mm mm*" is employed, the listener already has predicted that the following telling will be negative. This proves that, similar to the single "*mm*" token that displays the listener is "better" informed than before, the multiple saying "*mm mm mm mm*" also shows that she has successfully grasped the information offered in the prior turn while she is still on a path between K– and K+. As the repetition intensifies the impact of its affective stance, it seems its epidemic stance is also highlighted. In other words, the multiple saying token emphasizes the fact that the listener's knowledge level has been upgraded. This does not mean that the listener claims a higher level on knowledge, rather, she is more certain about her developed epistemic status, therefore assures the speaker that she is ready to receive the next information.

The second token "*e e e e e*" has the same impact as the previous one. However, through choosing the lexical token "*e*" and the higher number of repeats, the listener even raises the degree of her informedness and underlines her readiness. Also, as mentioned above, the token co-occurs with five nods. I claim that the accompanying nods emphasize the listener's realization. Thereby the nods upgrade the listener's epistemic stance along with the lexical choice and the number of repeats.

The third multiple saying token produced in line 14 is "*e e*". Compared to the information given before "*e e e e e*" that provides the gist of the telling, the turn that "*e e*" is responding to only appends a minor detail to it. Therefore the degree of improvement in the listener's knowledge is relatively little. This small jump in her epistemic status is presented through the "*e e*" token that is still stronger than the single "*e*" token but it shows a decrease in the number of repeats.

To sum up, through the multiple saying tokens, the listener emphasizes the fact that she has received the prior information and, therefore, has become more knowledgeable than before while she is still on a path between K– and K+. The listener manages the degree of informedness and readiness to obtain further information through the lexical choice, the number of repeats, and the co-occurring nods.

4.4. Summary and Discussion

In this chapter, the "yes" type tokens "*ung*", "*e*", "*mm*", and multiple saying "*e*" and "*mm*" as well as nods have been discussed with respect to affective and epistemic stance they display.

First, I demonstrated the affective stance displayed by the "yes" type tokens. The lexical token "*ung*" and "*e*" with a falling intonation display mild interest towards the prior turn, while the "*e*" token shows a higher level of engagingness than "*ung*" token. Also the nods that co-occur with or follow the verbal token can be used to display the listener's confirmation of the validity of the information besides her interest.

The vocalization "*mm*" is similar to the "*ung*" token, but it is semantically and phonetically weaker. Therefore, it expresses even milder interest than does the "*ung*" token. Nods are divided into two types; mid-turn nods and turn-completion nods. The mid-turn nods display a more intense degree of interest because of the less obligation imposed on a listener in the midst of the turns. However, the nods at turn completion points express mere attentiveness since they are weaker than any other type of tokens used at the same position.

The multiple saying tokens show enhanced interest towards the prior turn in accordance with the upgrade of the noteworthiness of the information given. Also, these tokens are usually accompanied with nods and facial expressions that raise the level of interest. Listeners show different levels of engagingness to the speaker's turns by managing the token type and the number of repeats.

I also demonstrated the epistemic stance the responses display. First, though the "*ung*" and "*e*" tokens, the listener can display her epistemic authority as well as epistemic stance. The analysis reveals that the listener claims her higher epistemic authority than the speaker for the telling by employing the "*e*" tokens, while she does not do so with the "*ung*" tokens. The listener can also display that she is "better" informed than before while she is on a path between K– and K+ through the "*ung*" and "*e*" tokens. The same stance is indexed by the "*mm*" tokens. The "*e*" token can also show the listener's upgraded epistemic stance, but, unlike "oh/really"-type tokens, it does not display "surprise".

The epistemic stance of nods are also depending on the nod type. Through the mid-turn nods, the listener positions herself as K-, because she has not been informed with the information because the turn is still on the way. However, with the turn-completion nods, the listener claims her prior knowledge and displays her already earned K+ status treating the prior turn as not informative.

Through the multiple saying tokens, the listener emphasizes the fact that she has received the prior information and, therefore, has become more knowledgeable than before while she is still on a path between K– and K+. The listener manages the degree of informedness and readiness to obtain further information through the lexical choice, the number of repeats, and the co-occurring nods.

Before finishing this chapter, it will be necessary to mention the distinction between "*mm/ung*" and "*e*" tokens. In my data, "*mm*" and "*ung*" tokens are deployed when the listener is given rather limited information about the speaker's telling. Using the concepts of

Stivers' (2008), the listener is not offered "access" to the speaker's stance and the mono-syllabic lexical token "*ung*" can be compared to the vocal continuer "*mm*" that structurally aligns with the speaker's turn. Also, Oh and Park (2017) treat the "*ung*" tokens as a typical continuer acknowledging that the preceding turn is in progress.

On the contrary, the "*e*" tokens are employed after the speaker's turns containing more specific information about the telling. The placement is analogous to the environment in Stivers' (2008) study where the teller has provided insight into her own stance and the recipient gets access to it. The recipient's affiliation with the speaker's stance displayed through nods can be compared to the "*e*" token here. Meanwhile, Oh and Park (2017) interpret the function of the "*e*" token as the listener's appreciation to the noteworthy story. The table below summarizes the discussion thus far (Table 5.)

Table 5. Findings on "*mm*","*un*g", "e" tokens and nods

	Stivers (2008)	**Oh and Park (2017)**	**My study**
Token type	*"mm"*	*"ung"*	*"mm/ung"*
Environment	Where no access is provided	After non-noteworthy turns	Where limited information is provided
Displayed aspect	The preceding turn is in progress	The preceding turn is in progress	Affective stance: mild interest Epistemic stance: Listener is "better" informed than before while he or she is on a path between K– and K+.
Function	Structural alignment	Acknowledgment	

	Stivers (2008)	Oh and Park (2017)	My study
Token type	Nods	*"e"*	*"e"*
Environment	Where access is provided after teller stance is displayed	After assessment, noteworthy or emphasized telling	Where detailed information is provided
Displayed aspect	Recipient's endorsement to teller's stance	Recipient's appreciation to the turn-so-far	Affective stance: mild interest (higher than "*mm/ung*") Epistemic stance: Listener' upgraded epistemic stance or Listener's higher epistemic authority than speaker
Function	Social affiliation	Appreciation	

My data show highly similar findings as that of Stivers (2008) and Oh and Park (2017) in terms of the environment of the responses. While "*mm*" and "*ung*" tokens appear when the preceding turn or the story-so-far does not provide noteworthy information, nods and "*e*" token are deployed in response to the more significant telling of the speaker. The previous studies, however, found that only the latter is related to social or affective stance of the participants. Through the detailed analysis, I demonstrated that each type of response displays their own affective and epistemic stance.

CHAPTER 5: OH / REALLY TYPE RESPONSES

In this chapter, "oh" and "really" type reactive responses will be examined. These types of tokens are used in responses to "informings". The term "informing" was introduced by Heritage (1984) and used by Thompson et al. (2015) to cover news, informings, announcings, and reportings. The studies on responses to an informing have been discovered their distinct interactional functions and projected trajectories in talk. In the early stage, Jefferson (1981) and Heritage (1984) investigated English "*oh*" and provided an insight into its interactional actions. According to their work, free-standing "*oh*" and "*oh*"-plus-newsmark (*oh really*)" function as a backward-looking "information receipt" and "news receipt" respectively, which are implicative of sequence closure, while "*oh*"-plus-partial repeat functions as a forward-looking "newsmark" projecting further talk. They conclude that different newsmarks project different trajectories. Later on, Maynard (1997) argued that newsmarks can be both restrospective and prospective as receiving an announcement as news and promoting development of the news. Schegloff (2007) demonstrated that "*oh really*" can invite more substantial elaboration "providing for a curtailed expansion of telling on a mentioned matter" (p.158).

Responses to informings have been also considered to convey affective stance (surprise, unexpectedness, or disbelief) and epistemic stance (less knowledgeable to more knowledgeable) through their syntactic structures and prosodical features (Selting, 1987; Jefferson, 1988; Sacks, 1992; Gardner 1997, 2001; Wilkinson and Kitzinger, 2006, Thompson et al., 2015). In the early study, Gardener (1997) introduced the term "newsmarker" referring to responses that mark the prior turn as "newsworthy" in some way. The "newsworthiness" however, is a rather vague

concept to capture the listeners' affective or epistemic stance. Therefore, Thompson et al. (2015) who although keep the term "newsmarker", steer away from the question whether the recipients find the prior turn "newsworthy" or not. Instead, they explain that these responses display a cognitive–affective stance towards the informing through their marked phonetic-prosodic delivery. Particularly, they argue that the relation of the response to the prior informing in terms of pitch height/range, volume, and timing is relevant for conveying stances such as interest and surprise (prosodic upgrading) or empathy and sympathy (prosodic downgrading) (p. 69). In respect to epistemic positioning, they discovered that a rising intonation displays the responders are not yet fully K+, whereas a falling intonation expresses they are now K+. Additionally, some particles such as "*really*" and some prefabs (semi-fixed forms) such as "*oh really*" with a rising intonation are found to index the responder's intermediary stage on a path between K– and K+.

In this chapter, the most frequently occurring response tokens in informing sequences in Korean conversation will be discussed; "*a*", "*cincca*", "*kulay*", and *a*-prefaced structures "*a cincca*" and "*a kulay*". In the analysis, I will not avoid the term "newsworthiness", but will specifically examine how it is presented as affective and epistemic stance through prosodic and multimodal features. Also, I will use the term "informing" to refer to newly introduced information by a speaker. This study will, however, only concentrate on "storytelling" sequences which is equivalent to "volunteered informings" where responses are produced in the second-position (Thompson et al., 2015).[12] Therefore, question-elicited informings where the responses are produced in the third-position will not be covered.

[12] Thompson et al. (2015) explains the different sequential positions occupied by the two response types: a response to a volunteered informing is a second-position response, while a response to a question is third-position response.

5.1. Free-standing "*a*"

In this section, free-standing "*a*" will be discussed. While the token "*a*" has been mentioned in several studies as the Korean counterpart of the change-of-state "*oh*" in English (Heritage, 1984), it has not been deeply delved into. The importance of investigations on "*a*" was emphasized by Ha (2018) who found that sequence-closing third "*kuleh-kwuna*" expresses realization, point-like connection, spontaneous reaction, or surprise and questioned whether the frequently co-occurring "*a*" would function in a similar way.

Before examining "*a*", I will briefly look into the change-of-state "*oh*" in English. While Heritage (1984) claims that free-standing "*oh*" is used to mark the receipt of the informing delivered in the preceding turn and close the sequence, Thompson et al. (2015) subdivide the token by its prosody; falling *oh* and rising *oh* (free-standing "*oh*" with a falling intonation and that with a rising intonation, respectively). According to them, the former accepts the validity of the informing and declines to pursue further expansion, whereas the latter expresses doubt about the informing and makes a further turn from the informer relevant. They also argue that with prosodic upgrading and downgrading, free-standing "*oh*" can add an "affective lamination[13]". Having this as the background, I will investigate Korean "*a*" responses in informing sequences.

5.1.1. Rising-falling "*a*"

[13] 'I'm interested, pleasantly surprised, horrified, disappointed, saddened, etc. by the information.' (p. 80)

In fragment 4, the speaker CY explains the steps of completing a dissertation in his academic department. In the midst of the story, the listener HS produces a free-standing "*a*" token with a rising-falling intonation.

<Fragment 4: Dissertation>

```
1 CY:    우리는 보통
         wuli-nun pothong
         we-TOP   usually
         "In our department usually,"

2        그냥 졸업할 때까지
         kunyang colepha-l   ttay-kkaci
         just    graduate-RL when-by
         "Until graduating,"

3        계속 그 컨퍼런스랑 저널이랑
         kyeysok       ku  khenphelensu-lang cenel-ilang
         continuously  the conference-and    journal-and
         "Continuously, to conferences and journals"

4 HS:    (nod)(nod)

5 CY:    submit하고
         submit-ha-ko
         submit-do-CNN
         "(You) submit (the papers) and,"

6 HS:    응
         ung

         (…)[14]

12 CY:   이제 세번째 거 submit 하면은
         icey seypenccay kke   submit ha-myen-un
         now  third      thing submit do-if-TOP
         "If I submit the third one,"

13 HS:   응
         ung
```

[14] In line 7 - 11, the speaker CY tells about how many papers he is required to write and the listener HS regularly produces "*ung*" tokens.

```
14 CY:      dissertation은 이제
            dissertatio-nun icey
            dissertation-TOP now
            “The dissertation is,”

15          그동안 했던 거 이제 다 summarize 해가지고
            kutongan ha-yss-ten ke    icey ta  summarize hay-kaciko
            so:far   do-PST-RT  thing now  all summarize do-and
            “You summarize everything you have done and,”

16          책 같이 그냥 만들어 가지고
            chayk kathi   kunyang mantul-e kaciko
            book  such:as just    make-and
            “Just make it into a book.”

17 HS:   →  아::[:::::::::]
            a ::[:::::::::]

18 CY:         [그- 그런   ] 식으로    하는 건데
               [ku- kulen] sik-ulo ha-nun ke-ntey
               [like:that] way-by  do-RL  thing-CIRCUM
            “It goes like that.”
```

Before this fragment, HS gave an explanation of the procedure of submitting dissertation in her department. In this fragment, on the other hand, CY describes how the process goes in his department. In line 1 - 3 and 5, the speaker CY explains what the people in his department usually do until they graduate beginning the turn with “*wuli-nu*n (we + topic particle *nun*)” that topicalizes “his department”. After the first predicate “submit”, the listener simply produces an “*ung*” token. In line 14 - 16, how a dissertation actually is created is narrated. After this information, the listener deploys an elongated “*a*” token carrying a rising-falling intonation. In line 18, the speaker finishes the telling with the sequence-closing remark “it goes like that”. After this fragment, the speaker CY shifts the topic to “conferences” and goes on.

(1) Affective stance

The affective stances of the responses are shown in the prosodic features of the "*a*" token and the co-occurring nod. First, when compared to the prior turn, the token's volume is stronger, pitch is higher and pitch range is greater, as presented in Figure 16 and 17[15].

