Are You Flirting, Objectifying or What?
a Conversation Analysis of "you're very sexy" Conversational Turn

Kadek Ratih Dwi Oktarini

School of Humanities, Nanyang Technological University Singapore
✉48 Nanyang Avenue, Singapore 639818
✉ rkadek@ntu.edu.sg

Article Info	ABSTRACT

Intent identification is one of the most critical components in conversational agent design. Conversational agent "is any dialogue system that not only conducts natural language processing but also responds automatically using human language." (Conversational Agent, 2019). The crux of designing human-like conversational agent is to mimic how human understands another human and then responds "naturally". The current study attempts to answer the fundamental question: how to model human processes of understanding another human? In order to answer that question, it starts from exploring some basic concepts relevant to intent identification from Conversation Analysis (CA). CA is a mature field that studies authentic human interaction. The basic concepts from CA are then synthesised into a model that potentially fit to existing framework and paradigm in conversational agent design, i.e. Natural Conversation Framework (NCF) and Intent-Entity-Context-Response (IECR) paradigm. Instead of using a made-up sentence, the model is then tested to an authentic conversational turn *seksi sekali dirimu* 'you're very sexy'. The test shows that the model is able to detect several possible intents contain in this authentic conversational turn. The model is also able to handle Conversational Indonesian and multi-modality. Considering the versatility of Conversation Analysis, in all likelihood the model will be able to handle any language and all kinds of modalities. Future study can be done to analyse more Conversational Indonesian data (to develop library of intent for Conversational Indonesian Language), as well as conversational data from different languages and conversational data containing diverse modalities.

INTRODUCTION

With the advancement of Natural Language Processing (NLP) and computational power in general, conversational agent becomes ubiquitous. Conversational agent "is any dialogue system that not only conducts natural language processing but also responds automatically using human language." (*Conversational Agent*, 2019). Apple Siri, Amazon Alexa, Google personal assistant, Microsoft Cortana, are some of the conversational agents that have recently become household names.

So far, we can celebrate the conversational agents' successes in searching for and retrieving information. The phrase "let's google it" has become a staple of our every life. Conversational agents are doing well in specific tasks. For example, in phone-banking, now we give our information to Conversational Agent before talking to human customer service. When we enter a website, we often have a "person" greet us through the small chat windows. That person is Conversational Agent of some sort. The downside is that we may still notice that those Conversational Agents are incapable of understanding complex messages. We may spend more time than needed if we talk to a Conversational Agent than speaking to human customer service. We are still far from having a conversational agent that can hold a conversation similar to a human. One vital issue is to design a Conversational Agent or system that can understand and respond to human the way human do.

To address this issue, we may turn to studies on human conversation. There is a mature field of studying authentic human interaction. Studies in this field may belong to interactional linguistics, social psychology, social interaction, linguistic anthropology, human-computer interaction (HCI), and so on. If we look closely, a considerable number will employ a method of Conversation Analysis (CA). CA is a branch of ethnomethodology, an approach within the field of sociology. CA is a way of thinking, approach, and method, that is sourced from observing the naturally-occurring conversation. It is dedicated to studying how human (or members of society) make sense of and conduct conversational interaction (Cf. Francis & Hester, 2004; Heritage, 1984; Sacks et al., 1974).

One framework for conversational agent that has incorporated Conversation Analysis' paradigm and rich findings is Natural Conversation Framework (NCF) (Moore et al., 2017; Moore & Arar, 2019). NCF is a promising framework. The conversational agent built based on this framework will be able to handle bite-sized, back and forth conversational interaction, much like human conversation. In Moore et al. (2017) and Moore and Arar (2019), within the NCF, there is Intent-Entity-Context-Response (IECR) paradigm. Intent identifies the conversational action of the input. While Entity and Context are the "context" part of the framework. Entity identifies the exact domain of the social action or "intent". Then context determines properties of the context consequential to the conversation.

Moore et al. and Moore & Arar in their volumes above did not explicate in detail how intent identification is performed. From their explanation (Moore & Arar, 2019, p. 49), it appears that "brainstorming" is the method to build the library of intent. If we turn to Conversation Analysis, we may be able to build a library of intent, based on authentic conversational data.

In addition to the above, to the best of my knowledge, there is no work on Indonesian Conversation within, or at least compatible, with Natural Conversation Framework (NCF) and Intent-Entity-Context-Response (IECR) paradigm. There is also no known accessible open library for intent in Conversational Indonesian compatible with the Intent-Entity-Context-Response (IECR) paradigm. In general, conversational Indonesian has received very little attention, though it is the language that Indonesian people use in their daily lives (Ewing, 2005; Sneddon, 2003). More studies are done on its counterpart, the Standard Indonesian (Bahasa Indonesia yang Baik dan Benar). Some Indonesians may not even realise that they speak differently than the ideal language they have in mind (Cf. Englebretson, 2003, 2007).

