

Gazing, pointing and modality expression in conversations

Mayumi Bono

Yasuhiro Katagiri

ATR Media Science Laboratories, Japan

bono@atr.jp

Motivations

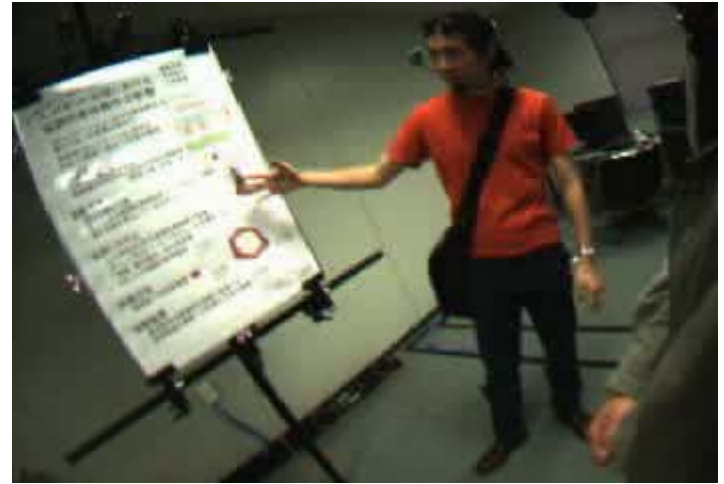
- How do gesture and speech jointly contribute to conversation organization?
 - Information delivery
 - Interaction management
- How are looking and pointing actions related with linguistic expressions in conversations?
 - Particularly interactions involving external resources
 - Poster presentations

In a Poster Presentation...



•Overview phase,

- <Gaze> Exhibitor almost looks at audience.
- <Gesture> Exhibitor expresses representational gestures.



•Explanatory phase,

- <Gaze> Exhibitor looks at audience and poster one after the other.
- <Gesture> Exhibitor expresses pointing gestures.

Looking Action & Pointing Action

- **Current studies**

- Gazing
 - Kendon (1967) studied the patterning of gaze direction and speech in conversational dyads
 - Monitoring function, regulatory function, expressive function
- Pointing
 - Some studies focus on pointing gesture as index function.

- **Target gestures**

- **Looking actions**
- **Pointing actions**
 - Looking at an object and pointing at an object are similar actions.
 - These have an index function in conversations.

- **We focus on...**

- Temporal relations of Looking and Pointing in conversations
- Differences between Looking and Pointing
- Conversational turn management function of Looking and Pointing
- Expressions of speech concerned with Looking and Pointing

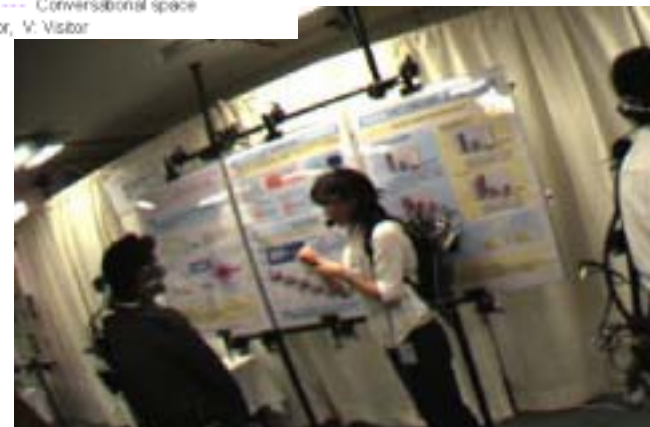
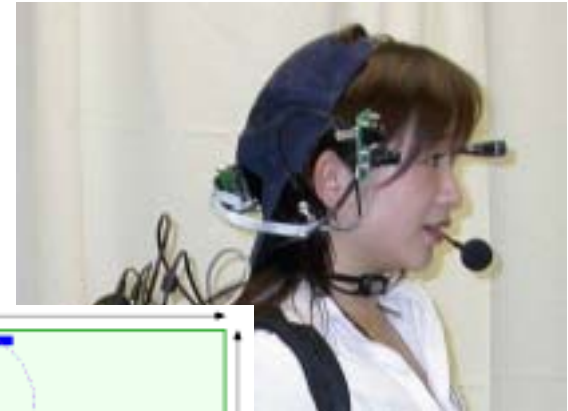
Data collection

- **Conversations in a poster presentation**

- Exhibitors explain their poster contents to visitors
- Exhibitor is always the dominant speaker
- Visitors sometimes have questions or responses