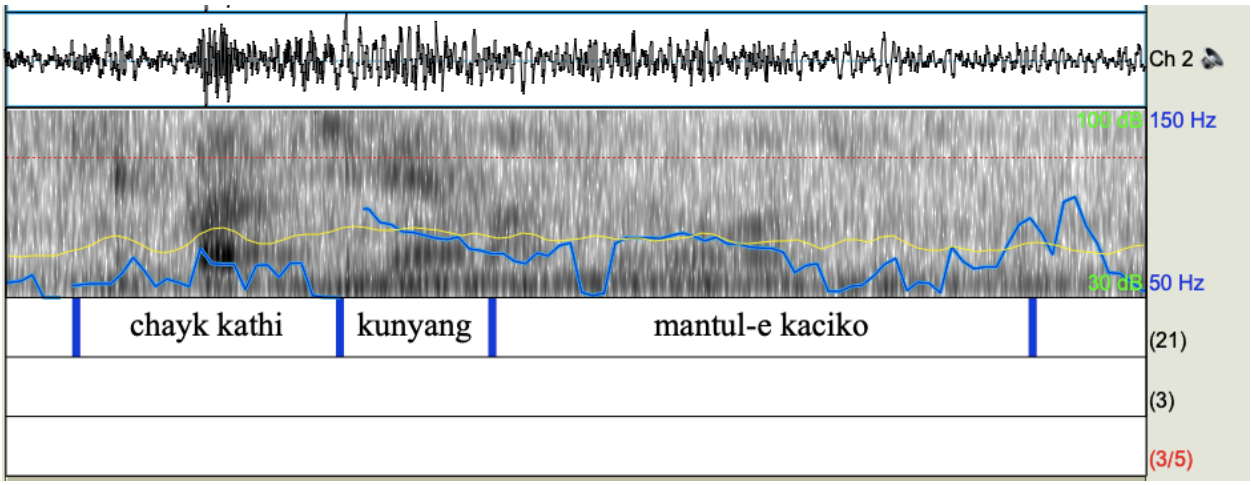


Figure 16. Praat information of the prior turn of the "*a*" token (CY)

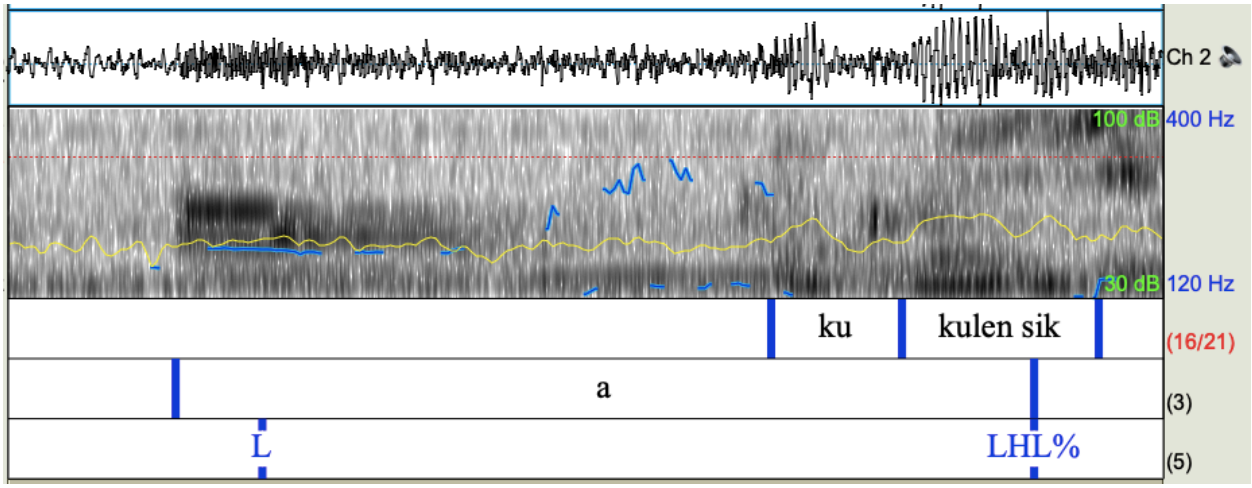


Figure 17. Praat information of the "*a*" token (HS)[16]

[15] The values used for CY voice range were: 50 - 150 Hz; those for HS's were 120 - 400 Hz.

[16] Since the last part of the "*a*" token overlaps with the speaker's next turn, the pitch contour is not clear. The original contour of the token would have shown a smoother fall at the end.

HS's "*a*" token draws a rising-falling intonation that is marked as the boundary tone LHL%. According to Park (2003), HL% tone expresses surprise (H) followed by realization (L) and LHL% tone is the same as HL% tone but has more emphasis on the realization (L). The realization is part of the listener's cognitive process and its discussion is more relevant to her epistemic stance. From the perspective of affective stance, however, I claim that the process of realization elicits the listener's interest. The listener first feels "surprised" (H) and her emotion shades off into "interested" (L) while registering the new information. Also the interest is emphasized by the first "L" tone in "LHL". In this fragment, however, the prior turn does not contain neither positive or negative connotation and therefore the listener's emotion is more of pure interest towards the "unexpectedness" of the newly received information.

In my data, the vocalization "*a*" token frequently appears with a rising-falling intonation at such an environment where the listener has been just given new information. It is inferred that, similar to English "*oh*" particle, the vocalization "*a*" itself indexes that the listener has been informed with new information. Therefore, when this "*a*" token is intoned with the rising-falling intonation that displays surprise and interest, the listener's stance seems to be emphasized.

The token shows an increased pitch height and pitch range compared to CY's informing and this can be considered a pitch upgrade (Curl, 2005; Ogden, 2006; Thompson et al., 2015). Also, it is also overly lengthened lasting for 1.6 seconds[17] to the extent the last 0.6 seconds overlap with the speaker's next turn. With the elongation, the listener takes up a longer time than usual highlighting her surprised feeling. Additionally, the listener produces a big slow nod while

[17] The average length of the mono-syllabic tokens is 0.45 seconds and that of all the "*a*" tokens is 0.55 seconds in my data.

producing the vocalization. The co-occurring nods also seem to reinforce the listener's emotional expression to the informing.

In sum, the upgraded rising-falling intoned "*a*" token displays the listener's surprised feeling through the high tone (H) and interest through the low tone (L) which is highlighted through the pitch upgrade, elongation, and big slow nod.

(2) Epistemic stance

The listener's epistemic stance is delivered through the token type, its prosodic features, and non-verbal reactions. First, the token type of the response is the vocalization "*a*" that signals the receipt of the informing in Korean conversation. Through this token, the listener displays that she has received the information and now she is K+.

Also, as mentioned above, the token's rising-falling contour is marked as the boundary tone LHL% denoting surprise (H) followed by realization (L) with extra emphasis on the realization (L) (Park, 2003). By displaying that she has adopted the informing through the falling part (L), the listener positions herself as K+. The positioning is reinforced by the lengthening of the token and signifies that the knowledge gap has been fully filled (Heritage, 1984).

This argument is supported by her non-verbal actions. With the beginning of the vocalization, the listener produces one big nod at a very slow pace along with the prolonged vocalization (1.6 sec). The slow nod seems to reflect the duration of the listener's cognitive process; registering new information. Additionally, when her head is lifted, she rolls her eyes upward as presented in Figure 19.

Figure 18. Seating arrangement of the participants; CY (left) and HS (right)

Figure 19. Listener HS's head and eye movement

The upward gaze is often considered a sign for searching or consulting memories, or contemplating (Jarque and Pascual, 2016). Gazing upward also occurs when the producer is adopting the previous participants' position (Bommel et al, 2013). In this case, the gaze can be interpreted as a signal that the listener is processing and adopting the informing.

To sum up, when the listener has been provided the complete information, she is positioning her epistemic stance as K+ through the vocalization "*a*" with a rising-falling intonation and the big nod with the upward gaze.

5.1.2. Falling "*a*"

In fragment 5, the speaker HS narrates the factual reality of her academic field. The listener CY uses free-standing "*a*" token with a falling intonation twice in the talk.

<Fragment 5: Reality>

```
1 HS:      또 막 field가 우리는 field가 되게 쫌 좁은 편이거든
           tto  mak  field-ka  wuli-nun field-ka  toykey ccom
           cop-un     phyen-iketun
           also just field-NOM we-TOP   field-NOM very   little
           narrow-RL  tendency-CORL
           "Our field is relatively small."

2 CY:      (nod)(nod)

3 HS:      그래서 이제 우리 교수는 field에서                  [잘 알려진
           분이라
           kulayse icey wuli kyoswu-nun    field-eyse  [cal  allyeci-n
           pwun-ila
           so      now  our  professor-TOP field-LOC  [well known-RL
           person.hon-because
           "Also, my professor is very well-known in the field."
                                                       [(nod)(nod)
4 CY:      (nod)(nod)

5 HS:      이 교수한테 약간 밉보이면은
           i    kyoswu-hanthey yakkan mippoi-mye-nun
           this professor-by   little be:hated-if-TOP
           "If you get on the wrong side of her,"

6 CY:  →   (nod)(nod) 아
           (nod)(nod) a
```

```
7 HS:       이제 내가 이 연구실을 나가더라도
            icey nay-ka i    yenkwusil-ul       naka-telato
            now  I-NOM  this research:lap-ACC leave-even:if
            “Even if I leave this research lab,”

8 CY:       (nod)(nod)

9 HS:       다른 데 가면은 이제
            talun     tey   kamyenun icey
            different place go-if    now
            “When I join a different lab,”

10          [평판이 안 좋아질                          [수가 있는
            그런 위험이 또 있고
            [phyengphan-i   an  cohaci-l       [swu-ka  iss-nun
            kulen      wihem-i    tto  iss-ko
            [reputation-NOM not become:good—RL [way-NOM exist-RL
            like:that  danger-NOM also exist-CNN
            “I could have a bad reputation.”

11 CY:  →   [(nod)(nod)                        [아 (nod)(nod)(nod)
                                               [a

12          (nod)(nod) 그렇겠네
                       kuleh-keyss-ney
                       be:so-CONJ-APP
            “It must be so.”
```

The speaker HS initiates the story with the background information that her academic field is relatively small and her adviser is well-known in the field (line 1 and 3). The speaker's two first turns are responded to by the listener's nods. In line 5, the speaker drops a hypothetical condition, "if you get on the wrong side of her" and in response to this turn, the listener produces the first falling intoned "*a*" token after two nods. In line 7, 9, and 10, the speaker HS depicts a possible situation where she could get a bad reputation even if she leaves her lab. As a response to this turn, CY deploys another "*a*" token with a falling intonation followed by several nods. In line 13, he adds a comment "I see (it must be so)" projecting a sequence closure.

(1) Affective stance

The listener uses the falling intoned "*a*" token twice during the storytelling. The affective stance of the first "*a*" tokens is mainly conveyed through their prosodic features and they are presented in Figure 20 and 21[18].

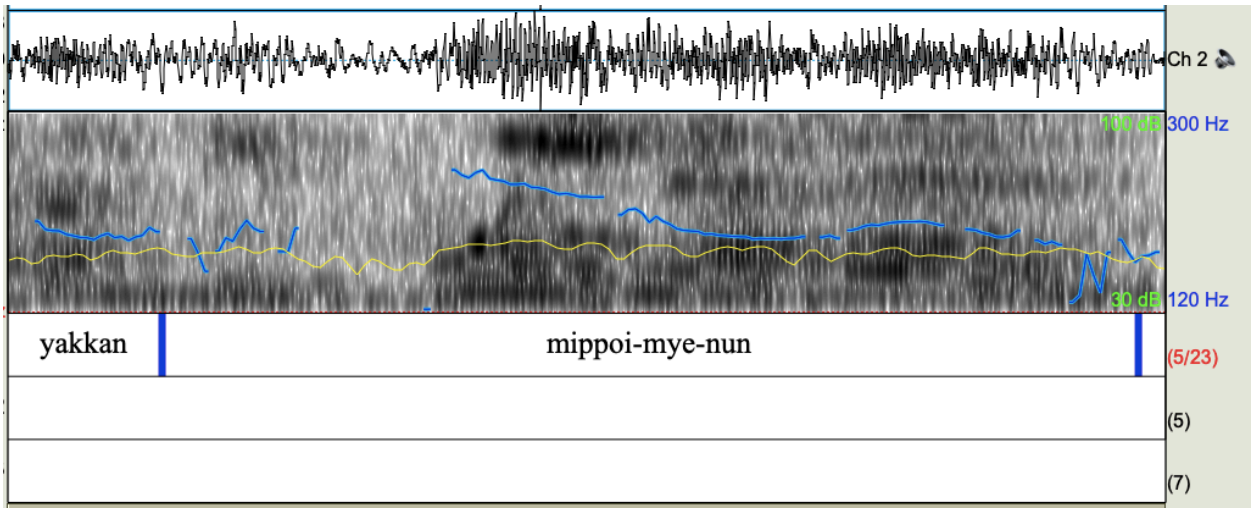


Figure 20. Praat information of the prior turn of the the "*a*" token (HS)

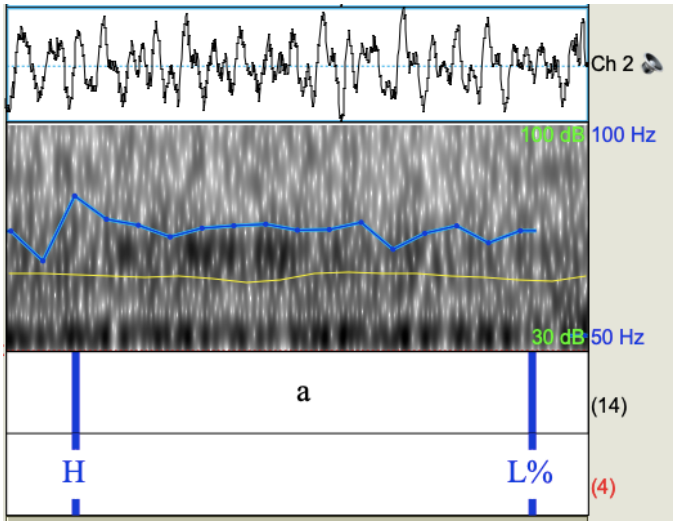


Figure 21. Praat information of the first "*a*" token (CY)

[18] The values used for HS voice range were: 120 - 300 Hz; those for CY's were 50 - 100 Hz.

The first "*a*" token appears after the speaker has offered a substantial hint that strongly projects a negative consequence by using the word "*mippoi-* (be hated)" in line 5. Although, at this point, the listener has not been fully informed with what might happen as a result of the hypothetical condition, he produces the vocalization "*a*" indexing his realization to some extent. However, instead of a rising-falling intonation that can intensify the listener's surprised emotional status (see 5.1.1.), the token carries a slightly falling intonation. Also, in relation to their prior turn, the tokens' volume is softer and pitch is lower. CY's pitch contour on "*a*" is lower than the overall pitch configuration of HS's informing. The decrease in pitch height on the "*a*" token compared to the informing can be described as a pitch downgrade (Curl, 2005; Ogden, 2006; Thompson et al., 2015). Through the downgraded falling "*a*" token, the listener seems to reserve excessive emotional expressions, but displays his empathy and sympathy towards the telling considering the seriousness of the situation.

The second "*a*" token occurs after the speaker has conveyed the possible consequence she might end up with. The token is deployed as soon as the speaker has said "*phyengphan-i an cohaci-l* (reputation becomes bad)". The information provided up to this point seems to be sufficient for the listener to understand the situation and his realization is shown through the "*a*" token. The second "*a*" token also shows pitch downgrade drawing the same pitch contour as the first one but it is followed by several nods[19]. Through the downgraded falling "*a*" token, the listener shows that the listener still treats the story as a serious and non-pleasant matter. Also the

[19] Because this token completely overlaps with the speaker's turn and the two voices were merged in one recording, the prosodic information of the token was unavailable to see on Praat.

nods following the token intensify his sympathetic attitude signifying that he understands the seriousness of the possible consequence.

The listener's affective stance is supported by his topic-closing utterance "*kuleh-keyss-ney* (it much be so)" that is best translated into "I see". "*Kuleh-keyss-ney*" consists of the anaphoric adjective "*kuleh-* (be so)", the conjecture expression "*-keyss*" (might/must be), and the suffix "*-ney*" (Sohn, 1999). The suffix "*-ney*" is used to express empathy (Strauss, 2005) and therefore, *ney*-marked utterances often show convergent alignment (Du Bois, 2007; Ha, 2018). The listener expresses that he can imagine the situation the speaker is facing and affiliates with her stance through "*kuleh-*" indexing the preceding turn, "*-keyss*" showing his conjecture, and "*-ney*" exhibiting his empathy. The utterance also shows similar prosodic features as the two "*a*" tokens shown above (Figure 22 and 23[20]).

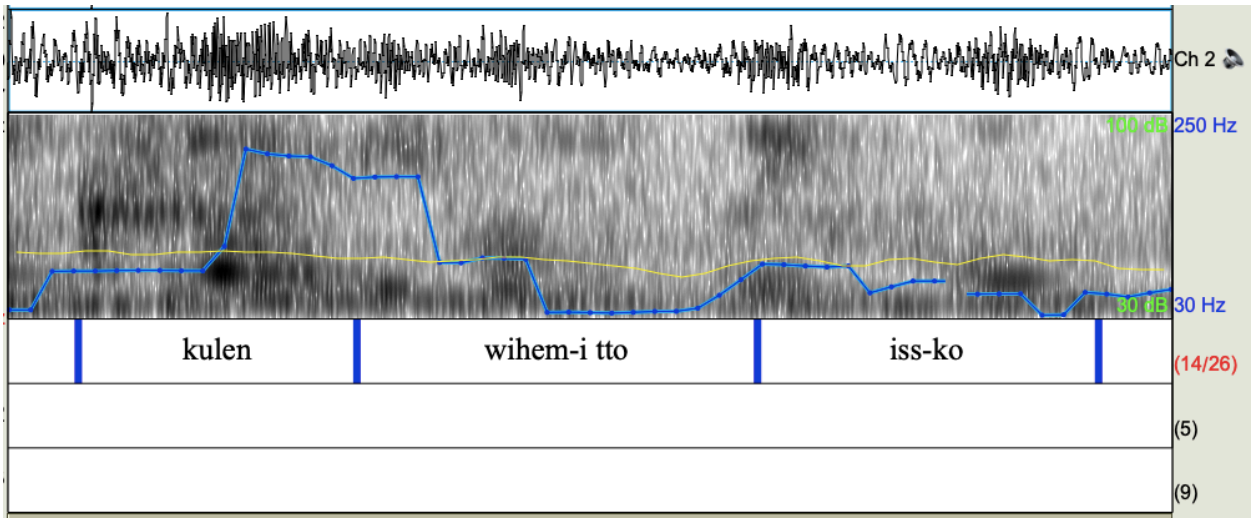


Figure 22. Praat information of the prior turn of "*kuleh-keyss-ne*" (HS)

[20] The values used for HS voice range were: 30 - 250 Hz; those for CY's were 40 - 120 Hz.