The current study is a small step towards addressing the gaps mentioned in the prior paragraph. It attends to the questions of, how to model human processes of understanding another human? In answering that question, the current study will propose an intent identification model for Conversational Agent, informed by Conversation Analysis method. In addressing the gap for Conversational Indonesian, the current study will employ the proposed model to a piece of data taken from authentic Indonesian conversation. In the future, a comprehensive library of "intent" for Indonesian Conversation can be built by analysing more (conversational) turns from authentic Indonesian Conversation. The model can also be continuously improved as more data is analysed.

Considering the versatility of Conversation Analysis, in all likelihood, the model will be able to handle any language and all kinds of modalities. Future study can be done to analyse more Conversational Indonesian data (to develop a library of intent for Conversational Indonesian Language), as well as conversational data from different languages and conversational data containing diverse modalities.

METHODS

Data
The specific (conversational) turn analysed in the current study is taken from a collection named "flirtatious sequence" (Oktarini, 2017). This collection consists of 8 sequences of flirtatious interaction, taken from ±50 minutes of video-recorded data of two-party conversation.

Flirting has been defined to be an ambiguous kind of activity (Hopper, 2003; Speer, 2017). Hence, "intent" identification is very challenging in this kind of sequential environment. The specific turn in question (i.e., the target turn) is one of the simplest and most straightforward turns in the collection. While the responsive turn is highly ambiguous. The target turn is chosen to illustrate the value and versatility of the proposed intent identification model. It has value since the proposed intent identification model can identify "hidden" intents even in, what appears to be, a simple and straightforward (conversational) turn. It is versatile because it can still operate even when the response to the target turn is ambiguous.

Ethical Considerations
The participants have given their consent for their image and voice to be used for the study. Still, in maintaining their anonymities, pseudonyms (FS for First Speaker and SS for Second

Speaker) are used instead of their real names. Then, line-drawings or sketches are in use in the place of the participant's original images.

Transcription

The data is transcribed using Conversation Analysis transcription convention or Jeffersonian transcription convention (Jefferson, 1985, 2004). An abridged version of the transcription convention is available as an appendix (Appendix: Transcription Convention).

Three-lines transcription is used in the current study to preserve crucial Indonesian grammatical information. The first line is the data in its original language; the second line is the word by word glossing or, "gloss", while the third line is idiomatic translation. The goal is to make sure that linguistic information of the original language does not get "lost in translation". Three lines transcription is a common way to present data from non-English language. See Extract 1 below:

```
1 C    :     =On' naji    yo    [eri mo              original language
              same         FP    [collar ADVP         gloss
             "((It))'s the same  [the collar too"     idiomatic translation

2 A    :                        [A! honto::          original language
                                 [oh really           gloss
                                 ["Oh! really?"       idiomatic translation
```

Extract 1: (Tanaka, 2000, p. 8) [source]

In the above example, for C's speech, we can see the first line (original language), *On' naji yo eri mo*. On the second line (gloss), there is a word by word translation from Japanese to English. Then, on the third line, there is an idiomatic translation "((It))'s the same the collar too".

"FP" in the gloss (second line) of C's turn is an abbreviation of Final Particle. Japanese Final Particle appears at the end of a sentence, and it commonly marks the end of a sentence. The open bracket "[" appear across C and A's talk in the transcript means that the talk following it is produced in an overlap. So C's talk, *eri mo* and A's talk *A! honto::* are produced in overlap or at the same time. Without having the second line, the gloss, reader who is not familiar with Japanese will not know that the overlap occurs right after the Final Particle (FP) and not after the word "same" as suggested by the English translation. This kind of information is crucial to the analysis. Hence, gloss is also provided for non-English data.

LITERATURE REVIEW

How Human Identify Intent (According to Conversation Analysis)?

The current study will propose an intent identification model for Conversational Agent inform by Conversation Analysis method and findings. In so doing, firstly it will review the existing literature in CA relevant to human "intent" identification mechanism. Then based on that review, the current study constructs and proposes an intent identification model for Conversational Agent.

The closest notion to intent in CA is "action". In CA, there is already a well-defined understanding of how human identifies or understands each other's action in conversation. The

relevant critical concepts for our discussion are Action Ascription and Adjacency Pair. Before outlining the two key concepts above, lets touch upon the basic analytical units in CA.