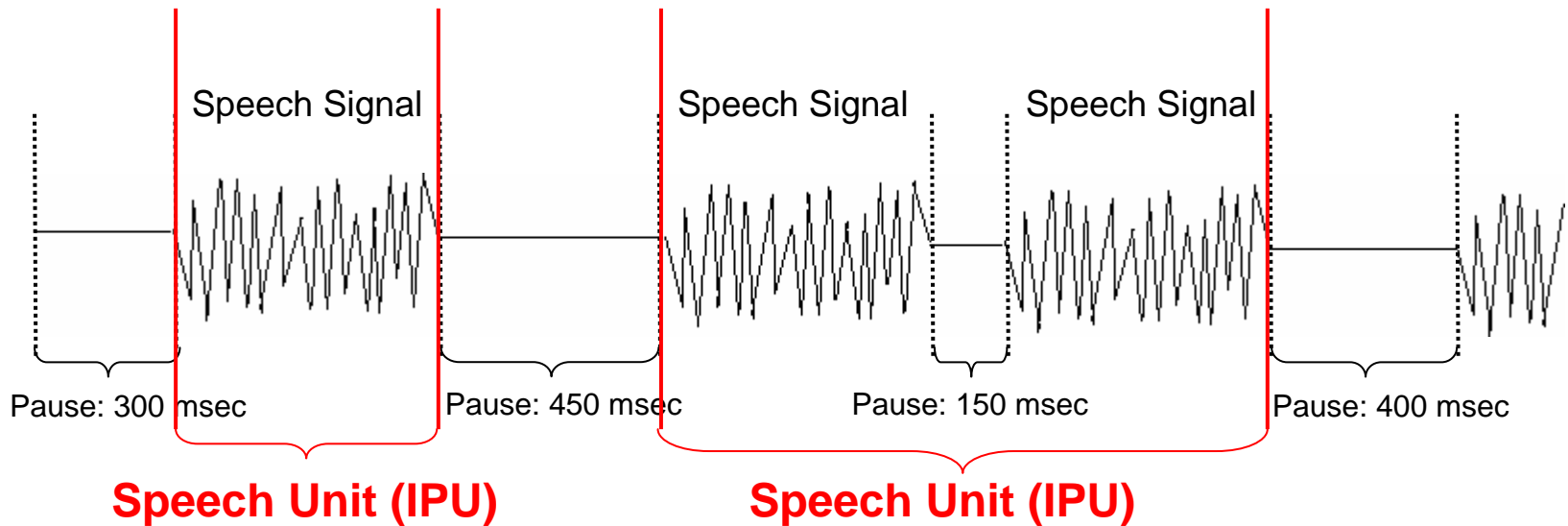
- **Investigation of Exhibitors' behaviors**

- Exhibitors' Looking
- Exhibitors' Pointing
- Exhibitors' Speech



Coding: Speech

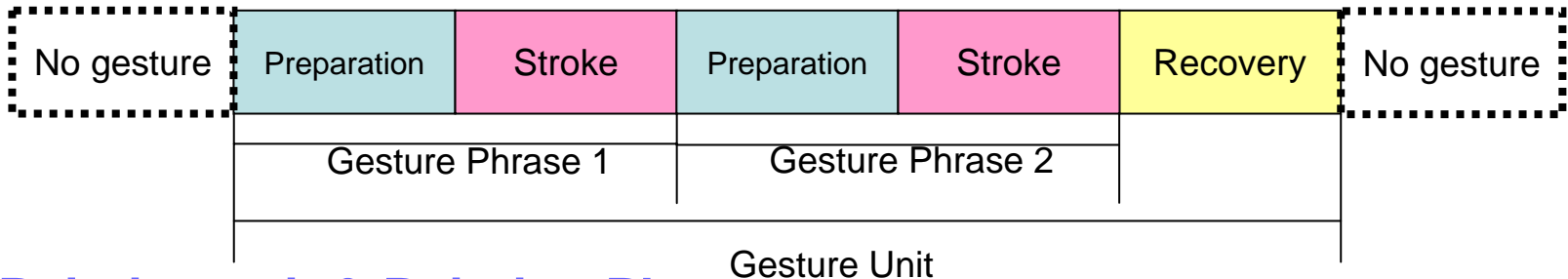
- **Speech: Inter-pausal unit (IPU)**
 - Speech units that are surrounded by pauses whose duration is longer than 300 msec