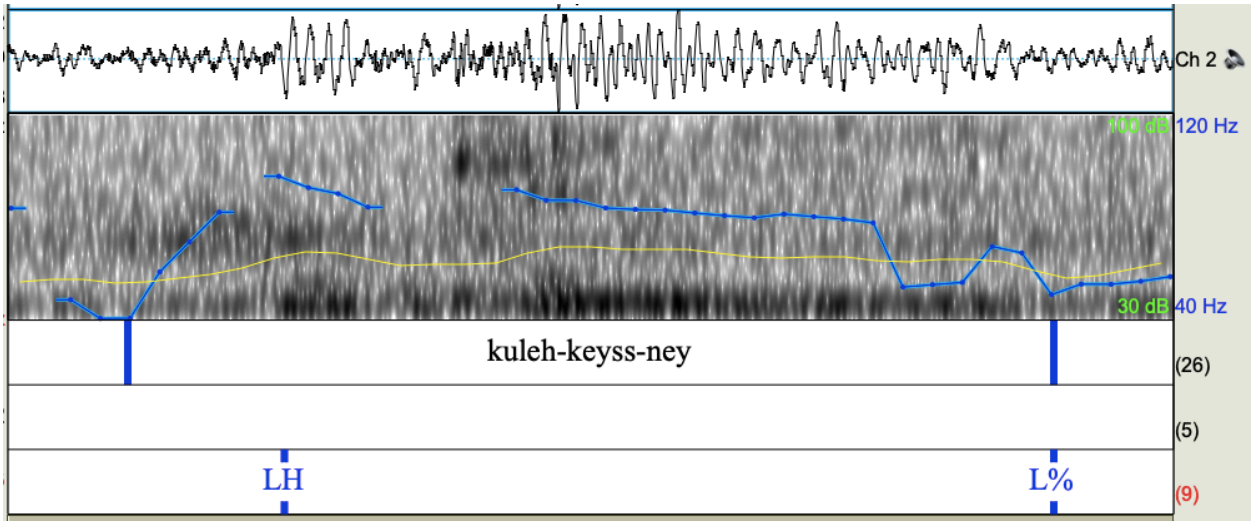


Figure 23. Praat information of "*kuleh-keyss-ne*" (CY)

The utterance shows a falling intonation as well as a pitch downgrade with weaker volume, and lower pitch compared to the prior turn. With its lexical meaning and the prosodic display, the downgraded "*kuleh-keyss-ney*" strongly reinforces the empathetic stance of the "*a*" tokens and their affiliation with the speaker's stance.

In sum, through the falling "*a*" token, the listener displays his empathy and sympathy towards the telling considering the seriousness of the situation. Also, the pitch downgrade and the following nods intensify his sympathetic attitude.

(2) Epistemic stance

The listener's epistemic stance is delivered through the token type, its intonation, and the following nods. First, both of the tokens shown in this fragment take the form of the vocalization "*a*" displaying that he has become K+ after having been informed with the informing.

Second, they both carry a falling intonation. According to Park (2003), in her study of Korean boundary tones, found that final L% tone exhibits its producer's realization. Thompson et al. (2015) also claim that the listener indexes he or she is now K+ through a final falling intonation. In this fragment, the vocalization "*a*" with a falling intonation marked as boundary tone L% indeed marks the listener's cognitive process that he registers new information and accepts its validity. Thereby he positions his epistemic stance as K+.

Third, the several nods following the second "*a*" token emphasizes the listener's realization. While the first informing to which the first "*a*" token responds is a "hypothetical" condition, although a negative consequence is highly predictable through it, the second informing "vividly" illustrates the possible result. In accordance with the enhanced information, the listener produces the "*a*" token with the multiple nods as if he is manifesting "now, I can even understand better". Therefore, the accompanying nods function to upgrade his already K+ stance even higher.

In sum, the listener positions himself as K+ through the vocalization "*a*" and the falling intonation, and this epistemic stance can be intensified by the following nods.

5.2. Lexical token "*cincca*" and "*kulay*"

In this section, "*cincca*" and "*kulay*" tokens in response to informings will be discussed. First, "*cincca*" is a lexical word meaning "real" whose part of speech is noun. The word is used in interrogative structures such as "*cincca-lo*?" ("for real?") or "*cincca-ya/yeyyo*?" ("is that real?"). In colloquial conversations, however, the interrogative sentences are abbreviated into

"*cincca*" and the word is used independently unless the polite suffix "*-yo*" is required. "*Cincca*" can be considered the Korean counterpart of the "*really*" type lexical tokens in other languages; English "*really*", Japanese "*honto*", German "*echt*", etc. (Heritage, 1984, More 1999, Wilkinson and Kitzinger, 2006; Gubina and Betz, 2021). Similar to the studies on English "*oh*", those on "*really*" also focus on its discourse function (newsmark), affective and epistemic stance it conveys, and how it develops the further sequence. Heritage (1984) maintains a distinction between free-standing "*oh*" and newsmarks, such as "*really*", in that the former functions as a backward-looking "information" receipt while the latter servers a forward-looking "news" receipt. The affective stance of this type of tokens have been described as "assertions of ritualized disbelief" referring to expressing "surprise". Sacks (1992) claims that expressions of surprise typically occur after news announcements and prefigure expressions of sympathy. Therefore, the display of surprise "reserves rights to future expression of emotion". Thompson et al. (2015) found that "*really*" with a final rising intonation marks "doubt" about the veracity of the informing. It also positions its epistemic status as an intermediary stage on a path between K– and K+ and prompts the informer to substantiate or support the informing. Gubina and Betz (2021), in their study of German "*echt*", studied how the token shape the sequential development and found that it functions as inviting a reconfirmation or a topic elaboration, and soliciting an account. They also explain that the recipient's epistemic positioning can change the function and have different interactional consequences. For instance, the higher epistemic access the recipient is claiming, the more it may sound as a pre-challenge.

On the other hand, "*kulay*" is a syntactically complete sentence composed of the anaphoric adjective stem "*kuleh-* (to be so)" and the intimate sentence ending "*-a*". It means

"that is so" in declarative contexts and "is that so" in interrogative contexts. While "*cincca*" is semantically equivalent to the "*really*" type lexical tokens in other languages, "*kulay*" can be classified as a minimal clausal response such as "*did you*" in English. In English, the clause can be formatted in two ways, as interrogative sentences (e.g., "*did you*"), or as declarative sentences, (e.g., "*you did*"). Thompson et al. (2015) demonstrate the discourse functions of the various minimal clausal responses formatted with either interrogative or declarative form along with two different intonation contours (rising and falling). According to them, while rising interrogative responses ask a verification of the information, rising declarative responses request a confirmation of the informing implying that it is contrary to their expectation. Also falling-intoned responses in both syntactic structures display that the recipient is not encouraging a further talk. In Korean, however, the syntactic structure of minimal clauses do not distinguish interrogative or declarative meaning. Rather, the meaning is mainly conveyed through their intonation; a rising intonation signifies "interrogative", while a falling intonation denotes "declarative". In my data, "*kulay*" tokens are deployed with two different intonation contours; rising intonation and rising-falling intonation. In this section, I will demonstrate how the listeners manage the details to engender a preferred response while displaying their own stance.

In fragment 6, the speaker PT narrates three examples about people around him and his fiancé who expected their marriage at the very beginning stage of their dating. The listener, who is surprised by the fact, produces several "*cincca*" and "*kulay*" tokens during the telling.

<Fragment 6: Prediction>

1 PT: 저랑 창엽이 형은 카운터에 뭐 디저트
시키러 카운터에 [갔는데
ce-lang changyepi hyeng-un khawunthe-ey mwe ticethu
sikhi-le khawunthe-ey [ka-ss-nuntey
I-with Changyep brother-TOP counter-LOC what dessert
order-in:order:to counter-LOC [go-PST-CIRCUM
"When Changyep and I went to the counter to order dessert"
(HB) [(nod) (nod)

2 현수한테
hyenswu-hanthey
Hyenswu-to
"To Hyenswu"

3 어 딱 처음 봤을 때 둘이 결혼할 것 같다 뭐 이런 말을 했대요
e ttak cheum pw-ass-ul ttay twul-i kyelhonha-l
kes kath-ta mwe ile-n mal-ul ha-yss-tay-yo
e right first see-PST-RL when two-NOM marry-RL
seem-QT what like:this-RL word-ACC do-PST-QT-POL
"(She) said, when she saw us for the first time she knew
we two were going to get married"

4 HB: → 진짜?
cincca?

5 PT: 에
ey
yes
"Yes"

6 HB: [(big slow nod)

7 PT: [그리고 막
[kuliko mak
[and just
"And like"

8 제 친구들도 막
cey chinkwu-tul-to mak
my friend-PL-also just
"My friends also say"

9 HB: (nod)[(nod)(nod)

10 [되게 결혼할 것 같은데 약간 이런
[toykey kyelhonha-l kes kath-untey yakkan ile-n

[very matty-RL seem-CIRCUM little like:this-RL
“(We) will definitely get married, like this”

11 HB: → 어 진짜:
e cincca:

12 PT: 네 그리고 막
ney kuliko mak
yes and just
“Yes, and just”

13 HB (nod)(nod)(nod)[(nod)(nod)

14 PT: [현수(0.3)가 같이 일하는 학생이 있는데
[hyenswu(0.3)-ka kathi ilha-nun haksayng-i iss-nuntey
[Hyenswu-NOM together work-RL student-NOM exist-CIRCUM
“There is a student Hyenswu is working with”

15 HB: 어
e

16 PT: 사귄 지 며칠 안 돼 가지고 [바로
sakwi-n ci myechil an tw-ay kaciko [palo
date-RL have:been a:few:days not become-because [right:away
“Only after a few days we started dating”
(HB) [(nod)(nod)

17 혹시 만나시는 분 있으세요 막
이러- 이러다가 [나중에
hoksi manna-si-nun pwun iss-use-yyo mak
ile- ile-taka [nacwungey
bu:any:chance meet-HON-RL person.hon exist-HON-PL just
this- like:this-TRNS [later
“Are you seeing someone? She says”
(HB) [(nod)(nod)

18 HB: (nod) (nod)

19 PT: 진지하게 만나시는 거 겠네요 막 이렇게 하다가
cinciha-key manna-si-nun ke keyss-ney-yo mak ileh-key ha-taka
serious-ADV meet-HON-RL thing may-APP-POL just like:this do-TRNS
“It must be something serious, also she says”

20 결혼 (0.2) [약간 되게
kyelhon (0.2) [yakkan toykey

```
            marriage      [little very
            “Marriage, a little bit”
  (HB)                    [(nod)(nod)

21          [주위에서 예상 조금 했던 거 같아요
            [cwuwi-eyse yeysang cokum ha-yss-ten ke katha-yo
            [around-LOC prediction little do-PST-RT seem-POL
            “People around us seem to have expected it.”
  (HB)      [(nod)(nod)(nod)(nod)

22 HB:  →   그래?
            kulay?

23 PT:      타이밍은 예상 못 했을 수도 있는데
            thaiming-un yeysang    mos    ha-yss-ul swu-to
            iss-nuntey
            timing-TOP  prediction cannot do-PST-RL way-also
            exist-but
            “(They) might not have predicted the timing but”

24 HB:      [어어어
            [e e e
            [(nod)(nod)

25 PT:      주위에서
            cwuwi-eyse
            around-LOC
            “Around us.”

26 HB:      음
            mm
```

Before this fragment, the listener HB told the speaker PT that she was surprised by PT and his fiancé’s engagement yet it was not unexpected. The speaker PT said “there were several people who had a good hunch” before he started to give the actual examples. In line 1 - 3, the speaker PT drops the first example that one of his colleagues told his fiancé that she had expected the two to get married. As a response, the listener HB deploys the first “*cincca*” token with a rising contour while opening her eyes widely. The speaker briefly responds to this token with “*ney* (yes)” in line 5, and after this, the listener gives one big nod at a slow pace. In line 8 and 10, the

speaker PT provides the second example that his friends also told him that they had predicted his marriage. This time the listener produces another "*cincca*" token drawing a rising-falling intonation preceded by an "*e*" token. In line 12, the speaker PT again responds to the token with *ney* (yes)" and prefaces the third example saying "and like". The listener reacts to the preface with nods. Through line 14 - 19, the speaker delivers another marriage prediction story with his fiancé's colleague. In line 20 - 21, he closes his turn by offering the topic sentence of the three examples; "people around us seem to have expected it". The listener HB reacts to this closing comment with a "*kulay*" token carrying a rising intonation. This time, the speaker does not respond to the token but goes on and add his opinion that the people might not have predicted the actual timing. The listener responds to this information, which is somewhat opposing to the topic of the telling, with a multiple saying token "*e e e*". In line 25, the speaker repeat a part of his preceding turn "*cwuwi-eyse* (around us)" and this turn is responded to with a vocalization "*mm*".

(1) Affective stance

In this fragment, two "*cincca*" tokens and and one "*kulay*" token are shown. The affective stances of these responses are displayed through the prosodic features of the verbal token and the non-verbal behavior. The prosodic features of the first "*cincca*" token is shown in Figure 24 and 25[21].

[21] The values used for PT voice range were: 50 - 250 Hz; those for HB's were 70 - 500 Hz.

