

# Fundamentals of Linguistic Interaction

Raquel Fernández

Institute for Logic, Language & Computation  
University of Amsterdam



NASSLLI 2016

# Linguistic interaction



Dialogue is the primary setting for language acquisition and use:

- *spontaneous and online*: disfluent, fragmentary

# Linguistic interaction



Dialogue is the primary setting for language acquisition and use:

- *spontaneous and online*: disfluent, fragmentary
- *multi-agent*: requires coordination (joint action)

# Linguistic interaction

A transcript fragment from the Switchboard corpus:

```
B.52 utt1: Yeah, /
B.52 utt2: [it's,+ it's] fun getting together with immediate family./
B.52 utt3: A lot of my cousins are real close /
B.52 utt4: {C and} we always get together during holidays and
           weddings and stuff like that, /
A.53 utt1: {F Uh, } those are the ones that are in Texas? /
B.54 utt1: # {F Uh, } no, # /
A.55 utt1: # {C Or } you # go to Indiana on that? /
B.56 utt1: the ones in Indiana, /
B.56 utt2: uh-huh. /
A.57 utt1: Uh-huh, /
A.57 utt2: where in Indiana? /
B.58 utt1: Lafayette. /
A.59 utt1: Lafayette, I don't know where, /
A.59 utt2: I used to live in Indianapolis. /
B.60 utt1: Yeah, /
B.60 utt2: it's a little north of Indianapolis, about an hour. /
```

# Overview of topics

- *timing* coordination – turn taking
- *meaning* coordination – dialogue acts and grounding
- *style* coordination - alignment and adaptation
- language *acquisition* in interaction

Slides, links to references, data sources, etc:

<http://www.illc.uva.nl/~raquel/teaching/nasslli2016/>

# Timing coordination: turn taking

Dialogue participants do not only need to make decisions about what to say, but also about *when to say it*  $\rightsquigarrow$  *timing*

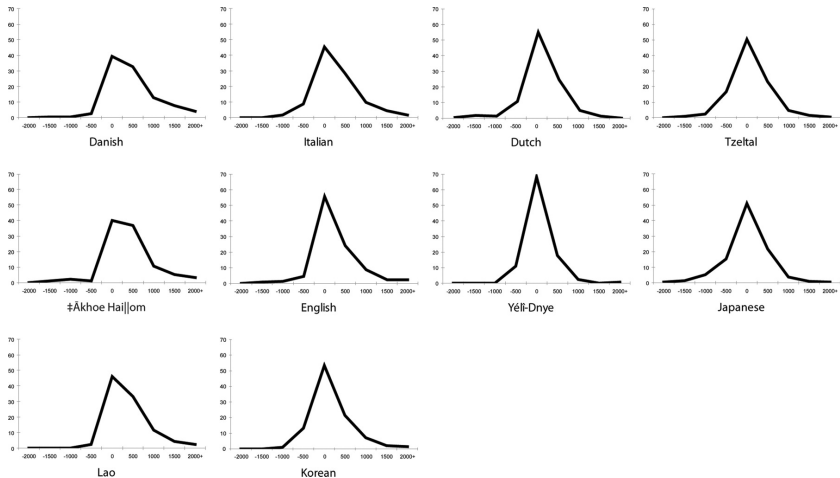
## Outline for this topic:

- Empirical observations: how turn taking works
- Models of turn taking
- Semiotics of timing
- Development and turn taking

# Turn taking: the facts

- Turn-taking is one of the fundamental organisational principles of conversation.
- Learned early: within the first 2 years of life
- There are some individual and cultural differences
- But also strong universal patterns: tendency to minimize both **overlaps** and **gaps** between turns

## Distribution of turn transition length in milliseconds in 10 languages:



Stivers et al. (2009) Universals and cultural variation in turn-taking in conversation, *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*.