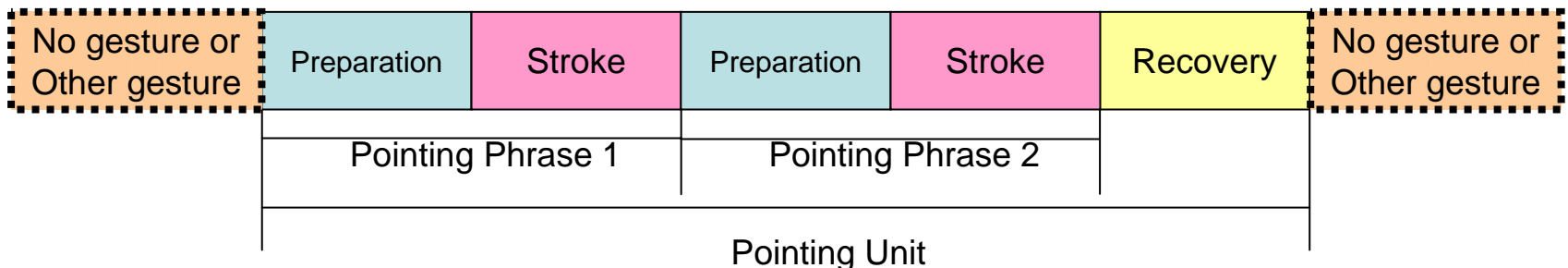
Coding: Pointing



- **Gesture unit (GU) & Gesture phrase (GP) (Kendon, 1972, 1980, 2004)**



- **Pointing unit & Pointing Phrase**



Coding: Looking (1/2)

- **Wearable camera**

- Exhibitors' view image data collected by wearable cameras
- The cameras are mounted on the right side of the head
- These data were used for coding *Looking action*

- **Features of Looking**

- Looking always has a direction.
- Looking does not have a *home position, recovery* and *nothing phases*.



Coding: Looking (2/2)

- **Looking unit consists of one preparation and one stroke.**
 - Start point at preparation
 - Exhibitor's view image starts to move toward the next target in order to look at it
 - Start point of stroke
 - Achieve exhibitor's view image of next target

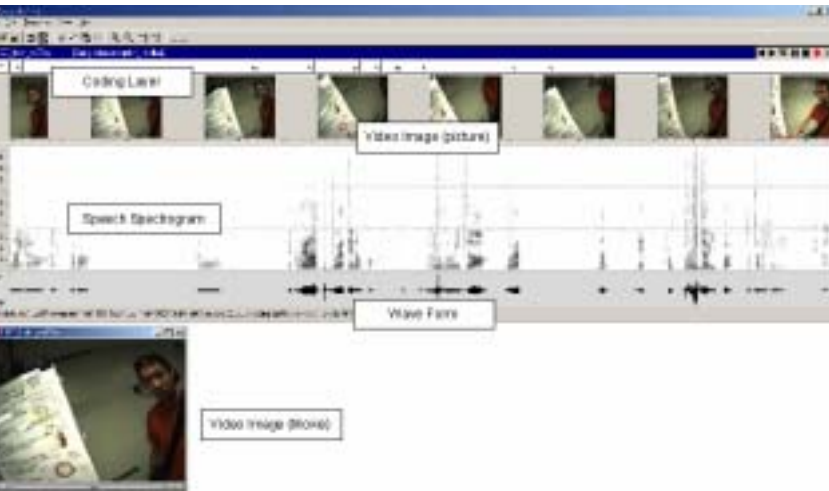


Preparation	Stroke	Preparation	Stroke	Preparation	Stroke	Preparation	Stroke
Look unit: poster		Look unit: audience		Look unit: poster		Look unit: audience	

Tool & Target Data

- **WaveSurfer**

- Open tool for sound visualization and video analysis



- The exhibitor explain about...
Humanoid robot study
- Number of looking and pointing

Time (min)	# of pointing units (mean interval)	# of pointing phrases (mean interval)	# of look units (to poster) (mean interval)
12.4	63 (12 sec)	254 (3 sec)	106 (7 sec)


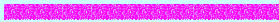
Analyses 1 & 2:

Temporal relations of looking and pointing

- When do we look and point at objects?
In this study,
 - **Looking actions (= look unit)**
 - Exhibitor's behaviors of looking at poster
 - **Pointing actions (= pointing unit)**
 - Exhibitor's behaviors of pointing at poster
- Analysis 1
 - Which action is the first to **start** indexing?
 - Look-start first pattern
 - Point-start first pattern
- Analysis 2
 - Which action is the first to **end** indexing?
 - Look-end first pattern
 - Point-end first pattern

Analysis 1 (a):

Which action is the first to *start* indexing?