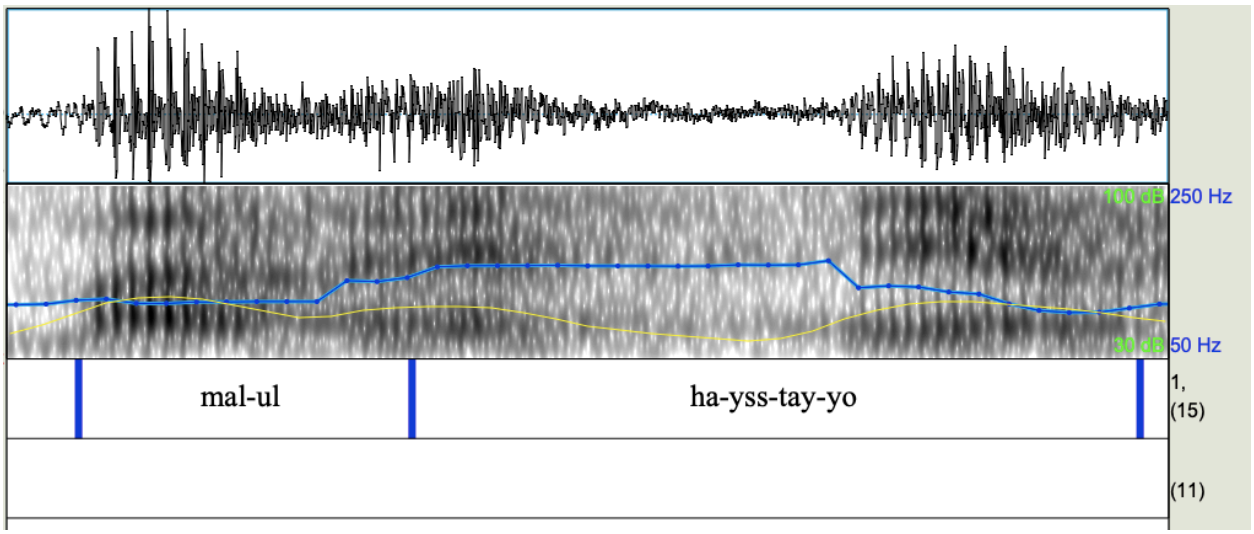


Figure 24. Praat information of the prior turn of the first "*cincca*" (PT)

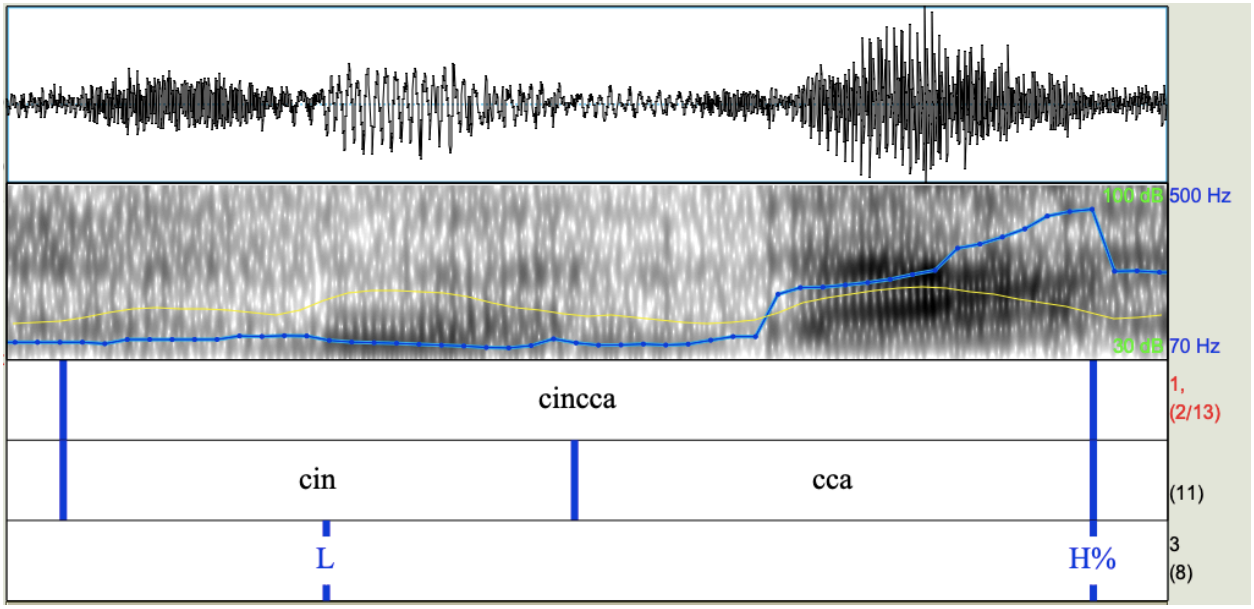


Figure 25. Praat information of the first "*cincca*" (HB)

The first "*cincca*" token is deployed when the speaker has introduced the first example of how people around them anticipated their marriage. The token draws a rising intonation that is marked as the boundary tone H% that expresses surprise (Park, 2003). Also, the "*cincca*" token shows a pitch upgrade compared to PT's informing through the increased pitch configuration. In

response to the speaker's first example, the listener is expressing the low degree of certainty and surprise through the final rising "*cincca*" token intensified by the pitch upgrade.

Also the listener's non-linguistic behaviors support her emotional display. While producing the rising "*cincca*" token, she raises her eyebrows. Right after the response, she takes her gaze to the right side and plucks her lips making a suspicious look on her face, and slowly nods once.

Figure 26. Seating arrangement of the participants; HB (left) and PT (right)

(before) → [*cincca?*] → (before nod) → (nod)

Figure 27. Listener HB's nods and facial expressions

Looking sideways and plucking lips show the person is having a doubt (Jarque and Pascual, 2016). Although after the verbal response she slowly nods once signaling she is accepting the validity of the received information, the suspicious look is still on her face. The listener HB's facial expression displays that she finds the informing hardly believable.

The second "*cincca*" token shows different prosodic features (Figure 28 and 29[22]).

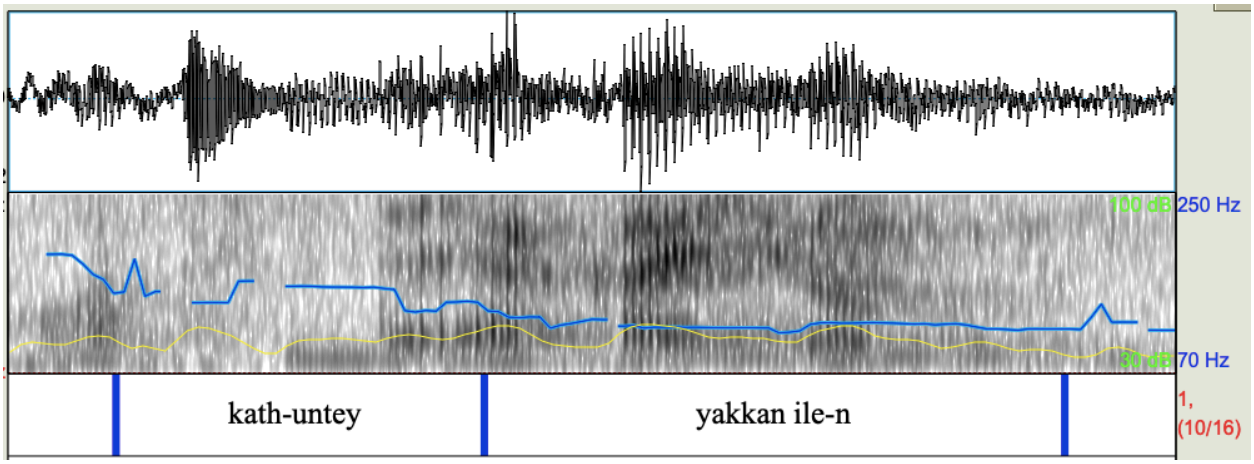


Figure 28. Praat information of the prior turn of the second "*cincca*" (PT)

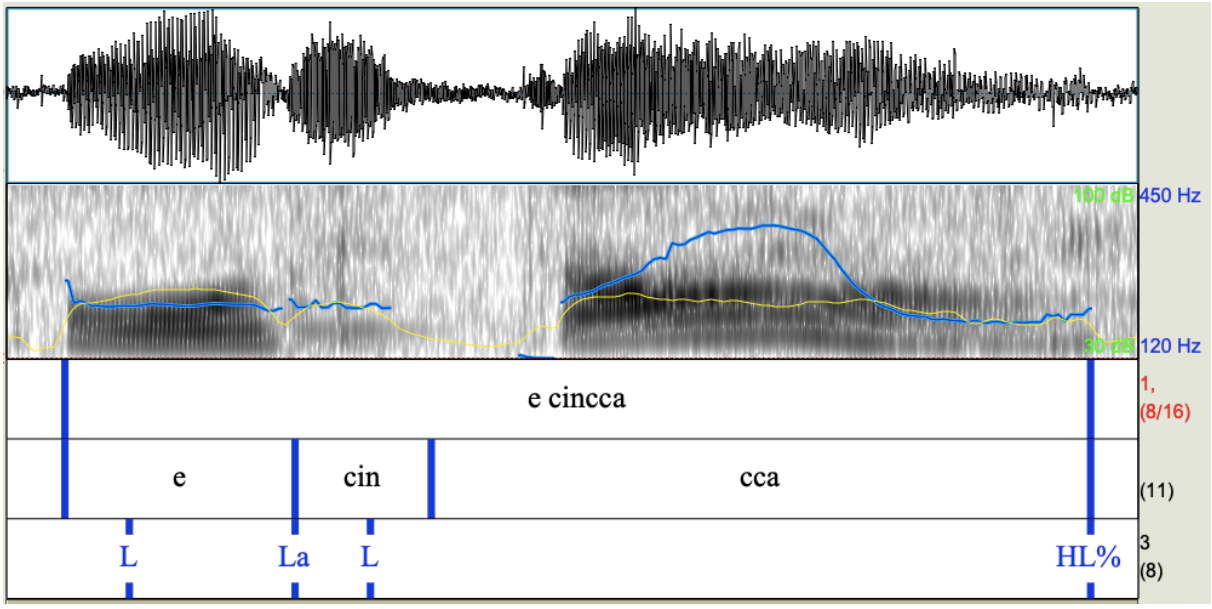


Figure 29. Praat information of the second "*cincca*" (HB)

[22] The values used for PT voice range were: 70 - 250 Hz; those for HB's were 120 - 450 Hz.

The second response appears after the second example about the speaker's friends who also predicted his marriage. The "*cincca*" token is preceded by an "*e*" token and carries a rising-falling intonation contour that is marked as the boundary tone HL%. Additionally, the "*cincca*" token also shows a pitch upgrade in relation to PT's informing. As mentioned earlier, HL% tone expresses surprise (H) followed by realization (L) (Park, 2003) and I suggested that the process of realization engenders the listener's interest. After hearing the second example of the speaker's experience, the listener still finds it surprising (H), but this time it is interesting (L) at the same time. Her affective stance is again emphasized by the pitch upgrade.

This time the "*cincca*" token also is preceded by an "*e*" token. When used independently, the "*e*" token with a falling intonation displays mild interest at the turn completion points (4.1.). Although the "*e*" token in this response draws a rather plateau contour than a falling one, the overall low tone seems to serve the same function here[23]. Therefore, by adding the "*e*" token before the "*cincca*" token, the listener's "being interested" stance is highlighted.

The "*kulay*" token, similar to the first "*cincca*" token, shows a final rising intonation marking the boundary H% tone and an overall pitch upgrade in relation to the prior informing (Figure 30 and 31[24]).

[23] In my data, "*e*" token with a plateau contour is not found when used independently. Further research on this should be conducted in the future.

[24] [24] The values used for PT voice range were: 40 - 200 Hz; those for HB's were 120 - 450 Hz.

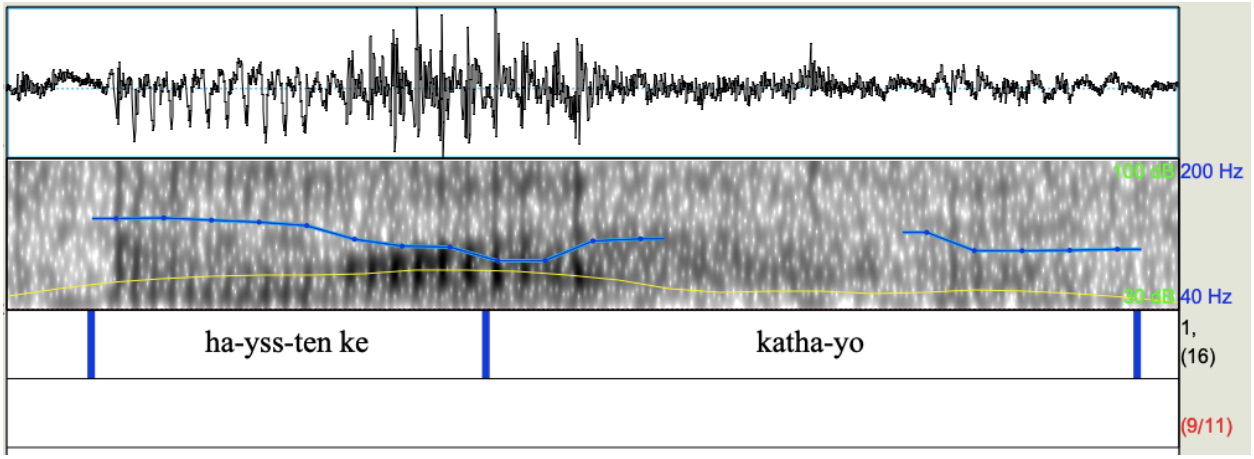


Figure 30. Praat information of the prior turn of the second "*kulay*" (PT)

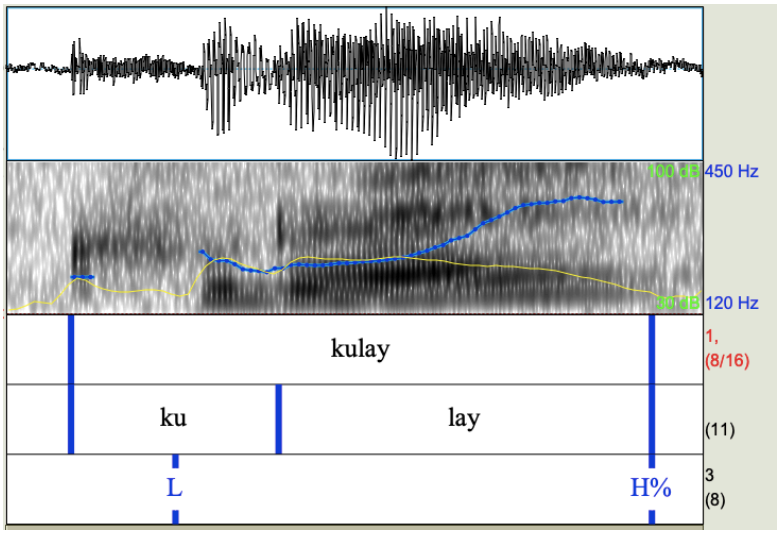


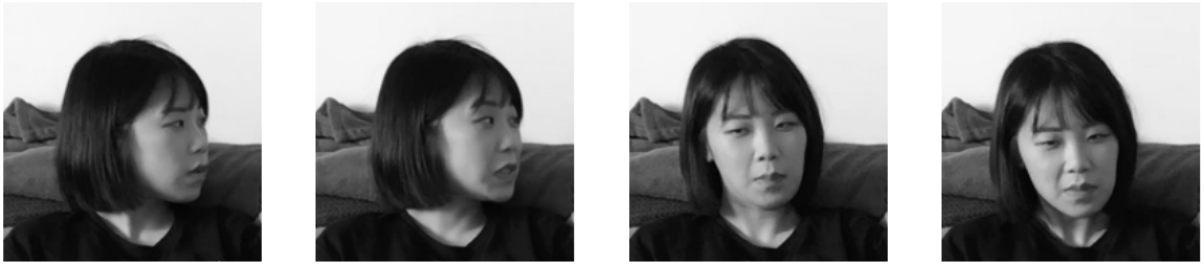
Figure 31. Praat information of the second "*kulay*" (HB)

The "*kulay*" token with a rising intonation is placed after the speaker's topic summarizing remark, "people around us seem to have expected it", which insinuates the completion of the entire telling. Even though the topic itself is not anymore "new" to the listener after hearing the three similar examples, she expresses her astonishment (H) to the entire story as higher commitment is preferred at this point. Therefore, the "*kulay*" token in this fragment displays enhanced surprise towards the telling, rather than suspicion or doubt.

In sum, the basic affective stances displayed through the lexical tokens "*cincca*" and "*kulay*" are surprise and suspicion. The non-linguistic behaviors and prosodic features can either enhance the emotional expressions or add another layer such as showing interest.

(2) Epistemic stance

The listener's epistemic stance is also displayed through the token type, prosodic features, and non-verbal behaviors. The first response consists of a single "*cincca*" token with a rising intonation. According to Thompson et al (2015), a rising intonation displays the responders are not yet fully K+, and particularly with some particles such as "*really*", it indexes the responder's intermediary stage on a path between K– and K+. In this fragment as well, until the point where the listener has been provided only one example of the speaker's experience, she has not been fully convinced, that is, not yet K+. Therefore, the listener produces a rising-intoned "*cincca*" token. The claim that she is not fully K+ at this point is supported by her facial expression as shown in Figure 32.



(before) → [*cincca?*] → (before nod) → (nod)

Figure 32. Listener HB's nods and facial expressions

The suspicious look on her face reveals that the first example alone is not sufficient to remove her doubt. However, through the following slow nods, the listener displays that she is at least accepting, or trying to accept, the validity of the information. Also, as long as a piece of information is given, the listener is more informed than before. Therefore, the listener expresses that she is not fully K+ but she is on a path between K– and K+ through the "*cincca*" token with a rising intonation, her facial expression, and the following nods.