# Turn taking: the facts

Turn-taking happens very smoothly:

- *Overlaps are rare*: on average, less than 5% of speech (although there can be a lot of variation).
- *Inter-turn pauses are very short*: ~ 200ms (less than 500ms.)
  - ▶ even shorter than some intra-turn pauses
  - ▶ shorter than the motor-planning needed to produce the next utterance

↪ Turn-taking can't be a reaction to silence

Sacks, Schegloff, & Jefferson, A simplest systematics for the organization of turn-taking in conversation, 1974.

Duncan, Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292, 1972.

Holler, Kendrick, Casillas & Levinson (editors), Turn-Taking in Human Communicative Interaction, *Frontiers in Psychology*, 2015.

# Turn taking: the facts

## Seminal work on turn taking within the framework of Conversation Analysis:

Sacks, Schegloff, & Jefferson (1974) A simplest systematics for the organization of turn-taking in conversation.

### Main empirical observations:

- (a) Overwhelmingly, one party talks at a time.
- (b) Occurrences of more than one speaker at a time are common, but brief.
- (c) Transitions (from one turn to the next) with no gap and no overlap are common. Together with transitions characterized by slight gap or slight overlap, they make up the vast majority of transitions.
- (d) Turn size is not fixed, but varies.
- (e) What parties say is not specified in advance.
- (f) Turn-allocation techniques are obviously used. A current speaker may select a next speaker (as when he addresses a question to another party); or parties may self-select in starting to talk.
- (g) Repair mechanisms exist for dealing with turn-taking errors and violations; e.g., if two parties find themselves talking at the same time, one of them will stop prematurely, thus, repairing the trouble.

# Models of turn taking

- Models based on *prediction*:  
anticipation of the end of the turn.
- Models based on *reaction*:  
response to signal indicating turn yielding.

# Turn taking models: prediction

The CA model by Sacks et al. (1974) emphasises anticipation:

- Turns consist of *turn constructional units* (TCUs) with *projectable* points that can be predicted beforehand.
- Such projectable points act as *transition relevance places* (TRPs) where turn transitions are relevant.

Three rules govern the expected behaviour at TRPs:

1. if devices to select a next speaker (e.g. questions, gaze, naming) are used, the current speaker stops and the selected speaker takes the turn;
2. else, any other speaker may take the turn (may *self-select*),
3. if no other party takes the turn, then the current speaker may continue.

# Turn taking models: reaction

Duncan and colleagues proposed a system of *turn-yielding clues*: the likelihood of a speaker change increases linearly with the number of indicators jointly displayed.

Duncan (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292.

- *turn yielding*: syntactic closure / pragmatic completion plus acoustic information (rising/falling intonation; faster speaking rate); . . .
- *turn-holding*: syntactic incompleteness plus prosodic patterns signal; word fragments and filled pauses.
- From the listener's side: *turn requesting* and *backchannelling* cues.

# Turn taking models: prediction vs. reaction

Recent research has aimed at making all these notions more precise: large scale studies and implementation in dialogue systems.

There is a large amount of literature . . .

Gravano and Hirshberg (2011) Turn-taking cues in task-oriented dialogue, *Computer Speech & Processing*, 5(3).

Magyari and de Ruiter (2012) Prediction of Turn-Ends Based on Anticipation of Upcoming Words, *Frontiers in Psych.*

Prediction-based models are the most common in psycholinguistics. But discussion is ongoing . . .

# Evidence for the models

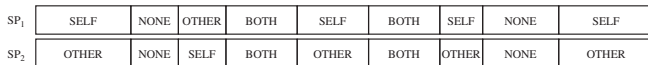
Mattias Heldner and Jens Edlund (2010), Pauses, gaps and overlaps in conversation, *Journal of Phonetics*, 38:555–568.

- Aim to challenge claims about precision timing in turn-taking
  - ▶ “no-gap / no-overlap”  $\rightsquigarrow$  turn-taking must rely solely on the ability to anticipate upcoming turn-endings
- Three corpora: Dutch, Swedish and Scottish English

## 1. VOICE ACTIVITY DETECTION



## 2. COMMUNICATIVE STATE CLASSIFICATION



## 3. SILENCE AND OVERLAP CLASSIFICATION



- Within-speaker overlap not considered in this study

Mattias Heldner and Jens Edlund (2010), Pauses, gaps and overlaps in conversation, *Journal of Phonetics*, 38:555–568.