Conversation Analysis is developed based on observing the authentic conversation, or naturally occurring conversation, i.e. recorded conversation with a minimum, close to zero, intervention. Though its analysis takes linguistics units into account, CA's fundamental analytical unit is conversational turn, or "turn". A turn is a unit of talk, produced by a single speaker before any speaker change occurs. A turn is constructed through at least a single Turn Constructional Unit (TCU). A TCU can be a linguistics and extra-linguistics unit. Linguistics units are such as sentence, clause, phrase, lexical construction, and lexicon, etc. (Cf. Couper-Kuhlen & Selting, 2017; Sacks et al., 1974). Extra-linguistics units are such as laughter (Cf. Glenn, 2003; Glenn & Holt, 2013; Jefferson et al., 1987), prosody (Cf. Barth-Weingarten et al., 2010; E. A. Schegloff, 1998), eye-gaze, smile, facial expression, bodily movements, etc. (Cf. Charles Goodwin, 2000; Mondada, 2019). Conversation analysis is a considerably mature field of research. Hence there has been a sizeable body of research on both linguistics and extra-linguistics TCU.

Each of the TCU may be identified as a vehicle of a (social) action or actions. Consequently, a turn may be recognised as a vehicle of a (social) action or actions. The action or actions may be derived from the action of each of its TCU. The action or actions can also be derived from some combination(s) of some, or all, of its TCUs. In CA the relationship between a TCU or turn with a social action(s) is often written simply as "the TCU or turn is doing X, Y, Z, etc. action", with X, Y, Z as the name of the action.

Action Ascription refers to the mechanism used by a (human) speaker in a conversation to ascribe at least one action to the turn that they heard in a conversation. This mechanism is the one that ties the conversational unit of Adjacency Pair. Adjacency Pair is a unit of talk consisting of at least two conversational turns. Adjacency Pair may be extended to consist of more than two-turns, but to limit the complexity handled in the current paper, we will only focus on two-turns Adjacency Pair. The two turns in an Adjacency Pair have chronological and type-specific (or action-response) relationship. The current discussion is a cursory discussion on Action Ascription and Adjacency Pair. For more in-depth discussions, see Levinson (2013) for Action Ascription and Schegloff (2007) for Adjacency Pair.

As mentioned above, a unit of two turns with a sequential and type-specific action-response relationship is termed as Adjacency Pair. The first turn in an Adjacency Pair is termed First Pair Part (FPP), and the speaker is referred to as the First Speaker (FS). The second turn is termed as Second Pair Part (SPP), and the speaker is referred to as the Second Speaker (SS). Some examples of Adjacency Pair are the pairing of FPP-SPP of Greeting-Greeting, Request-Granting, Invitation-Acceptance/Rejection, Assessment-Assessment, and so on.

The relationship between FPP and SPP can be illustrated through Chart 1 below. Chart 1 exemplifies the relationship between an invitation FPP and an Acceptance SPP.

```
1    FS:    Invitation                          1st Turn (FPP)
2    SS:    Acceptance                          2nd Turn (SPP)
```
Chart 1: Invitation-Acceptance Adjacency Pair Chart

Action Ascription is the process that enables the Second Speaker (SS) to respond to First Speaker (FS) turn. It is not a simple encoding-decoding process. It is a process of SS actively ascribing a certain action or actions based on FPP's design and structure. For instance, an invitation as a social action may be produced verbally through a multitude of linguistic structures. Some turns and TCUs are more transparent, while some are opaquer than others. Some turn and TCUs may be understood as doing more than one action. If the action of the FPP is opaque or complicated, as an analyst, we can identify FS's action through SS's response. We can deduce that SS ascribes that the FPP is doing an invitation if the action done through the SS' response is acceptance to invitation. As an analyst, we have empirical evidence that such turn, the FPP, which is structured in that specific way, is understood as doing invitation.

Now let us continue with an authentic conversational (Extract 2).

```
1      J: Let's feel the water. Oh, it …              FPP
2      R: It's wonderful. It's just right. It's like  SPP
3         bathub water.                               SPP
4         ((feeling the water))                       SPP
```
Extract 2: (Anita Pomerantz, 1984, p. 57) [source]

J's whole talk (Line 1, Extract 2) is a single turn. J's turn consists of two TCUs; both are verbal:
1. "Let's feel the water",
2. "Oh, it …".
3. (Feeling the water)

Provided by our understanding of the English language, the first TCU can be recognised as an "invitation to feel the water". The second TCU "Oh, it" potentially indicates that J is feeling the water (TCU number 3). The source does not provide J's behaviour while producing Line 1. However, from the second TCU, we may deduce that she is touching and feeling the water when she produces her second TCU. Here again, to note the difference between CA and linguistics analysis, the incomplete sentence can still be analysed. It is not "defective" data. Whether the sentence is complete, or not, the main concern is whether the linguistic structure has a potential to be a vehicle of (an) action(s), or indicative of (an) action(s). If it can be identified as doing even a single action, then we can take it as a TCU.

R's talk (Line 2 -3, Extract 2) is a single turn. R's turn consists of four TCUs.
1. Feeling the water
2. "It's wonderful."
3. "It's just right."
4. "It's like bathtub water."