The second "*cincca*" token, on the other hand, carries a rising-falling intonation. Park (2003) explains that a rising-falling contour marked as the boundary tone HL% denotes surprise (H) followed by realization (L). Thompson et al. (2015) also claim that a final falling intonation indicates the listener is now K+. As for the second "*cincca*" token, its intonation marks that the listener is now registering the information after the second example. The preceding "*e*" token also solidifies this epistemic positioning. The "*e*" token independently can display that the listener is better informed than before. Therefore, this response exhibits the listener's knowledge level has become upgraded and now she is more convinced.

The epistemic positioning is also displayed by the listener's nods and gaze. As explained earlier, nods can be used as a tool to display the listener's cognitive process when she is registering new information and accept its validity. Additionally, when producing nods, the listener is gazing downward as shown in Figure 33.

[*e cincca*] (before nod) (nod)

Figure 33. Listener HB's nods and eye movement

According to Navarro (2018), downward gazing is one of the eye movements that happen when a person is processing a thought or an emotion. The listener's gaze direction and nods tell that she is processing the information given while adopting the validity of the telling at the same time.

The third response consists of a "*kulay*" token with a rising intonation. Before going in to it, I will retrieve the discussion of English minimal clausal responses. According to Thompson et al. (2015), English minimal clausal responses show different functions depending on their syntax; with a rising intonation, interrogative responses (e.g., "*did you?*") ask a verification of the information, while declarative responses (e.g., "*you did?*") request a confirmation of the informing implying that it is contrary to their expectation. While with interrogative responses, the listeners mark relatively low epistemic stance, through declarative responses, they bring some expectation to the sequence and claim to have prior knowledge. This is attributed to the fact that English speakers express a K+ position through declarative syntax (Thompson et al., 2015).

In Korean, on the other hand, the distinction is not made syntactically because of the same word order of interrogative and declarative forms. Therefore, the different epistemic positioning displayed through syntax is not visible in Korean. Although the two types can be differentiated by intonation, usually shaping interrogative sentences with a rising intonation and declarative sentences with a falling intonation, the interpretation can be different depending on the context. Therefore each contingency should be closely examined to understand their actual meaning.

In this fragment, since the "*kulay*" token is delivered in a rising intonation, it can be seen that the listener denotes a low level of knowledge while delivering the interrogative meaning "is that so?". Considering the location, where the listener has been informed with the three consecutive examples, the listener is not claiming her own knowledge or finding it contrary to her expectation. Instead, she marks that she is still not yet fully K+, and still on a path between K- and K+, even after the three informings.

To sum up, through "*cincca*" and "*kulay*" tokens, the listener displays her epistemic status. With a rising intonation, she positions herself as not yet K+ but somewhere between K- and K+. On the other hand, through a rising-falling intonation, the listener shows she is now registering the information and more informed than before.

5.3. *a*-prefaced Structure "*a cincca*" and "*a kulay*"

The "*a*" + "*cincca*" and "*a*" + "*kulay*" structure consists of two components; prefacing "*a*" and a lexical token. The two components frequently co-occur as a response in Korean

conversations and this type of configuration is shown in other languages too. As for the treatment of structure, scholars have had different views. Schegloff (2007) uses the terms "composite" or "two-component" responsive turns referring to turns consisting of two particles and explains each component has its own prosodic contour doing its own action. On the other hand, Thompson et al. (2015) treat the "particle combinations" (e.g., "*oh really*?") as having one intonation contour performing a single action. In this section, I treat these structures as two-component responses composed of two independent intonations. However, the data show that the responses perform a single action that results from the interplay of two intonations. I will illustrate how the configuration of responses affects the display of stances.

In fragment 7, the speaker HS explains why and how she will get Hanbok, Korean traditional costume, for her wedding reception. The listener GA reacts with "*a cincca*" and "*a kulay*" tokens several times throughout the telling.

<Fragment 7: Hanbok>

```
1 GA:     그러면 한복도 빌리는 데로 또 가야 되겠네?
          kulemyen hanpok-to  pilli-nun tey-lo      tto
          ka-ya toy-keyss-ney?
          then    Hanbok-also borrow-RL place-LOC  again
          go-should-may-APP
          "Then you also have to go somewhere to borrow Hanbok?"

2 HS:     아 근데    [한복은 뭔가 내께 하고 싶기도 해서
          a kuntey [hanpok-un   mwenka  nay-kkey ha-ko siph-ki-to
          ha-yes
          a but    [Hanbok-TOP somewhat my-thing do-want-NOM-also
          do-because
          "A, but, as for Hanbok, I want to have my own one."
 (GA)               [(nod) (nod)

3 GA:     [음:: 그래        그래        그래
```

[mm:: kul-ay kul-ay kul-ay
mm be:so-INT be:so-INT be:so-INT
“Mm, sure sure sure.”

4 HS: 근데 엘에이가 되게 비싸다 그래서
kuntey eyleyi-ka toykey pissa-ta kul-ayse
but LA-NM very expensive-QT-because
“But, I heard that it is very expensive in LA.”

5 GA: → 아 진짜?=
a cincca?=

6 HS: =어 그래서 아마 한국(0.1)에 >한복 하는 데가
[있는데<
=e kulayse ama hankwuk(0.1)ey >hanpok ha-nun tey-ka
[iss-nuntey<
yes so maybe Korea-LOC >Hanbok do-RL place-NM
[exist-CIRCUM<
“Yes, so maybe, there is one Hanbok place in Korea.”
(GA) [(nod)(nod)(nod)

7 HS: [우리 오빠가 결혼을
[wuli oppa-ka kyelhon-ul
[my brother-NM marriage-ACC
“My brother got married,”
(GA) [(nod)(nod)(nod)(nod)

8 GA: 어:: (nod) (nod)
e:: (nod) (nod)

9 HS: 2년 전인가 해 가지구
2nyen cen-inka ha-y kacikwu
2year before-Q do-because
“2 years ago.”
10 GA: 어 어
e e

11 HS: [어 거기에서 했던 거 알아봐서
[e keki-eyse ha-yss-ten ke alapw-ase
[yes there-LOC do-PST-RT thing check-and
“Yes, we will check that place and,”
(GA) [(nod)(nod)(nod)

12 우리 이모가 올 때 가지고 오는 걸로
wuli imo-ka o-l ttay kaci-ko o-nun kel-lo
my aunt-NM come-RL when bring-and come-RL thing-by
“My aunt will bring it to me when she comes.”

13 GA: → [아,: [진짜?

[a,: [cincca?
[(nod)(nod)(nod)

14 HS: [아마도
[amato
[maybe
“Maybe.”

15 어
e
yes
“Yes.”

16 GA: 아마 너가 사이즈는 재서 가서?
ama ne-ka saicu-nun ca-yse ka-se?
maybe you-NM size-TOP measure-and go-and
“You should measure your own size?”

17 HS: 어 [근데 엄마가 그러는데 되게 쉽대 한복은
e [kuntey emmaka kule-nuntey toykey swip-tay hanpok-un
yes [but mom-NM be:so-CIRMUN very easy-QT Hanbok-TOP
“Yes, but my mom says, Hanbok is easy (to order).”
(GA) [(nod)

18 GA: → [아: 그래=
[a: kulay=
[(nod)(nod)(nod)

19 HS: =어 치수만 그냥 딱딱 얘기해 주면은
되게 [쉽게 한다고
=e chiswu-man kunyang ttakttak yaykiha-y cwu-myen-un
toykey [swip-key ha-ntako
=yes size-only just correctly tell-CNN give-if-TOP
very [easy-ADV do-QT
“Yes, if you only let them know your size correctly,
then they will do the job easily (she says).”

20 GA: → [아: 그래
[a: kulay
[(nod)(nod)(nod)

21 HS: [(nod)(nod)

22 GA: [(nod)(nod)(nod) 좋네
[(nod)(nod)(nod) coh-ney
good-APP
“That’s nice.”

Before this fragment, the speaker HS explained how she would get a wedding dress in LA. In line 1, the listener GA asks a *ney*-marked question whether HS also has to go somewhere to borrow Hanbok for her wedding reception. *Ney*-marked utterances in the first actions are used to indicate the action of first assessment and/or confirmation requests (Ha, 2018). The listener GA's question is seeking for a confirmation on her assumptions;(1) HS will borrow Hanbok from a store; (2) the store will be located in LA. The speaker HS begins the telling with correcting the listener's first assumption that she will not *borrow* Hanbok, but will *buy* one. This turn receives a multiple saying token "*kulay kulay kulay* (sure sure sure)" proceeded by an elongated vocalization "*mm*" token carrying a rise-fall contour. The response expresses the listener's news receipt followed by her agreement on HS's opinion with a high level of engagingness. In line 4, the speaker HS gives a preface projecting another correction for the listener's second assumption. In response to this, the listener produces the first "*a cincca*" token with a rising contour. In line 6 - 12, the speaker HS describes her plan for getting Hanbok from Korea. After the speaker's multiple turns, the listener deploys another "*a cincca*" token with co-occurring nods. In line 17, the speaker provides an additional piece of information and this turn is responded to with the first "*a kulay*" token carrying a rise-fall contour and a few nods. In the next turn, the speaker HS rephrases the previous turn with more details. The listener responds to this turn with the second "*a kulay*" token with nods and continuously nods until the her next turn. In line 21, the speaker does not produce any more turn but simply nods twice reciprocating with the listener's response. Without the continuation of the telling, the listener offers a short assessment "*coh-ney* (that's nice)".

(1) Affective stance

In this fragment, the "*a cincca*" structure is used twice. The affective stance of the first response "*a cincca*" is shown in the prosodic features as presented in Figure 34[25].

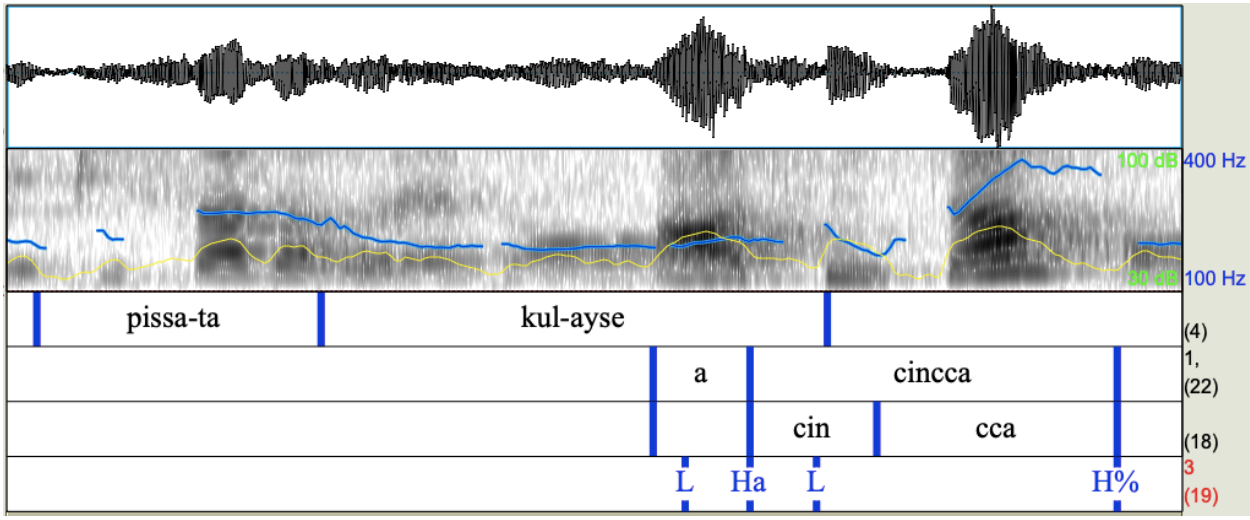


Figure 34. Praat information of the first "*a cincca*" structure (GA) and its prior turn (HS)[26]

The first "*a cincca*" is used in response to the speaker's turn "but, I heard that it is very expensive in LA" in line 4. The new information provided by the speaker prefaces the correction for the listener's second assumption that the Hanbok store HS will visit is located in LA. After hearing this, the listener, who has become surprised at and interested in the fact she did not know, expresses her strong unexpectedness and interest through the "*a cincca*" structure consisting of the preceding "*a*" token and the lexical token "*cincca*".

[25] The values used for GA and SU's voice range were: 100 - 400 Hz.

[26] Although the syllable "*cin*" seems to have a high pitch on the pitch contour, it is an error triggered by the overlapping sound of the speaker's turn. The syllable is produced with a low pitch in the original recording.

As demonstrated above, the rising "*cincca*" token displays the listener's surprised feeling through its lexical meaning and boundary tone (H%). Here, the preceding "*a*" token also shows a rising contour. However, the duration of "*a*" token is short and it forms an AP[27] which is not considered to convey discourse/pragmatic meanings in Korean. Although this is not a traditional interpretation, I propose to draw a distinction between the AP-final boundary tones of the preceding vocalization "*a*" and those of the lexical words. As shown in the responses in this section, when "*a*" is followed by another component, it is often shortened and forms an AP (Fig. 34 and Fig. 40), while it also occurs as an IP (Fig. 37 and Fig. 42). However, regardless of whether it forms an AP or IP, the "*a*" tokens with the same shape of intonation contour seem to display a similar meaning (Compare the two "*a cincca*" structures in Fig. 34 and Fig. 37, and the two "*a kulay*" structures in Fig.40 and Fig. 42). This seems to be triggered by the unique trait of "*a*", that it is a mono-syllabic exclamation which can be produced quickly when preceding a lexical token while not losing its discourse meaning. Therefore, I consider the preceding "*a*" tokens with an AP boundary tone to serve a similar function as those with an IP boundary tone, although the difference in the lengths can generate differences in their meanings. As for the first "*a cincca*" response, while the "*a*" token is an AP, it shows a rising intonation which shows the same intonation as the following "*cincca*" token. Here, the listener's affective stance displayed by the rising "*cincca*" token is intensified by the rising "*a*" token. Compared to the single "*a*" token or the single "*cincca*" token, the juxtaposition of these two rising tokens enhances the level of surprise expressed by the listener.

[27] An IP boundary can be changed to an AP-final boundary due to the fast speech rate (Fougeron and Jun 1998).

The listener also displays her interest through this response. It has been shown the form of the vocalization "*a*" denotes that the listener has been informed with new information (See 5.1.1.). Although it is produced with a rising intonation in this structure, the token form alone indexes the listener's realization which in turn elicits her interest towards the informing. However, compared to the rising-falling "*a*" token marked as LHL%, the level of interest in the rising "*a*" token is mild due to the lack of "L" tone. Therefore, the "*a cincca*" structure displays the listener's intensified surprise and mild interest. The overall response is reinforced by its pitch upgrade in relation to the speaker's informing.

The listener's facial expression also supports her display of affective stance. When producing the response, her eyebrows are raised and eyes are dilated as presented below.

(Before the response)

[*a cincca?*]

Figure 35. Listener GA's facial expressions

Eyebrow raising happens when a person is presented with a surprise (Navarro, 2018). Eye opening also indicates surprise and genuine interest (Smart, 1986). The listener's stance that she is being surprised and interested is presented through this facial expression along with her verbal response.

Figure 36. Seating arrangement of the participants; GA (left) and HS (right)

To sum up, the "*a cincca*" structure displays the listener's surprise and interest with extra emphasis on surprise. The listener's eyebrow raising and eye opening also supports her stance.

The second "*a cincca*" structure shows similar prosodic features as the first response (Figure 37[28]).