**Table 5**

Frequencies and percentages of the different types of between-speaker intervals in the Spoken Dutch Corpus (CGN), the HCRC Map Task Corpus (MTC), and the Swedish Map Task Corpus (SMTC).

	Frequency			Percent		
	Gap	Overlap	No-gap-no-overlap	Gap	Overlap	No-gap-no-overlap
CGN	25,844	17,361	169	59.6	40.0	0.4
MTC	8915	6457	115	57.6	41.7	0.7
SMTC	1225	824	11	59.5	40.0	0.5

- striking similarities across languages
- no-gap / no-overlap: less than 1% of between-speaker intervals
- overlaps: about 40% of all between-speaker intervals
- the most frequent kind of between-speaker interval is a slight gap
- 41–45% of intervals with clearly noticeable gaps (> 200ms)
- more pauses than gaps, and with longer durations



# Consequences for turn-taking models

Mattias Heldner and Jens Edlund (2010), Pauses, gaps and overlaps in conversation, *Journal of Phonetics*, 38:555–568.

## Evidence for both projection and reaction theories:

- clear cases where the decision to start speaking cannot be a reaction to interaction control signals near the end of the current speaker's speech (overlaps, gaps shorter than 200ms).
- substantial share of gaps long enough to be reactions to cessation of speech, or to prosodic information just before the silence.

## Interesting speculations:

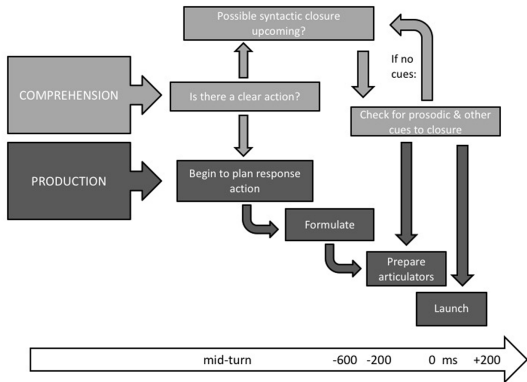
*“reaction to interaction control signals is a plausible explanation for a significant proportion of all speaker changes in human-human conversation”*

*“projection is about content and understanding, rather than about timing.”*

*“reaction is used when the continuation of an utterance is not predictable, or, when the next speaker for some reason wishes to wait until the current speaker has finished and stopped talking.”*

# Prediction + reaction

Stephen C. Levinson and Francisco Torreira (2015) Timing in turn-taking and its implications for processing models of language, *Frontiers in Psychology*.



Sketch of the interleaving of comprehension and production in the recipient of an incoming turn.

# Pragmatic significance of the unexpected

Norm: little overlap, short gap.

- Lengthy *silences* carry semiotic significance (undesired or unexpected response; rhetorical effect)
- *Overlaps* (or interruptions) may be socially loaded (sign of dominance and authority).

# Political debate in Oct 2004

## From LanguageLog ▶ The Rhetoric of Silence

**Jim Lehrer:** Do you believe you could do a better job than President Bush in preventing another 9/11-type terrorist attack on the United States?

**John Kerry:** [pause 0.278] Yes, I do. [pause 1.268] But before I answer further, let me thank you for moderating. [pause 0.588] I want to thank the University of Miami [pause 0.564] for hosting us.

▶ Audio

**Jim Lehrer:** Mr. President, you have a ninety-second rebuttal.

**George W. Bush:** [pause 0.055] uh uh I- [pause 0.165] I, too, thank the University of Miami, and [pause 0.454] and uh [pause 2.116] and say our prayers are with [speeds up] the good people of this state, who've suffered a lot.

▶ Audio

- Empirical facts
- Models: prediction vs. reaction, prediction + reaction
- Semiotics of timing (e.g., rhetoric and social significance)
- Development and turn-taking