Each of the TCU can be identified as R's evaluation of the water. The evaluation is sourced from her act of touching and feeling the water. There is a specific CA term for evaluation, i.e. assessment. Assessment evaluates an object or entity in interaction (Cf. C. Goodwin & Goodwin, 1992; Lindstrom & Mondada, 2009). To date, there is already sizeable research on assessment and second assessment (response to an assessment) in CA. As a whole, R's verbal turn (Line 2 and 3) can be identified as assessing the water.

Now, let us move on with analysing the two turns as parts of an Adjacency Pair. The FPP is J's turn (Line 1), and the SPP is R's turn (Line 2-3). Based on the analysis above, the action of the FPP is an invitation to feel the water, while the SPP is an assessment of the water (verbal) and feeling the water (non-verbal). Now we can see that Extract 2 is an invitation-acceptance Adjacency Pair. It is an invitation to feel the water. The invitation is verbal, while the acceptance is non-verbal (R feeling the water). The potential non-verbal behaviour of SS (J) of feeling the water may add extra weight on the verbal invitation.

As an analyst, we have empirical evidence that such turn (Extract 2, Line 1), that consists of the two TCUs is understood as doing invitation. Besides, we also identify that invitation to feel the water can be responded with a cluster of actions, consisting of feeling the water (non-verbal) and three assessments of the water (verbal). Those three assessments of the water can be understood as sourced from the non-verbal action of feeling the water.

SPP Action FPP Action	Feeling the water	Assessment: "It's wonderful."	Assessment: "It's just right."	Assessment: "It's like bathtub water."
Invitation to feel the water	X	X	X	X
(Feeling the water)	X			

Table 1:   Extract 2 FPP-SPP Action Pairing

Proposed CA Based Intent Identification Model
Above, we have analysed a conversational turn through Conversation Analysis (CA) method. The steps are as below:
1. Splicing the two turns (FPP and SPP) into its TCUs (Turn Constructional Units)
2. Analysing actions done through
   a. Each of the TCUs
   b. Some combinations of the TCUs in each of the turns
   c. All TCUs in each the turns as a single turn
3. Pairing actions of the FPP to the actions of the SPP

If our target of intent identification is the FPP, we can propose an intent identification model for Conversational Agent informed by Conversation Analysis as illustrated in Chart 1 below. The target turn is the FPP, while the responsive turn is the SPP.
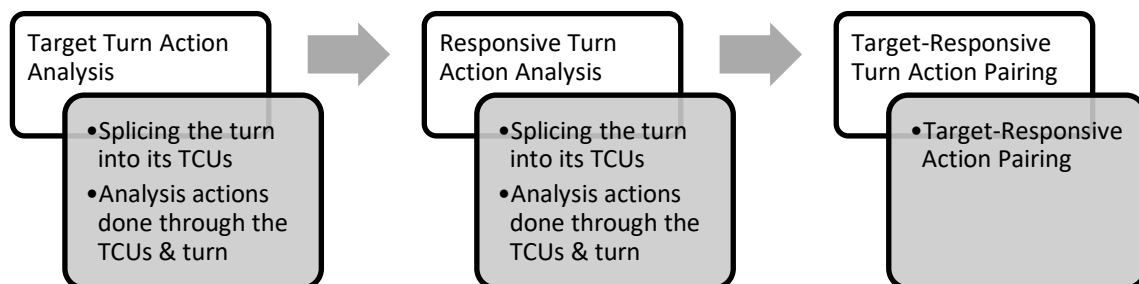

Chart 1: Intent Identification Model for Conversational Agent informed by Conversation Analysis

ANALYSIS

Now that we have a preliminary intent identification model, we can apply the model to analyse our data. This section will be divided into three sub-section, following the model: Target Turn Action Analysis, Responsive Turn Action Analysis, and then Target-Responsive Turn Action Pairing. Implication and input to the proposed model will be discussed in the next section, i.e. Discussion.

```
1 → TS            : Seksi seka-[li dirimu]          Target Turn
                    Sexy  very  self-you
                    You('re) very sexy
2   RS            :            [ eh   ehm]hem hem    Responsive Turn
```


Extract 3: Target Turn and its Immediate Response
        Flirtatious sequence collection (Oktarini, 2017) [source]

FS (First Speaker) produces his turn, the FPP (Line 1) while looking at Second Speaker (SS) intently. SS's turn, the SPP (Line 2) is soft laughter, done while turning her head upward and away from FS. The line drawing in Extract 3 is rendered from the still image taken from the moment when SS produces Line 2. We can see from the drawing that the SS produces her laughter while she closes her mouth tightly and turns her head slightly away from FS. We can see FS looks at SS intently in the line-drawing. He maintains this gaze behaviour from the moment he produces his turn (Line 1) up to SS's response (Line 2).