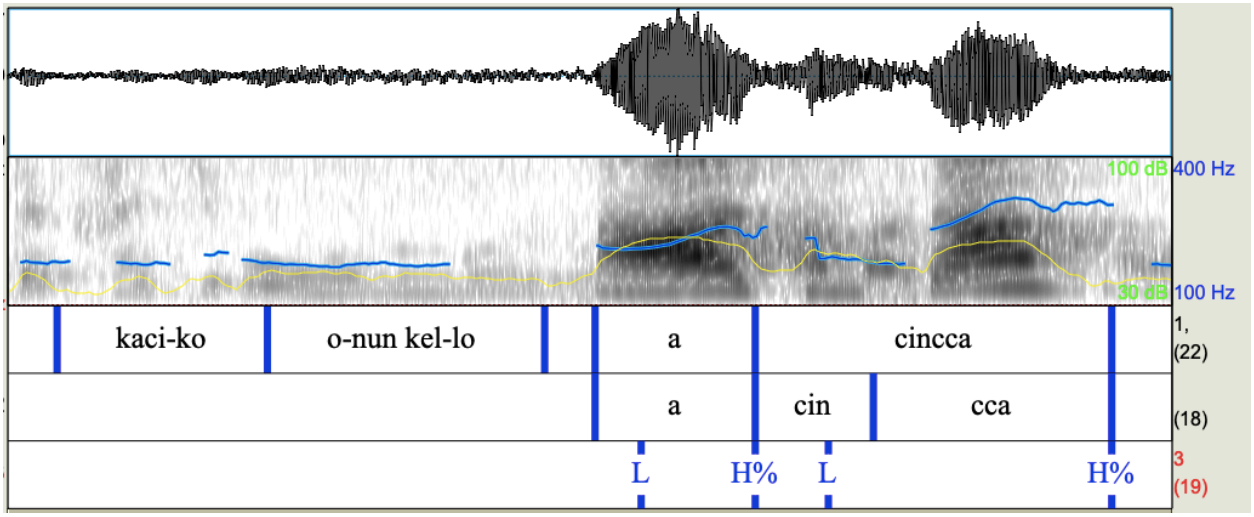


Figure 37. Praat information of the second "*a cincca*" structure (GA) and its prior turn (HS)

[28] [28] The values used for GA and SU's voice range were: 100 - 400 Hz.

The second "*a cincca*" is shown after the speaker has said her aunt will bring her Hanbok from Korea. This informing explains how the speaker HS will get her Hanbok for her wedding, and it is also contrary to what the listener initially expected. Similar to the first response, the listener displays her surprise through the "*a cincca*" structure with the final rising intonation. This time the preceding "*a*" token is produced longer than the one in the previous response and forms an IP. The repetition of the same intonation of the two consecutive tokens highlights the listener's affective stance. However, the pitch upgrade of the lexical item of this response is not as high as that of the same token in the first response.

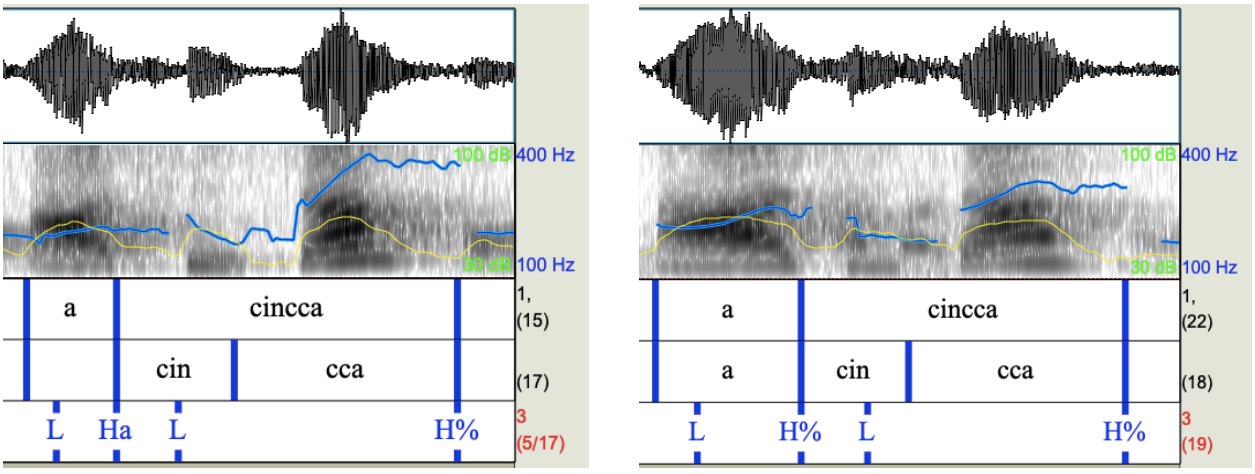


Figure 38. Praat info of the first "*a cincca*" (left) and second "*a cincca*" structure (right)

By managing the intensity of pitch upgrade of the lexical item, the listener seems to differentiate the second response from the first response. The "*cincca*" token in the second response is delivered in a lower pitch and lower intensity compared to the same lexical token in the first response. After the first informing, the listener has been informed that the speaker HS will not make Hanbok in LA, which is contrary to her expectation. For the listener who now knows this information, the impact of the second informing is not as big as the first one. Therefore, the

listener slightly lowers the degree of pitch upgrade to display that she is less surprised than before.

This is also displayed in her nods accompanied with the verbal response (Figure 39). The listener nods five times and each of the first three nods is produced along with each syllable of the verbal response "*a cincca*" with the first one consisting of only a head up movement. As explained above, nods can indicate that the listener accepts the validity of the new information while registering it.

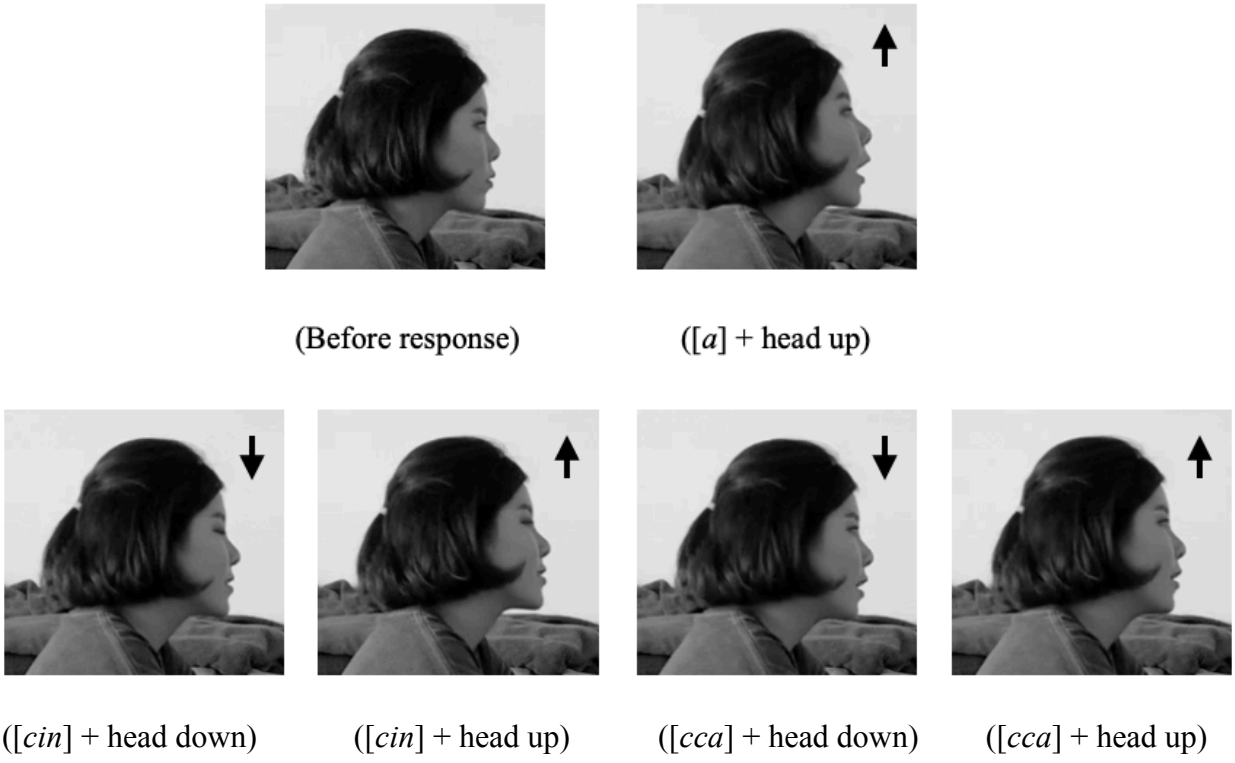


Figure 39. Listener GA's head movement

With respect to affective stance, I claim that co-occurring nods display empathy. Although the listener is displaying her surprise through the final rising intonation of the verbal response, this

time through the co-occurring nods, she also shows that she can understand more of the speaker's situation after having received the detailed plan.

In sum, while the listener displays her surprise through the second "*a cincca*" response, she also shows that this time she is less surprised (lower pitch) and more understanding and empathizing with the speaker (co-occurring nods).

The "*a kulay*" structure is also shown twice in the fragment. The prosodic features the first "*a kulay*" are shown in Figure 40[29].

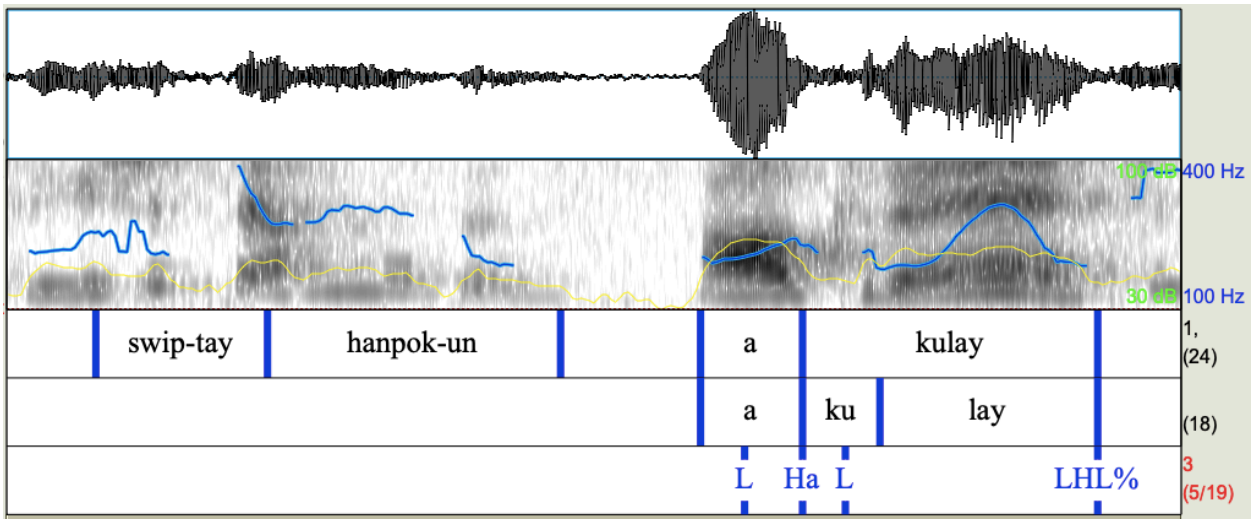


Figure 40. Praat information of the first "*a kulay*" structure (GA) and its prior turn (HS)

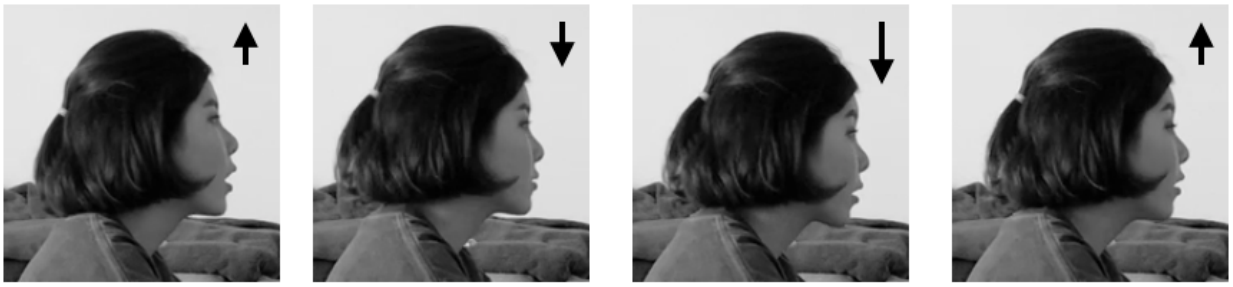
Unlike the "*a cincca*" responses, the lexical token in this response carries a rising-falling intonation. The LHL% tone displays the listener's "surprise" (H) and her becoming "interested" (L) with more emphasis on the interest (L). By adding the rising "*a*" token before the "*kulay*" token, the listener highlights both emotions through the token type ("*a*") and the boundary tone (Ha). While the "*a cincca*" responses with the boundary tone H% mainly

[29] The values used for GA and SU's voice range were: 100 - 400 Hz.

expresses the listener's “surprise” and its display of “interest” is relatively weak, the boundary tone LHL% presents both emotions equally strong.

In line 16, the listener brings up an informing-related question whether HS will measure her size by herself. In line 17, the speaker answers the question with a type-conforming “*e* (yes)” and voluntarily appends additional information that the ordering process of Hanbok is actually easy. This appendix justifies the speaker’s choice and therefore reduces any face-threatening impact that can be possibly triggered by having an idea that is contrary to the listener’s assumption. Here, the listener displays that she is surprised and interested through the “*a kulay*” structure. Compared to the “*a cincca*” responses with which she mainly expresses her unexpectedness, this time she displays more of “interest”. As I argued earlier, “interested” feeling is elicited by the process of realization. After hearing the speaker’s extra information that adds more validity to her decision, the listener shows that she is now more convinced, and therefore more interested. At the same time, she expresses surprise because the information is also new to her. A notable thing is that, unlike the previous “*a cincca*” responses, “*a kulay*” response shows a pitch downgrade compared to the speaker’s informing. The listener seems to display that she is less “surprised” than before by downgrading the pitch configuration, thereby differentiates the response from the previous ones.

Additionally, the response occurs with a few nods (Figure 41).



[*a*] (1st nod - head up) | [*ku*] (2nd nod - head down) | [*lay*] (2nd nod - head down) | [*lay*] (2nd nod - head up)

Figure 41. Listener GA's head movement

As I claimed above, the accompanying nods display the listener's empathy. The listener's nods enhance her stance that she is accepting the speaker's informing and therefore empathizing with her stance.

In sum, through the "*a kulay*" response, the listener expresses both of her surprise and interest equally strong. However through the pitch downgrade and co-occurring nods, the listener also displays that she is less surprised but more empathizing with the speaker.

The second "a kulay" structure shows similar prosodic features as the first response (Figure 42[30]). The second "*a kulay*" is shown after the speaker has basically rephrased her previous turn that ordering Hanbok is easy "if you only let them know your size correctly" in line 19. Here the listener again deploys "*a kulay*" structure. This time however, the preceding "*a*" token is produced for a longer time to the point where it is treated as an independent IP marked as the boundary tone LH%.

[30] The values used for GA and SU's voice range were: 100 - 400 Hz.