	Looking action	<look> Start point of looking at poster
	Pointing action	</look> End point of looking at poster
		<point> Start point of pointing at poster
		</point> End point of pointing at poster

えつまり人間の方/道を教えていただく<look><point><pre>という実験にな<p>ります<re> 0.62

え<pre>と実験の条件の方<p>は<pre> 0.39

(A)

えー教える側が人間で教<p>わる側も</look>人間の場合<look></point> 0.80

教える側が人<point><pre>間で教わる<p>側が協調</look>的</point>身体動作をするロボットの<look>場合<point><pre> 0.67

えーそして教える側<p>が人間で教わる</look></point>側が何も動かない静止しているロボットの場 0.33

(A)

この三つについてそれぞれどのように違うのか<look> 0.44

と<point><pre>いうことをえー比<p>較してみました<re><pre> 0.62

(B)

[Look-start first pattern] Looking action is the first to start

- (A) Looking and Pointing actions occur at the end of IPU
- (B) Looking actions occur at the end of IPU and Pointing actions occur at the beginning of IPU

Analysis 1 (a):

Results & Discussions

Look-start first pattern <look>-<point> sequence	42 cases (71%)
Point-start first pattern <point>-<look> sequence	17 cases (29%)

- Results
 - (A) (B) Looking action is the first to **start**
 - <physical reason> Exhibitors have to catch the area where they want to point before finger pointing gesture.
 - <pragmatic reason> Starting looking at poster near the end of IPU, exhibitors could hold turns.

- Discussion
 - Looking actions are concerned with not only index function but also interactive or interpersonal functions.
- Question
 - Do Pointing actions have index and interactive functions?
- To Analyses 1 (b) & 1(c)
 - To show the difference between Looking and Pointing in conversations

Analysis 1 (b):

Connecting Function in Topic Boundary

えー</look>とその50人の被験者に 0.19

まずある道を歩いてもらってその道を覚えてもらいました 0.61

えーそして次にある部屋に戻ってその部屋の中を人が1人に道を教えていただく 0.51

ということをやっていただきました 0.32

1st Topic

Overview of experime

えつまり人間の方/道を教えていただく<look><point><pre>という実験にも<p>ります</pre> 0.02

え<pre>と実験の条件の方<p>は<pre> 0.39

えー教える側が人間で教<p>わる側も</look>人間の場合<look></point> 0.80

教える側が人<point><pre>間で教わる<p>側が</pre></look>的</point>身体動作をするロボットの<look>場合<point><pre> 0.67

えーそして教える側<p>が人間で教わる</look></point>側が何も動かない静止しているロボットの場合 0.33

この三つについてそれぞれどのように違うのか<look> 0.44

え<point><pre>いうことをえー比較してみました</pre> 0.02

えーそしてその条件の中を自分が歩いた 0.50

えーこのように<pre></look><pre>たのか<look><pre>の条件の中を自分が歩いた</pre> 0.50</look>きたいんですけど

(continued)

3rd Topic

Experiment conditions
(analysis 1(a))

System used in experiment

Analysis 1 (b):

Connecting Function in Topic Boundary

- **At topic boundary**

- Exhibitor used Looking and Pointing actions at topic boundaries (ex. from the end of the 1st topic to the beginning of the 2nd topic)

- **Function**

- Both Looking and Pointing actions have ***Connecting function***
- Exhibitor keeps holding his explanation event by using looking and pointing actions.

Analysis 1 (c):

Connecting Function in Speech Unit Boundary

え<pre>e 実験の条件の方<p>は<pre> 0.39

e<pre>to jikken no jouken no hou<p>ha<pre>0.39

Well, the experimental conditions are...

え<pre>教える側が人間で教<p>えられる側は</look>人間の場合</point> 0.80

e- oshierugawa ga ningen de osowaruh gawa mo ningen no baai

Well, **in the case** where the giver is a human and the follower is also a human...