Target Turn Action Analysis

```
1   FS:           Seksi seka-[li dirimu]            FPP
                  Sexy  very  self-you
                  You('re) very sexy
```


Extract 4: Target Turn
        Flirtatious sequence collection (Oktarini, 2017) [source]

TCU:
1. Verbal   : *Seksi sekali dirimu* 'you're very sexy'
2. Behaviour   : gazing at SS intently

TCU 1 Recognisable Action 1: Assessment (Assessing Eva)
Line 1 consists of a single syntactic TCU, *seksi sekali dirimu* 'you're very sexy'. The turn is produced while looking at SS. Hence, *dirimu* 'you' is understood to be directed at SS, thus *seksi sekali* 'very sexy' is to be understood as evaluating SS. Hence, the first recognisable action of Line 1 is an evaluation or assessment, specifically, an assessment on SS self.

TCU 1 Recognizable Action 2&3: Compliment and Objectification
One of the meanings enlisted in OED for the word compliment (n) is, "…a neatly-turned remark addressed to anyone, implying or involving praise…" ('Compliment, n.', 2020). For some, being sexy is positive. Also, here in the target turn, SS is not only evaluated as being sexy but "very" sexy. She is being evaluated as having a high degree of sexiness. For some, this kind of evaluation can be considered praise. In that sense, the target turn can be identified as doing a compliment.

On the other hand, an object of evaluation is also prone to be objectified. One of the meanings enlisted in OED for the word objectification (n) is, "The demotion or degrading of a person or class of people (esp. women) to the status of a mere object…" ('Objectification, n.', 2020). In that sense, the mere evaluation of SS's sexual quality may be understood as an objectification.

TCU 1 Recognizable Action 4&5: Expression of Existing Intimacy and Invitation to Subsequent Intimate Talk
The keyword is "sexy". Saying a female interlocutor very sexy is different than, from example, saying that her hair is wavy. The first one possibly involves "noticing" (Sacks, 1992) of one's sexual quality, while the second one involves noticing one's hair type.

For some, sexually assessing someone may be considered as an "improper". The introduction of the improper topic in the conversation has long been identified as both indications of existing intimacy and an invitation to do more intimate talk (Coupland & Jaworski, 2003; Jefferson et al., 1987). By producing "improper" talk, the speaker indicates that he and the addressee has a sufficiently intimate relationship to produce such talk. At the same time, it may also occasion or invite more intimate talk.

TCU 2 Recognizable Action: Observation (observing Eva)
FS looks at SS throughout the production of his turn. He directs his gaze at SS attentively as if he is closely observing SS's facial expression. In that sense, a non-verbal action done in concurrent in Line 1 is FS observing SS.

Responsive Turn Action Analysis
```
1    FS:           Seksi seka-[li dirimu]              FPP
                   Sexy  very  self-you
                   You('re) very sexy
2    SS:                      [ eh   ehm ]hem hem      SPP
```

Extract 5: Subsequent Turn
Flirtatious sequence collection (Oktarini, 2017) [source]

Before FS completes his turn (Line 1), SS turns her head away and produces four beats of soft-voiced exhalation while smiling. This behaviour is marked in the transcription (Extract 5) as Line 2. SS's turn (Line 2), is produced in overlap with the tail-end of the target turn (Line 1). It is produced even before the target turn is completed. Also, SS does not show any sign of confusion. Considering its placement and absence of any sign of confusion, SS may have sufficient understanding of the action being done in the target turn even when Line 1 is yet to be completed. Hence, it is fair to analyse SS's turn (Line 2) as a response to the target turn (Line 1).

The units of behaviour (TCU) exhibit in Line 2 are soft laughter, head-turning and non-Duchenne smile. Duchene smile is the kind of smile that involves the raising of the corners of the mouth and the cheek, as well as creasing of the corners of eyes (Ekman et al., 1990). SS's smile does not involve the creasing of the corners of her eyes. Hence, her smile can be characterised as a non-Duchenne smile. Then, on laughter, laughter as a conversational object can be used to perform a multitude of actions (Glenn, 2003; Glenn & Holt, 2013; Greer et al., 2005; Jefferson, 1985, 2010). Below are the TCUs in SS's turn.

TCU:
1. Verbal   : 'eh   ehm hem hem' (soft laughter)
2. Behaviour   : head-turning (away)
3. Behaviour   : non-Duchenne smile

TCU 1 Recognisable Action: Mild Resistance
Though laughter is generally produced in response to something funny, it has been identified that laughter may also be produced in problematic situations (Drew, 1987; Glenn, 2003; Jefferson et al., 1987; Shaw et al., 2013). Glen (2003) observes that laughter can be produced as a responsive action to resist the relevance sets by the initiating action. While Drew (1987) identifies that a target of a tease may laugh in response to the tease, and then subsequently deny the tease. If we see different forms of resistances in conversation as a cline, SS's laughter in Line 2 would be a mild one. The strong one would be, for example, a verbal expression of disagreement alongside some kinds of hostile behaviour. In this sense, SS's response can be recognised and categorised as mild resistance.

