Disambiguating Sounds through Context

Maria Niessen, Leendert van Maanen and Tjeerd Andringa

Artificial Intelligence, Sony CSL

August 6, 2008
What do you hear?
What do you hear?

- people describe sound sources, not acoustic properties
- sound perception = sound + knowledge + context

audio retrieved from http://www.psychonomic.org/archive, Marcell et al. (2007)
contents

- background
- experiment
- model
- conclusions
background

- Grouping: (computational) auditory scene analysis
- Context: spreading activation semantic network
background

Dynamic network model for the use of context in sound identification

• spreading activation semantic network can also be applied in a dynamic domain
• what is the role of context in human sound identification
Does context facilitate one of more possible interpretations of a sound?

Known effects of context on human sound identification:

- context facilitates one interpretation \(^1\)
- incongruent context improves correct response \(^2\)
- incongruent context suppresses one interpretation \(^3\)

\(^1\) Ballas and Howard (1987); \(^2\) Gygi and Shafiro (2006); \(^3\) Ballas and Mullins (1991)
experiment: background

Does context facilitate one of more possible interpretations of a sound?

Known effects of context on human sound identification:

- **context facilitates one interpretation** ¹
- incongruent context improves correct response ²
- incongruent context suppresses one interpretation ³

¹ Ballas and Howard (1987); ² Gygi and Shafiro (2006); ³ Ballas and Mullins (1991)
Does context facilitate one of more possible interpretations of a sound?

Known effects of context on human sound identification:

- **context facilitates one interpretation** \(^1\)
- incongruent context improves correct response \(^2\)
- incongruent context suppresses one interpretation \(^3\)

\[
\text{context sound} + \text{homonymous sound} = \text{interpretation}
\]

\(^1\)Ballas and Howard (1987); \(^2\)Gygi and Shafiro (2006); \(^3\)Ballas and Mullins (1991)
Does context facilitate one of more possible interpretations of a sound?

Known effects of context on human sound identification:

- **context facilitates one interpretation**  
  \(^1\)
- incongruent context improves correct response  
  \(^2\)
- incongruent context suppresses one interpretation  
  \(^3\)

\[ \text{context sound} + \text{homonymous sound} = \text{interpretation} \]

\(^1\) Ballas and Howard (1987); \(^2\) Gygi and Shafiro (2006); \(^3\) Ballas and Mullins (1991)
experiment: design

basketball/door + cheering + walking
experiment: design and results

basketball/door  + cheering  + walking

% identification A and B

isolation  context A  context B
model: data

annotations in time-frequency space
model: design

progress

context A
model: design

progress

cheering

context A
model: design

pop concert

cheering

context A

progress

sports game
model: design

progress

pop concert
cheering
context A

sports game
basketball
model: design

context A

cheering

pop concert

sports game

basketball

progress 1
model: design

progress

pop concert —— 0.5 —— cheering —— 1 —— context A

sports game —— 0.5 —— basketball —— 0.8 —— context A
model: design

progress

pop concert

.5

.5

cheering

1

context A

basketball

.8

homonym

sports game
model: design

progress

pop concert

.5

cheering

.5

context A

1

sports game

.8

basketball

homonym

door closing
model: design

progress

pop concert

context A

cheering

1

basketball

homonym

0.5

door closing

0.5

sports game

0.5

0.8
model: design

- pop concert
- context A
  - 1
- cheering
- basketball
  - .5
- sports game
  - .5
- homonym
- door closing

progress

- .5
- .8
model: design

context A

progress

pop concert

cheering

context A

sports game

basketball

homonym

door closing

.5

.5

.8

.5

.5
model: current work

Include more (non-acoustical) contextual factors:

GPS → location

sound signal → loudness

clock → time of day
conclusions

experiment

- context in part determines the perception of sound events
- the effect of context in sound perception is not straightforward

model

- dynamic semantic spreading activation network for automatic sound identification
- context solves ambiguities in dynamic network model

future

- show validity model
- connect to bottom-up grouping
Questions?

(maria@ai.rug.nl)