(1) In a pause

教える側が人<point><pre>間で教<p>えられる側は</look>的</point>身体動作をするロボットの</look>場合<point><pre> 0.07

Oshieru gawa ga ningen de osowaru gawa ga kyouchouteki shintaidousa wo suru robotto no baai

In the case where the giver is a human and the follower is a collaborative-gesture humanoid...

(2) At the end of IPU

え<pre>て教える側<p>が人間で教</look></point>側が何も動かない静止しているロボットの</look>場合 0.33

E-soshite oshieru gawa ga ningen de osowaru gawa ga nani mo ugokanai seishi shiteiru robotto no baai

Well, **in the case** where the giver is a human and the follower is an immovable collaborative-gesture-humanoid...

(3) No looking

この三つについてそれぞれどのように違うのか</look> 0.44

Kono mittu ni tuite sorezore donoyouni chigau no ka

How are three cases different?

え<point><pre>いこうとええ、比べ<p>教してみました</pre><pre> 0.02

To iu koto wo e- hikaku shite mimashita

I compared them.

Semantic Boundary



Analysis 1 (c):

Connecting Function in Speech Unit Boundary

- **At speech unit boundary**
 - **<within semantic unit>** Exhibitor used a looking action in the following pause (ex. (1)) or at the end of speech unit (ex. (2)).
 - **<at the end of semantic unit>** Exhibitor did not use looking and pointing actions at the end of speech (ex. (3)).
- **Function**
 - Looking action has ***connecting function***
 - Exhibitor keeps holding his speech turn within semantic unit by using looking actions .

Analysis 2 (a):

Which action is the first to *end* indexing?

えつまり人間の方/道を教えていただく<look><point><pre>という実験にな<p>ります<re> 0.62

え<pre>と実験の条件の方<p>は<pre> 0.39

(C)

えー教える側が人間で教<p>わる側も</look>人間の場合<look></point> 0.80

教える側が人<point><pre>間で教わる<p>側が協調</look>的</point>身体動作をするロボットの<look>場合<point><pre> 0.67