TCU 2& 3 Recognisable Action: Display of Embarrassment
SS performs a combination of modalities that can be understood as a display of embarrassment. She turns her head away from FS while performing what can be characterised as a non-Duchene smile. Also, she turns her head away and presses her lips. Display of embarrassment is found to be marked by a non-Duchene smile, lip press, and head movement (Keltner, 1995, 1996). By

combining our observation on SS's smile and her head position, we can identify SS's action here as a display of embarrassment.

FPP-SPP Action Pairing

Based on the recognisable actions of the target turn (Line 1) and the subsequent turn (Line 2), we can map three initiating-responsive action pair of action (Adjacency Pair). The identified pairing can be summarised in Table 2 below:

SPP Action / FPP Action	Mild Resistance	Display of Embarrassment
Assessment	X	X
Compliment	X	X
Objectification	X	
Expression of Existing Intimacy	X	
Invitation to Subsequent Intimacy Talk	X	
Observation	X	X

Table 2:   FPP-SPP Action Pairing

Recognisable Adjacency Pair 1: Assessment – Display of Embarrassment

The next possible Adjacency Pair is assessment – display of embarrassment. According to Keltner and Buswell (1997, p. 250), there are three possible causes of embarrassment: loss of self-esteem, concern for others' evaluations, or absence of scripts to guide interactions. SS embarrassment may be produced in orientation to concern for other's evaluation. Hence, the action pairing may be Assessment – Display of Embarrassment.

Recognisable Adjacency Pair 2: Observation – Display of Embarrassment

The next possible Adjacency Pair is observation – display of embarrassment. SS may not know what to do in response to FS's observing her. Her embarrassment may result from Keltner and Buswell (1997, p. 250) the third scenario mentioned above, i.e. absence of scripts to guide interactions.

Recognisable Adjacency Pair 3: Compliment – Display of Embarrassment

A different possible source of SS's embarrassment is that if she takes FS's turn as a compliment. Responding to a compliment has been identified as a tricky business (Golato, 2002; Pillet-Shore, 2015; A. Pomerantz, 1978). There are two conflicting preferences at work: preference for agreement and avoid self-praise. When one blatantly disagrees with a compliment, one may be considered as unfriendly to the speaker, just as any other blatant disagreement with an assessment. However, when one directly agrees to a compliment, one may be understood as doing self-praise. Pomerantz (1978) observed that an addressee commonly dodges the compliment while displaying his or her understanding that the speaker has just produced a compliment. Therefore, the two preferences are satisfied.

SS's display of embarrassment may be a means to satisfy the two preferences. By being embarrassed, she may indicate that she does not fully embrace her entitlement to the compliment; While, subtly displays her understanding, acceptance, and affiliation to FS's compliment.

Recognisable Adjacency Pair Action 4: (Action) – Mild Resistance

Laughter produced in a second position instead of a type-fitting action may indicate a resistance to the relevance sets by the first action. It can be produced in response to any of the actions identified in the first step. In the context of interaction, this kind of action is useful since it signals that the second speaker is not willing to conform to whatever action being done in the prior turn. However, this response is not a type-specific response. Hence, it does not give enough clue to the type of action its response to.

DISCUSSION

There are several recognised actions in the target turn (FPP that do not have a direct pairing with the responsive action (SPP), they are objectification, statement of intimacy, and an invitation to intimate talk. Considering the wide relevant first action a mild resistance can respond to, SS may respond to any or all of those actions through her mild resistance. Though we do not identify "new" action, there is one crucial thing that we can learn from the current analysis. The target turn, *sexy sekali dirimu* 'you're very sexy' for the reason explored in the analysis section, may cause embarrassment in SS and resistance to produce type-fitting response in SPP. For intent identification, this turn, and possibly other turn with similar characteristics, can then be identified as doing or containing the intent of causing embarrassment and causing resistance.

So the intents of the target turn can be written in Intent-Entity-Context-Response (IECR) paradigm (pseudocode) (Moore & Arar, 2019, p. Appendix B) as:
Verbal:

    # Assessment
    # Compliment
    # Expression of Existing Intimacy
    # Invitation to Subsequent Intimacy Talk
    # Causing embarrassment
    # Causing resistance

Non-Verbal

    # Observation
    # Embarrassment causing action
    # Resistance causing action

This analysis and the first version of Library of Conversational Indonesian Intent is accessible through the link:
https://docs.google.com/document/d/11dHrNIW1qgyuX3rp11mTXkdFCp7G5yevFx7-k-7i1uI/edit?usp=sharing

CONCLUSION

The current study attends to the question of how to model human processes of understanding another human? In answering this question, the current study has proposed an intent identification model for Conversational Agent, informed by the empirically grounded Conversation Analysis method. The current study has also employed the model to analyse, thus

identify the different intents of and authentic conversational turn. It is evident that the model works and is able to handle Conversational Indonesian and multi-modality.