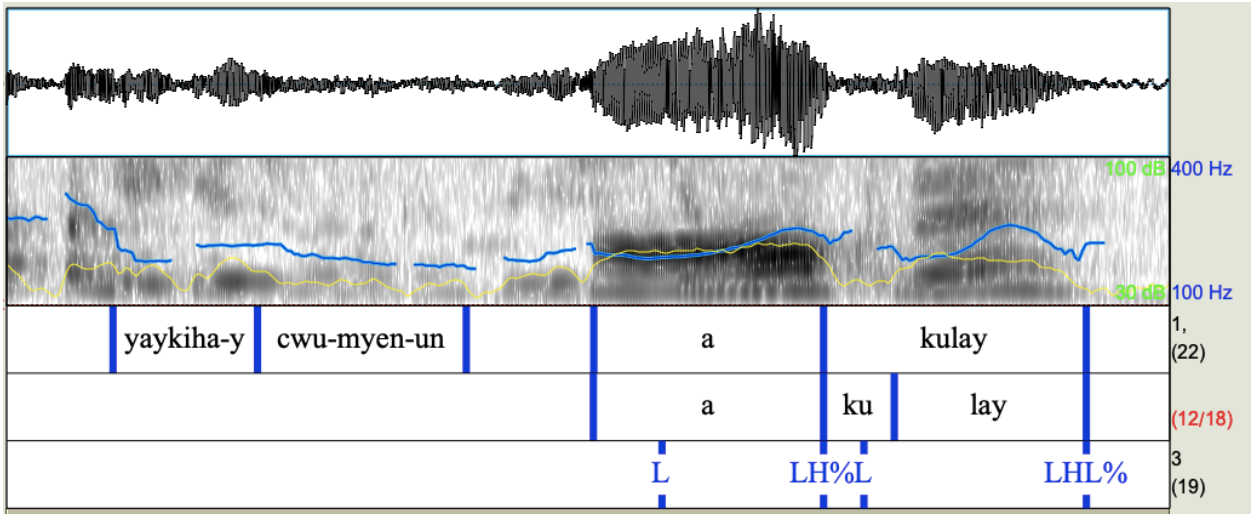


Figure 42. Praat information of the second "*a kulay*" structure (GA) and its prior turn (HS)

According to Park (2003), LH% expresses surprise highlighting the shift from the previous belief. Here, the lengthening of the "*a*" token in the second "*a kulay*" response intensifies the stance it delivers. On the other hand, the following "*kulay*" token is produced relatively short. The stance conveyed through the "*kulay*" token is weakened as its length is shortened and pitch is lowered compared to the same token in the previous response. The weakening of the response indicates that the listener treats the informing as relatively less significant with respect to how informative it is. It is seemingly because the speaker's informing in line 19 is merely a rephrase of the previous turn with only a minor new detail added.

Consequently, in relation to the first "*a kulay*" response, the "*a*" token is upgraded whereas the "*kulay*" token is downgraded. When considering the environment of the first and second response, the informing of the first "*a kulay*" response provides more substantial information, while the second informing is simply a repetition with a trivial detail. Therefore, I consider that the lexical token plays a more significant role in determining the overall affective

stance. In other words, the downgraded "*kulay*" token displays the listener's reduced surprise and interest treating the informing as less significant, while the impact of "*a*" token is relatively weak even though it is upgraded.

To sum up, similar to the first "*a kulay*" response, the second one also displays the listener's strong surprise and interest. However, by managing the length and pitch configuration of each component, the listener displays reduced intensity of her stance. Also it seems that, when placed together, the lexical token plays a more important role in displaying affective stance than does the vocalization.

(2) Epistemic stance

The *a*-prefaced "*a cincca*" and "*a kulay*" structures also display the listener's epistemic stance. The first and second "*a cincca*" response in this fragment consists of a rising "*a*" token and a rising "*cincca*" token.

After the listener has realized that the speaker's first informing is contrary to her previously held expectation, she employs the final rising "*a cincca*" structure. The lexical token "*cincca*" with a rising intonation, which serves the main function in this response, shows that she is still not yet fully K+. However, through the preceding vocalization "*a*", although it is produced with a rising contour, the listener displays that she is going through some kind of realization process and accepting the validity of the information to some extent. The combined structure of these two components signifies that the listener is placing herself somewhere between K- and K+ and the registration process of the new information is in progress.

The second "*a cincca*" response, however, compared to the first one, is delivered in a lower pitch (Figure 38").

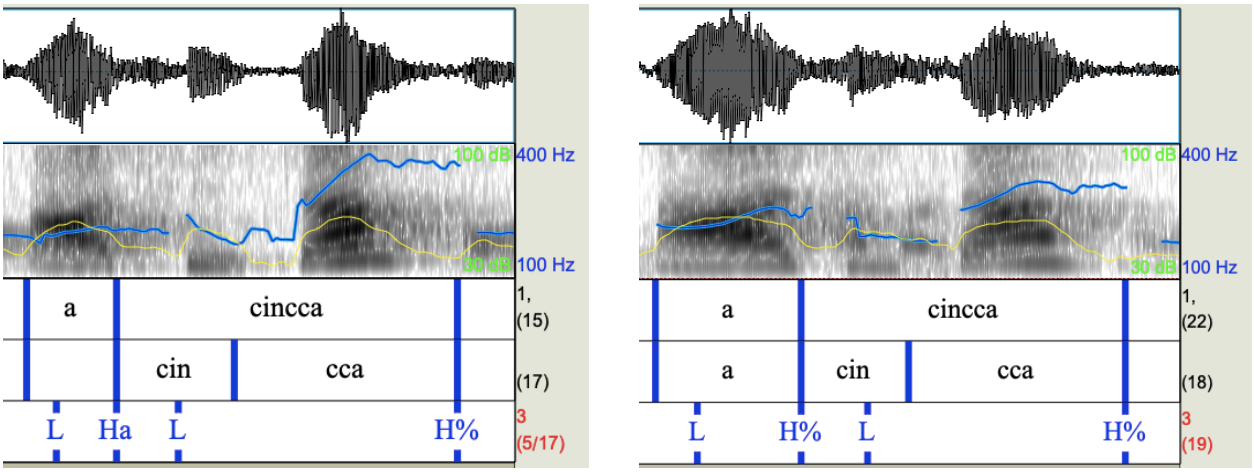


Figure 38". Praat info of the first "*a cincca*" (left) and second "*a cincca*" structure (right)

Although both responses carry a final rising intonation, it is observed that the higher the final pitch goes, the lower the epistemic stance is positioned. In other words, the listener can indicate her K- position in a stronger way with a higher final pitch. In this response, on the other hand, by lowering the final pitch of the second response, the listener is positioning herself as better informed than before, although she is still not fully K+. Also, as shown in Figure 39, the second "*a cincca*" response is accompanied with several nods. As I argued earlier, nods also display that the listener is registering new information and accepting the validity of it. By adding nods to the verbal response, the listener displays that she is registering it well, understanding it better, therefore she is more knowledgeable than before.

In sum, by managing the final pitch and co-occurring nods, the listener can make changes and add details in her epistemic positioning.

The first and second "*a kulay*" responses are composed of a rising "*a*" token and a rising-falling "*kulay*" token. In section 5.1.1, I claimed that with a rising-falling intonation, the listener positions herself as K+ by displaying that she has adopted the informing. Also, as explained above, the vocalization "*a*" token itself displays that the listener has been informed with new information and registering it. The juxtaposition of these two tokens emphasizes the listener's K+ positioning. Therefore, the first and second "*a kulay*" responses display that the listener is now fully K+ and the knowledge gap has been completely filled.

However, the second "*a kulay*" response is different from the first one in terms of its configuration. In the second one, the preceding "*a*" token is longer than the following "*kulay*" token. With respect to the affective stance of this response, I claimed that the impact of the "*kulay*" token is dominant because the lexical token plays a more important role in determining the discourse meaning. Therefore, the response displays the listener's downgraded surprise and interest towards the less informative informing, despite the upgrade of the "*a*" token.

As for the epistemic stance, on the other hand, the difference seems to be not so significant because both of the tokens display the listener is now K+. However, I argue that the configuration of the response contributes to the epistemic stance in a similar way. Although the "*a*" token is lengthened and therefore its meaning seems to be emphasized, because of the weakened following lexical token "*kulay*", the listener expresses that the informing is not as informative as the previous one. Here, the listener is showing that she has been fully K+ since the previous informing and there is only so much room left to upgrade her knowledge level. Through the downgraded "*a kulay*" response, the listener claims her prior knowledge she received from the immediately preceding turn, and thereby, she claims her already earned K+

status, while she still shows a slight upgrade in her knowledge level triggered by the minor new detail.

To sum up, the listener displays that she is now fully K+ with "*a kulay*" responses. Also by downgrading it, she claims that she has already held K+ status even before the current informing.

5.4. Summary and Discussion

In this chapter, the "oh/really" type tokens "*a*", "*cincca*", "*kulay*", and *a*-prefaced structures "*a cincca*" and "*a kulay*" have been discussed with respect to affective and epistemic stance they display.

First, I illustrated affective stance that an "*a*" token with a rising-falling intonation displays. The rising-falling intonation marked as the boundary tone LHL% shows the listener's surprise (H) followed by realization (L) (Park, 2003). Although the realization is part of the listener's cognitive process and its discussion is more relevant to her epistemic stance, I claimed that the process of realization elicits the listener's interest. Therefore, the rising-falling intonation displays the listener's feeling "surprised" followed by being "interested". I also argued that the vocalization "*a*" itself indexes that the listener has been informed with new information, and the stance is emphasized when the rising-falling intonation is carried on an "*a*" token. On the other hand, with a falling intonation, the "*a*" token can express the listener's empathy and sympathy by reserving excessive emotional expressions.

As for "*cincca*" and "*kulay*" token, with a rising intonation they display the low degree of certainty and surprise, while they show the listener's surprise and interest with a rising-falling intonation. The "*a cincca*" and "*a kulay*" structure consist of two components, an "*a*" token and a lexical token. The "*a cincca*" response composed of a rising "*a*" token and a rising "*cincca*" token displays the listener's surprise and interest with extra emphasis on surprise. On the other hand, the "*a kalay*" response consists of a rising "*a*" token and a rising-falling "*kulay*" token presents both surprise and interest equally strong.

Second, I demonstrated epistemic stance displayed by these responses. First, through the vocalization "*a*" signaling the receipt of the informing, the listener can display that he has received information and now he is K+. Also with both of a rising-falling and falling intonation, the listeners positions themselves as K+ by presenting the process of realization in the final falling intonation. This positioning is reinforced by nods and gaze that show the listener has registered the information and accepted the validity of it. Through "*cincca*" and "*kulay*" tokens with a rising intonation, the listener positions herself as not yet K+ but somewhere between K- and K+. With a rising-falling intoned "*cincca*" token, on the other hand, the listener shows she is now registering the information and more informed than before.

The "*a cincca*" and "*a kulay*" structures display more complex stances. In the "*a cincca*" response, the listener reveals her cognitive process through the vocalization "*a*", while she positions herself as not yet fully K+ through the "*cincca*" token. Therefore, the "*a cincca*" response displays that the listener is somewhere between K- and K+, while she is going through some kind of realization process. Also, I observed that the higher the final pitch goes, the lower the epistemic stance is positioned. By managing the final pitch of the intonation, the listener

makes changes and adds details in her epistemic positioning. On the other hand, as for the "*a kulay*" response, the juxtaposition of the vocalization "*a*" and the rising-falling "*kulay*" token highlights the listener's K+ positioning. Therefore, the listener displays that she is now fully K+ and the knowledge gap is completely filled. Also, by downgrading its pitch configuration, the listener claims that she has already held K+ status even before the current informing.

Before finishing the chapter, it will be worth noting one interesting phenomenon I encountered in this chapter. In my data, when the vocalization "*a*" is produced alone, it carries either rising-falling or falling intonation with its boundary tone marked as LHL% or L% and never occurs with a final rising intonation. Again, the boundary tone H denotes "surprise" and L indexes "realization" (Park, 2003). LHL% tone indexes surprise followed by realization with more emphasis on realization. As I illustrated in 5.1.1., the vocalization "*a*", through its token form itself, can display that the listener has been informed with new information and registering it. Therefore, the flow of the cognitive process displayed by the "*a*" token exactly accords with L% which indexes realization or LHL% which adds a short surprising moment before the realization.

However, when the "*a*" token is followed by another component, it often draws a rising contour marked as H% or LH% as shown in 5.3. Although more data and research will be needed, this seems to be a strong tendency. Then, why is "*a*" produced with a rising contour, not rising-falling contour, before a lexical item? Also, what is the contribution of the preceding "*a*" to the meaning of the entire "*a*" prefaced structure? To explain this, I will first divide the types of

the "*a*" prefaced structure into two groups depending on the boundary tone of the lexical token; H% and LHL%.

First, when the boundary tone of the lexical token is H%, the preceding "*a*" seems to copy the tone. This phenomenon seems to be related to "tone copy" referring to the repetition of the same boundary tone in two or more consecutive IPs (Gussenhoven, 1992). By adding a rising "*a*" token before the lexical token, the stance it conveys (surprise) is strengthened through the repetition of the same intonation.

On the other hand, when the lexical token is delivered in LHL% tone, two explanations are possible; (a) the preceding "*a*" token is intentionally produced with H% (or Ha) tone to emphasize surprise feeling, or (b) the preceding "*a*" token copies LHL% tone from the following token, but the final L is not realized. (a) is plausible in that the preceding "*a*" token can also be produced with full LHL% tone forming a clear repetition of the same intonation ([*a*]LHL% [*kulay*]LHL%). With this tone, the stance it conveys could be more vividly reinforced. Since there should be a difference between the meanings of H% tone and LHL% tone on the "*a*" token, listeners should be able to differentiate those two tones depending on their intention. Therefore, the H% tone on "*a*" token could have been deliberately chosen over the LHL% tone to convey certain stances.

However, although it is possible for listeners to produce the LHL% tone on the "*a*" token before a lexical token, so that they can emphasize the stance more clearly, the data show that the final L tone is less likely to be produced. Here, (b) can be the reason for the tendency. The reason why listeners tend to produce the "*a*" token with H% tone before a lexical token with LHL%

tone is that (1) there is not enough space and time for the final L tone to be realized, and (2) the omission of the L tone on the "*a*" token does not critically harm the display of its stance as long as the tone is produced through the lexical token. Because after all, what determines the core meaning of the entire response is the lexical token and its boundary tone.

To summarize, as for the "*a*" prefaced structures with the lexical token having LHL% tone, listeners technically have two options to choose from (final H tone or final L tone) with respect to the boundary tone of the preceding "*a*" token. Nevertheless, the "*a*" token is highly likely to be delivered with a final rising intonation (H%). This is because the time and space is limited before the immediately following token, and also because the full realization of LHL% tone on the "*a*" token is not necessary as it is the the boundary tone of the lexical token that mainly conveys the meaning of the entire response.

CHAPTER 6: CONCLUSION

6.1. Summary of Findings

This dissertation explored the display of listenership in Korean conversation. It aimed to demonstrate conversational meanings of reactive responses that are achieved through the sophisticated manipulation of prosodic and multimodal features. Particularly, it was investigated that how listeners display their affective and epistemic stance towards the prior turn in the course of telling. The analysis focused on the two major types of Korean reactive responses; yes/nod type and oh/really type.

Chapter 4 has illustrated the conversational meanings of the "yes" type tokens and nods with respect to affective and epistemic stance they display. As for affective stance, it has been discovered that the lexical token "*ung*" and "*e*" with a falling intonation display mild interest towards the prior turn, while the "*e*" token shows a higher level of engagingness than "*ung*" token. Also the co-occurring or following nods can display the listener's confirmation of the validity of the information besides her interest. The vocalization "*mm*", which is similar to the "*ung*" token but is semantically and phonetically weaker, expresses milder interest than does the "*ung*" token. Among the two types of nods, the mid-turn nods display a more intense degree of interest because of the less obligation imposed on a listener in the midst of the turns. On the other hand, the turn-completion nods express mere attentiveness since they are weaker than any other type of tokens used at the same position. The multiple saying tokens show enhanced interest towards the prior turn in accordance with the upgrade of the noteworthiness of the information given. Also, these tokens are usually accompanied with nods and facial expressions

that raise the level of interest. It has been proven that listeners show different levels of engagingness to the speaker's turns by managing the token type and the number of repeats.

The analysis has also shown epistemic stance displayed through "yes" type tokens and nods. Though the "*ung*" and "*e*" tokens, the listener can display her epistemic authority as well as epistemic stance. The listener claims her higher epistemic authority than the speaker for the telling by employing the "*e*" tokens, while she does not do so with the "*ung*" tokens. The listener can also display that she is "better" informed than before while she is on a path between K– and K+ through the "*ung*" and "*e*" tokens. The same stance is indexed by the "*mm*" tokens. The "*e*" token can also show the listener's upgraded epistemic stance, but, unlike "oh/really"-type tokens, it does not display "surprise". Through the mid-turn nods, the listener positions herself as K-, because she has not been informed with the information because the turn is still on the way. However, with the turn-completion nods, the listener claims her prior knowledge and displays her already earned K+ status treating the prior turn as not informative. With the multiple saying tokens, the listener emphasizes the fact that she has received the prior information and, therefore, has become more knowledgeable than before while she is still on a path between K– and K+. The listener manages the degree of informedness and readiness to obtain further information through the lexical choice, the number of repeats, and the co-occurring nods.