えーそして教える側<p>が人間で教わる</look></point>側が何も動かない静止しているロボットの場合 0.33

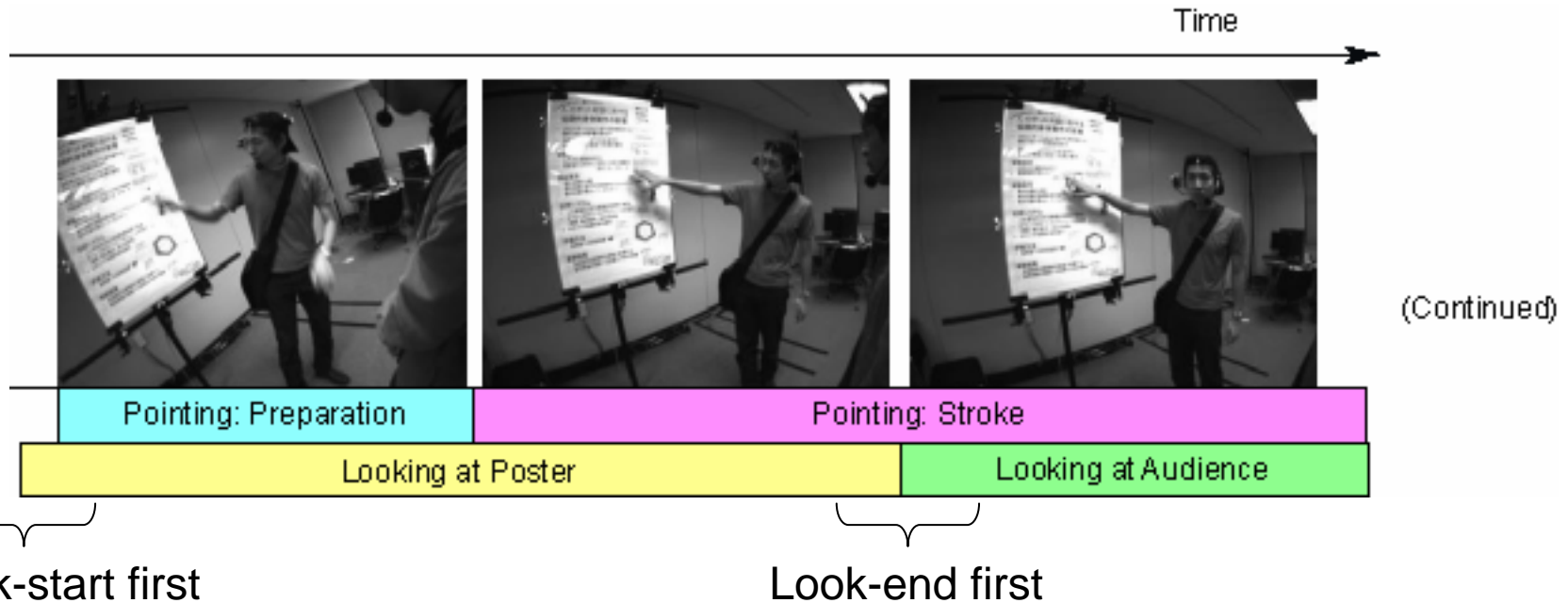
この三つについてそれぞれどのように違うのか<look> 0.44

(C)

と<point><pre>いうことをえー比<p>較してみました<re><pre> 0.62

- [Look-end first pattern] **Looking action is the first to end**
 - (C) Looking and pointing actions occur at the middle of IPU.

Analysis 2 (a): Results & Discussions



- **Results**
 - In some cases, Look-end appears in the middle of speech.
- **Questions**
 - How is Look-end concerned with speech expressions?
- **To Analysis 2 (b)**
 - To show expressions of speech concerned with Looking



Analysis 2 (b):

Looking action as a Boundary signals of Speech Information Structure

- **Hypothesis**

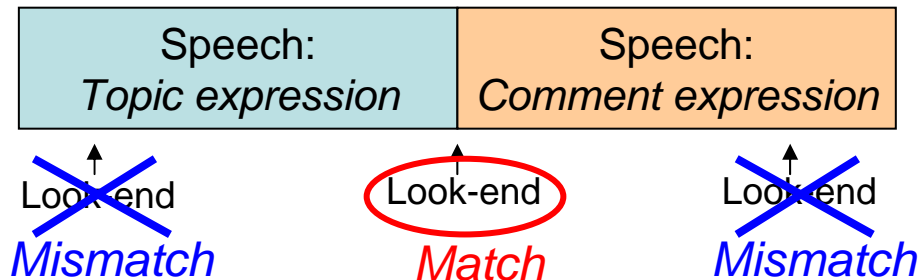
- The “looking at **poster**” phase is concerned with **topic** expression in speech
- The “looking at **audience**” phase is concerned with **comment** expression in speech

- **Procedures**

- Two coders
 - independently examine the transcript
 - Investigate whether “looking at the end point” expresses a boundary of **topic** and **comment** expression in speech.

Analysis 2 (b):

Looking action as a Boundary signals of Speech Information Structure



•EX) Match

この結果は </look>変わってくる かもしれない
Kono kekka ha </look> kawatte kuru kamoshirenai
This result NOM </look> change will might
'This result might change'

EX) Mismatch

この点 協調 </look>ロボットについて...
Kono ten kyouchou </look> robotto nituite...
This point collaborative </look> robot about
'In this point, about collaborative robot....'

• Results

- Looking actions are concerned with speech information structure in 76% of the cases

	Number of Looking	Ratio
Match	55	76%
Mismatch	17	24%

Confidence rate in two coders: 72%

Summary

- **Analysis 1 (a)**
 - Looking action is the first to **start** indexing.
- **Analysis 1 (b)**
 - Both Looking and Pointing actions have **Connecting function** at topic boundary.
- **Analysis 1 (c)**
 - Looking action has **connecting function** at speech unit boundary.
- **Analysis 2 (a)**
 - Looking action is the first to **end** indexing.
- **Analysis 2 (b)**
 - Looking at the audience is a **boundary signals** of the Speech Information Structure.

Concluding Remarks

- *Looking, Pointing* and speech are temporally coordinated in explanatory conversations.
- <Interaction management> Connecting function
 - *Looking* and *Pointing* are used for event-holding.
 - *Looking* is used for turn-holding.
- <Information delivery> Boundary signal
 - The timing of *Looking* is concerned with the information structure in speech expressions

Future directions

- Comparisons with other languages
- Investigate listeners' reactions, which are influenced by differences in speakers' Looking and Pointing expressions
- Investigate individual differences in the use of Looking and Pointing in conversations.
- Compare the use of pointing gestures toward posters with the use of pointing gestures toward Gesture space

Acknowledgment

The research reported here was supported in part by a contract with the National Institute of Information and Communications Technology (NICT) entitled, "A study of innovational interaction media toward a coming high functioned network society".

Thank you for your attention.