Considering the versatility of Conversation Analysis, in all likelihood the model will be able to handle any language and all kinds of modalities. Future study can be done to analyse more Conversational Indonesian data (to develop library of intent for Conversational Indonesian Language), as well as conversational data from different languages and conversational data containing diverse modalities. The intent identification model developed in the current study is built with Natural Conversation Framework (NCF) and Intent-Entity-Context-Response (IECR) paradigm in mind. However, considering the kind of data it able to handle and the versatility of Conversation Analysis, it may still hold the potentials to inform other conversational agent frameworks and paradigms.

REFERENCES

Barth-Weingarten, D., Reber, E., & Selting, M. (Eds.). (2010). *Prosody in Interaction*. John Benjamins Publishing Co.

Compliment, n. (2020). In *OED Online*. Oxford University Press. https://www.oed.com/view/Entry/37717

Conversational Agent. (2019, May 17). DeepAI. https://deepai.org/machine-learning-glossary-and-terms/conversational-agent

Couper-Kuhlen, E., & Selting, M. (2017). *Interactional linguistics: An introduction to language in social interaction*.

Coupland, J., & Jaworski, A. (2003). Transgression and Intimacy in Recreational Talk Narratives. *Research on Language and Social Interaction*, *36*(1), 85–106. https://doi.org/10.1207/S15327973RLSI3601_5

Drew, P. (1987). Po-faced receipts of teases. *Linguistics*, *25*(287), 219–253.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne Smile: Emotional Expression and Brain Physiology II. *Journal of Personality and Social Psychology*, *58*(2), 342–353.

Englebretson, R. (2003). *Searching for structure the problem of complementation in colloquial Indonesian conversation*. J. Benjamins.

Englebretson, R. (2007). *Stancetaking in Discourse Subjectivity, evaluation, interaction*. John Benjamins Publishing Company.

Ewing, M. C. (2005). Colloquial Indonesian. In K. A. Adelaar & N. Himmelmann (Eds.), *The Austronesian Languages of Asia and Madagascar*. Routledge.

Francis, D. W., & Hester, S. (2004). *An invitation to ethnomethodology: Language, society, and social interaction*. Sage Publications. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=251450

Glenn, P. J. (2003). *Laughter in interaction*. Cambridge University Press.

Glenn, P. J., & Holt, E. (2013). *Studies of laughter in interaction*.

Golato, A. (2002). German compliment responses. *Journal of Pragmatics*, *34*(5), 547–571. https://doi.org/10.1016/S0378-2166(01)00040-6

Goodwin, C., & Goodwin, M. H. (1992). Assessments and the construction of context. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon* (Vol. 11, p. 147). Cambridge University Press.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, *32*(10), 1489–1522. https://doi.org/10.1016/S0378-2166(99)00096-X

Greer, T., Usui, Y., Kato, T., Taniguchi, H., & Bussinguer, V. (2005). Suppressing laughter in the display of (dis) affiliation. *Kobe Daigakui Kokusai Komyunikeeshon Sentaa Ronshuu (Kobe University International Communication Centre Bulletin)*, *2*, 27–42.

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Polity Press.

Hopper, R. (2003). *Gendering talk*. Michigan State University Press.

Jefferson, G. (1985). An exercise in the transcription and analysis of laughter. In T. V. Dijk (Ed.), *Handbook of Discourse Analysis* (pp. 25–34). Academic Press.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–31). John Benjamins Pub.

Jefferson, G. (2010). Sometimes a frog in your throat is just a frog in your throat: Gutturals as (sometimes) laughter-implicative. *Journal of Pragmatics*, *42*(6), 1476–1484. https://doi.org/10.1016/j.pragma.2010.01.012

Jefferson, G., Sacks, H., & Schegloff, E. A. (1987). Notes on laughter in the pursuit of intimacy. In G. Button & J. R. E. Lee (Eds.), *Talk and Social Organization* (pp. 152–205). Clevedon, Avon, England ; Philadelphia : Multilingual Matters, c1987. http://psycnet.apa.org/psycinfo/1987-98425-006

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, *68*(3), 441–454. https://doi.org/10.1037/0022-3514.68.3.441

Keltner, D. (1996). Evidence for the Distinctness of Embarrassment, Shame, and Guilt: A Study of Recalled Antecedents and Facial Expressions of Emotion. *Cognition and Emotion*, *10*(2), 155–172. https://doi.org/10.1080/026999396380312

Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin*, *122*(3), 250–270. https://doi.org/10.1037/0033-2909.122.3.250

Levinson, S. C. (2013). Action Formation and Ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 103–130). Wiley-Blackwell.