Chapter 5 has demonstrated the pragmatic meanings of "oh/really" type tokens in informing sequences. The analysis has shown that the rising-falling intoned "*a*" token displays the listener's feeling "surprised" followed by being "interested" because the process of realization elicits the listener's interest. In fact, the vocalization "*a*" alone indexes that the listener has been informed with new information, and the stance is emphasized when the rising-

falling intonation is carried on an "*a*" token. On the other hand, with a falling intonation, the "*a*" token can express the listener's empathy and sympathy by reserving excessive emotional expressions. As for "*cincca*" and "*kulay*" token, with a rising intonation they display the low degree of certainty and surprise, while they show the listener's surprise and interest with a rising-falling intonation. The "*a cincca*" and "*a kulay*" structure consist of two components, an "*a*" token and a lexical token. The "*a cincca*" response composed of a rising "*a*" token and a rising "*cincca*" token displays the listener's surprise and interest with extra emphasis on surprise. On the other hand, the "*a kalay*" response consists of a rising "*a*" token and a rising-falling "*kulay*" token presents both surprise and interest equally strong.

The finding has also revealed epistemic stance of "oh/really" type tokens. The listener can display that he has received information and now he is K+ through the vocalization "*a*" signaling the receipt of the informing. Also with both of a rising-falling and falling intonation, the listeners positions themselves as K+ by presenting the process of realization in the final falling intonation. Nods and gaze can reinforce this stance because they show that the listener has registered the information and accepted the validity of it. Through "*cincca*" and "*kulay*" tokens with a rising intonation, the listener positions herself as not yet K+ but somewhere between K- and K+. On the other hand, with a rising-falling intoned "*cincca*" token, the listener shows she is now registering the information and more informed than before.

The "*a cincca*" and "*a kulay*" structures display more complex stances. In the "*a cincca*" response, the listener reveals her cognitive process through the vocalization "*a*", while she positions herself as not yet fully K+ through the "*cincca*" token. Therefore, the "*a cincca*" response displays that the listener is somewhere between K- and K+, while she is going through

some kind of realization process. Also, I observed that the higher the final pitch goes, the lower the epistemic stance is positioned. By managing the final pitch of the intonation, the listener makes changes and adds details in her epistemic positioning. On the other hand, as for the "*a kulay*" response, the juxtaposition of the vocalization "*a*" and the rising-falling "*kulay*" token highlights the listener's K+ positioning. Therefore, the listener displays that she is now fully K+ and the knowledge gap is completely filled. Also, by downgrading its pitch configuration, the listener claims that she has already held K+ status even before the current informing.

6.2. Suggestions for Future Research

Some related topics can be suggested for the future research. First, it will be interesting to see the sequential development triggered by reactive responses. Although sequential development is one of the main topics dealt with in conversation analytic studies, it was not in the scope of this study due to its main focus on the stance display of reactive response. How the responses further the sequence will show the importance of listener's role more evidently. Second, the correlation between the lexical tokens and their boundary tones will be worth researching. For example, although both of the two lexical tokens ("*cincca*" and "*kulay*") can possibly occur with either H% or LHL%, there seems to a tendency that "*cincca*" is more likely to be produced in H% and "*kulay*" tends to show LHL%. With more abundant data, this tendency will be more clearly shown and the correlation can be discovered.

6.3. Implication of the Study

This dissertation has implications with respect to a few agendas. First, it suggests a new definition and classification of reactive responses. In the previous literature, the criteria used for the categorizations lack consistency and the terminologies themselves are vague and disputable. To solve this problems, I suggested a new perspective on identifying reactive responses by considering the diverse aspects they display and treating them as different layers that can be activated simultaneously. Second, the study attempts to discover listeners' stance by analyzing reactive responses. By the close examination on the form, prosody, and multimodal aspects of each response, it has shown that even a very minor feature of reactive response can make changes in the delivery of the listener's intention. This study will not only contribute to the field of functional linguistics, but also can be a pedagogical foundation for language education.

APPENDIX A

TRANSCRIPTION CONVENTIONS

[	Overlapping talk begins
]	Overlapping talk ends
=	Latching between lines (no break or gap)
(0.1)	Length of silence in tenths of a second
-	Cut-off
.	Falling, or final intonation
?	Rising intonation
,	Continuing intonation
:	Sound stretch
> <	Compressed or rushed talk
hh	Laughter, or hearable exhalation or aspiration (outbreaths)

APPENDIX B

ABBREVIATIONS

ACC	Accusative	PL	Plural suffix
ADV	Adverbial suffix	POL	Polite speech level
APP	Apperceptive	PST	Past tense suffix
CIRCUM	Circumstantial	Q	Interrogative
COMM	Committal	QT	Quotative particle
CONJ	Conjecture	LOC	Locative marker
DC	Declarative suffix	RL	Relativizer
HON	Honorific	RT	Retrospective
INT	Intimate speech level	REPORT	Reportive
NOM	Nominative	TOP	Topic marker

REFERENCES

Angles, Jeffrey., Nagatomi, A., & Nakayama, M. (2000). Japanese responses hai, ee, and un: Yes, no, and beyond. *Language and Communication, 20*, 55–86.

Aoki, H. (2008). Hearership as Interactive Practice: A multi-modal analysis of the response token Nn and head nods in Japanese casual conversation. PhD Dissertation.

Berger, I., & Rae, John. (2012). Some uses of gestural responsive actions. *Journal of Pragmatics, 44*, 1821–1835

Bommel, T., Boasso, A., & Ruscher, J. B. (2013). Looking up for answers: Upward gaze increases receptivity to advice. *Current Research in Social Psychology, 22*, 60-70.

Clancy, P. M., Thompson, S. A., Suzuki, R., & Tao, H., (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics, 26*, 355–387.

Curl, T. S. (2005). Practices in other-initiated repair resolution: The phonetic differentiation of 'repetitions.' *Discourse Processes, 39(1)*, 1–44.

Drummond, K., & Hopper, R. (1993). Back channels revisited: Acknowledgment tokens and speakership recipiency. *Research on Language and Social Interaction, 26*, 157-177.

Du Bois, J. W. (2007). The stance triangle. In: Englebretson, R. (Ed.), *Stancetaking in Discourse*. John Benjamins, Philadelphia, 139–182.

Duncan, S., & Fiske, D. (1977). Face-to-face interaction : research, methods, and theory, illsdale, N.J. : L. Erlbaum Associates ; New York : Distributed by Halsted Press.

Enfield, N. J. (2011). Sources of asymmetry in human interaction: Enchrony, status knowledge and agency. In T. Stivers, L. Mondada & J. Steensig (Eds.), *The morality of knowledge in conversation*. Cambridge: Cambridge University Press.

Fishman, P. M. (1983). Interaction: The work women do, Revised version in B. Thorne, C. Kramarae, & N. Henley, (Eds.), *Language, gender, and society,* 89–101. Cambridge, England: Newbury House.

Fridlund, A. J. (1991). Sociality of Solitary Smiling: Potentiation by an Implicit Audience. *Journal of Personality and Social Psychology, 60*, 229-240.

Gardner, R. (1997). The listener and minimal responses in conversational interaction. *Prospect, 12(2)*, 12–32.

Gardner, R. (2001). *When listeners talk: Response tokens and listener stance*. Amsterdam: John Benjamins.

Goffman, E. (1961). Asylums: Essays on the social situations of mental patients and other inmates. Doubleday (Anchor).

Goodwin, C. (1986). Between and within: alternative sequential treatments of continuers and assessments. *Human Studies, 9*, 205–217.

Goodwin, M. H. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry, 50*, 303–317.

Gubina, A., & Betz, E. (2021). What do newsmark-type responses invite? The response space after German echt. *Research on Language and Social Interaction, 54(4)*.

Gussenhoven, C., & Rietveld, T. (1992). A target-interpolation model for the intonation of dutch. Proc. 2nd International Conference on Spoken Language Processing (ICSLP 1992), 1235-1238

Ha, K. (2018). The Social Actions Conveyed by the Sentence-ending Suffixes –Ney, –Ci, and –Kwuna in Korean Conversation. Ph.D. Dissertation, University of California at Los Angeles.

Heritage, J. (2012). The epistemic engine: Sequence organization and territories of knowledge, *Research on Language and Social Interaction, 45(1)*, 30–52.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis.* 299-345. Cambridge: Cambridge University Press.

Heritage, J., & Raymond, G. (2005). The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly, 68(1)*, 15-38.

Iwasaki, S. (1997). The Northridge earthquake conversations: the floor structure and the 'loop' sequence in Japanese conversation. *Journal of Pragmatics, 28*, 661–693.

Jarque, M. J. & Pascual, E.(2015). Direct discourse expressing evidential values in Catalan Sign Language. In J. Martines & V. Miglio (guest eds.). Special issue on 'Approaches To Evidentiality in Romance'.

Jefferson, G. (1981). "Caveat Speaker": A preliminary exploration of shift-implicative recipiency in the articulation of topic. End of grant report to the (British) Social Science Research Council on the articulation of topic in conversation.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens 'yeah' and 'mm hm'. *Papers in Linguistics, 17(2)*, 197-216.

Jun, S. A. (2000). K- ToBI (Korean ToBI) Labelling Convention Version 3.1

Jun, S. A. (1990). The Prosodic Structure of Korean – in terms of voicing in E-J. Baek (ed.) Proceedings of the Seventh International Conference on Korean Linguistics Vol. 7: 87-104, Univ. of Toronto Press.

Jun, S. A. (2005). *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press.

Kamio, A. (1997). *Territory of information*. Amsterdam: John Benjamins.

Kawanishi, Y., & Sohn, S.-O. (1993). Grammaticization of Korean negation: A semantic-pragmatic analysis of canh-a(yo). In S. Kuno, J. Whitman, I.-H. Lee, J. Mailing & Y.-J. Kim (Eds.), The fifth Harvard international symposium of Korean linguistics, 552– 561, Cambridge, MA: The Harvard ISOKL Committee.

Kim, K. H. (1999). Phrasal unit boundaries and organization of turns and sequences in Korean conversation. *Human Studies, 22*, 425-446.

Kim, K. H., & Suh, K. H. (2010). The Sentence-Ending Suffix -Ketun in Korean Conversation with Reference to -Nuntey : Sequence Organization and Management of Epistemic Rights. *Discourse and Cognition 17(3)*, 1–38.

Lee, H. S. (1999). A discourse-pragmatic analysis of the committal -ci in Korean: A synthetic approach to the form-meaning relation. *Journal of Pragmatics, 31(2)*, 243-275.

Mandelbaum, J. (1987). Speakership recipiency in the accomplishment of a telling. Ph.D dissertation, University of Texas, Austin, TX.

Maynard, D. (1997). The news delivery sequence: bad news and good news in conversational interaction. *Research on Language and Social Interaction, 30(2)*, 93-130.

Maynard, S. K. (1990). Conversation management in contrast: listener response in Japanese and American English. *Journal of Pragmatics, 14,* 397–412.

McCarthy, MJ (2003). Talking back: "Small" interactional response tokens in everyday conversation. *Research on Language and Social Interaction 36(1*), 33–63.

Müller, FE. (1996). Affiliating and disaffiliating with continuers: prosodic aspects of recipiency. In: Couper-Kuhlen, Elizabeth, Selting, Margret (Eds.), *Prosody in Conversation*. Cambridge University Press, Cambridge, pp. 131–176.

Nagata, R. (2004). Kaiwa ni okeru aizuchi no kinoo: hatsuwa tochuu ni utareru aizuchi ni chakumokushite (Functions of backchannels in conversation: with special focus on utterance-internal backchannels). *Nihongo Kyoiku, 120*, 53–62.

Navarro, J. (2018). *The Dictionary of Body Language A Field Guide to Human Behavior.* HarperCollins Publishers

O'Keeffe, A., and Adolphs, S. (2008). Response Tokens in British and Irish Discourse: Corpus, Context and Variational Pragmatics', in K. Schneider and A. Barron (eds) V*ariational Pragmatics: A Focus on Regional Varieties in Pluricentric Languages,* 69–98, Amsterdam: John Benjamins.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambering University Press.

Ogden, R. (2006). Phonetics and social action in agreements and disagreements. *Journal of Pragmatics, 38,* 1752–1775.

Oh, S. Y., & Park, Y. Y. (2017). Interactional Uses of Acknowledgment Tokens: ‘ung’ and ‘e' as Responses to Multi-Unit Turns in Korean Conversation. *In Enabling Human Conduct: Studies of Talk-in-Interaction* in Honor of Emanuel A. Schegloff, ed. by Geoffrey Raymond, Gene H. Lerner, and John Heritage, 145-166. Amsterdam: John Benjamins.

Park, M. J.(2003). The meaning of Korean prosodic boundary tones. Ph.D. Dissertation, University of California at Los Angeles.

Peräkylä, A., & Ruusuvuori, J. (2006). Facial expression in an assessment. In H. Knoblauch, B. Schnettler, J. Raab, & H-G. Soeffner (Eds.), *Video analysis: Methodology and methods,* 127–142. Frankfurt am Main, Germany: Peter Lang.

Pyun, D. O. & Yoon, K. E. (2022). Discourse functions of Korean ‘yes’ words. *Korean Linguistics*, 48–75.

Young, R., & Lee, J. (2004). Identifying units in interaction: Reactive tokens in Korean and English conversations. *Journal of Sociolinguistics, 83*, 380–407.

Sacks, H. (1974). An analysis of the course of a joke’s telling in conversation. In R. Bauman & J. Sherzer (Eds.), *Explorations in the ethnography of speaking,* 337–353. Cambridge, UK: Cambridge University Press.

Schegloff, E. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences In D Tannen (ed) *Analyzing Discourse: Text and Talk* Washington: Georgetown University Press.

Schegloff, E. (1997). ‘Narrative analysis’ thirty years later. *The Journal of Narrative and Life History, 7*, 97-107.

Selting, M. (1987). Verständigungsprobleme. Eine empirische Analyse am Beispiel der Bürger-Verwaltungs- Kommunikation [Communication problems: An empirical analysis using the example of citizen administration communication]. Niemeyer.

Sidnell, J. (2011). *Conversation analysis: An introduction (Vol. 45)*. John Wiley & Sons.

Smart, William E. (1986), Eye-Opening Body Language. The washington post. September 17, 1986.

Stivers, T. (2004). 'No no no' and other types of multiple sayings in social interaction. *Human Communication Research, 30(2)*, 260–293.

Stivers, T. (2008). Stance, alignment and affiliation during story telling: Nodding as a token of preliminary affiliation. *Research on Language & Social Interaction, 41*, 31–57.

Stivers, T. (2012). *Sequence organization*, In The Handbook of Conversation Analysis (Jack Sidnell, Tanya Stivers, eds.), Oxford, U.K., Wiley-Blackwell, 191–209.

Thompson, S. A., Fox, B. A., & Couper-Kuhlen, E. (2015). *Grammar in everyday talk: Building responsive actions*. Cambridge: Cambridge University Press.

Tsuchiya, K. (2013). *Listenership behaviours in intercultural encounters : a time-aligned multimodal corpus analysis*. Benjamins.

Wilkinson, S., & Kitzinger, C. (2006). Surprise as an interactional achievement: Reaction tokens in conversation. *Social Psychology Quarterly, 69(2)*, 150–182.

Xu, Jun. (2014). Displaying status of recipiency through reactive tokens in Mandarin task-oriented interaction. *Journal of Pragmatics, 74*, 33--51.

Yang, J. (2013). A multimodal study of response token Dui Dui Dui in Mandarin conversation. Master Thesis, University of Alberta, Canada.

Yngve, V. H. (1970). On getting a word in edgewise. In Papers from the sixth regional meeting of the Chicago Linguistic Society, 567-578. Chicago, IL, USA: Chicago Linguistic Society.

Zimmerman, D. H. (1993). Acknowledgment Tokens and Speakership Incipiency Revisited, *Research on Language and Social Interaction, 26(2)*, 179-194.