Lindstrom, A., & Mondada, L. (2009). Assessments in Social Interaction: Introduction to the Special Issue. *Research on Language & Social Interaction*, *42*(4), 299–308. https://doi.org/10.1080/08351810903296457

Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, *145*, 47–62. https://doi.org/10.1016/j.pragma.2019.01.016

Moore, R. J., & Arar, R. (2019). *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. Morgan & Claypool.

Moore, R. J., Arar, R., Ren, G.-J., & Szymanski, M. H. (Eds.). (2017). *Conversational UX Design*. ACM Press. http://dl.acm.org/citation.cfm?doid=3027063.3027077

Objectification, n. (2020). In *OED Online*. Oxford University Press. https://www.oed.com.remotexs.ntu.edu.sg/view/Entry/129623

Oktarini, K. R. D. (2017). *Flirtatious Sequence in Indonesian: A Conversation Analytic and Membership Categorization Analytic Approach* [Doctorate, Nanyang Technological University, Singapore]. https://repository.ntu.edu.sg/handle/10356/72240

Pillet-Shore, D. (2015). Compliments. In *The International Encyclopedia of Language and Social Interaction*. John Wiley & Sons, Inc. http://onlinelibrary.wiley.com.ezlibproxy1.ntu.edu.sg/doi/10.1002/9781118611463.wbielsi127/abstract

Pomerantz, A. (1978). Compliment responses: Notes on the co-operation of multiple constraints. *Studies in the Organization of Conversational Interaction*, 79–112.

Pomerantz, Anita. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 57–95). Press Syndicate of the University of Cambridge.

Sacks, H. (1992). *Lectures on conversation, volumes I and II* (G. Jefferson, Ed.). Blackwell.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735. https://doi.org/10.2307/412243

Schegloff, E. A. (1998). Reflections on studying prosody in talk-in-interaction. *Language and Speech*, *41*(3–4), 235–263.

Schegloff, E. A. (2007). *Sequence Organisation in Interaction, Volume 1: A Primer in Conversation Analysis*. Cambridge University Press.

Shaw, C., Hepburn, A., & Potter, J. (2013). Having the last laugh: On post-completion laughter particles. In P. J. Glenn & El. Holt (Eds.), *Studies of laughter in interaction* (pp. 91–106). https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/15482

Sneddon, J. N. (2003). Diglossia in Indonesian. *Bijdragen Tot de Taal-, Land-En Volkenkunde*, *159*(4), 519–549.

Speer, S. A. (2017). Flirting: A Designedly Ambiguous Action? *Research on Language and Social Interaction*, *50*(2), 128–150. https://doi.org/10.1080/08351813.2017.1301297

Tanaka, H. (2000). Turn Projection in Japanese Talk-in-Interaction. *Research on Language & Social Interaction*, *33*(1), 1–38. https://doi.org/10.1207/S15327973RLSI3301_1

Appendix: Transcription Convention

Below are some relevant symbols employ in the transcription and its explanation. The List and explanation are adapted from Jefferson's transcription convention (Have, 2007; Jefferson, 1985, 2004).

Temporal and Sequential Relationship:

[....]	separate bracket on two successive lines with utterances by different speaker indicates an overlap or the point where an embodiment starts and end with regards to the utterance.
[....	open bracket indicates the start of an overlap. To alleviate the transcriber burden, for the first round of transcription, we can use an open bracket and forego the close bracket.
=	equal signs, one at the end of one line and one at the beginning of a next, indicate no 'gap' between the two lines (latching). The sign comes in pairs. One sign is placed at the end of a line and another at the start of the next line.

Characteristics of Speech Production word

:	colons indicate a prolongation of the immediately prior sound.
::	multiple colons indicate a more prolonged sound.
.	a period indicates a stopping fall in tone.
,	a comma indicates a continuing intonation, like when you are reading items from a list.
?	a question mark indicates a rising intonation.

Transcriber's doubts and comments

()	empty parentheses indicate the transcriber's inability to hear what was said. The length of the parenthesised space indicates the length of the untranscribed talk.
(())	double parentheses contain transcriber's description rather than, or addition to, transcription.

References:

Have, P. ten. (2007). *Doing conversation analysis*. SAGE.

Jefferson, G. (1985). An exercise in the transcription and analysis of laughter. In T. V. Dijk (Ed.), *Handbook of Discourse Analysis* (pp. 25–34). London: Academic Press.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: studies from the first generation* (pp. 13–31). Amsterdam; Philadelphia: John Benjamins Pub.