Adaptation

Vera Demberg

Saarland University

ESSLLI, August 21st, 2014

Today's topic: Adaptation

1. Evidence for **humans adapting** their language production to user and situation.
2. **Adapting to a user**: language comprehension in younger vs. older adults.
3. Ways for a dialog system to **adapt: content and structure**

   (tomorrow: how to adapt natural language generation once decided what to say)

Table of Contents

Background: Language and Driving

Research on driving and language:

- mobile phone usage
  → negative effect on driving
  → “inattention blindness”
- hands-free speaking system
  → negative effect on driving, similar to mobile
- fellow passenger
  → ok

Background: Language and Driving

Research on driving and language:

- mobile phone usage
  → negative effect on driving
  → "inattention blindness"
- hands-free speaking system
  → negative effect on driving, similar to mobile
- fellow passenger
  → ok

Why?

What's the difference between fellow passenger vs. hands-free phone?

Passenger *adapts* to traffic situation

- complexity of speech of driver and passenger lower in difficult driving (Drews et al., 2008)
- shift topic to traffic in difficult driving situation (Villing 2009a,b)
- fewer utterances when driving on city course as opposed to rural route (Crundall et al., 2005)

What's the difference between fellow passenger vs. hands-free phone?

Passenger *adapts* to traffic situation

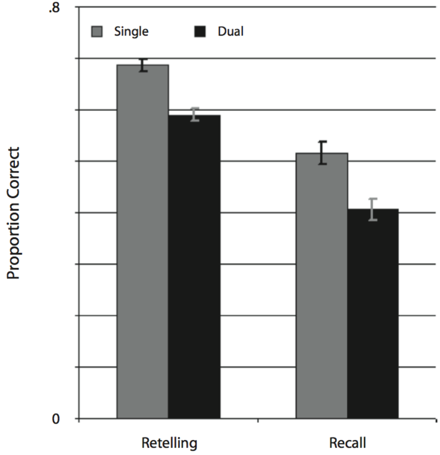
- complexity of speech of driver and passenger lower in difficult driving (Drews et al., 2008)
- shift topic to traffic in difficult driving situation (Villing 2009a,b)
- fewer utterances when driving on city course as opposed to rural route (Crundall et al., 2005)

Dialog systems should be more like passenger drivers and less like the remote conversational partner.

Language comprehension also impaired by driving

When driving, both language comprehension and production are negatively affected. (Becic et al., 2010)

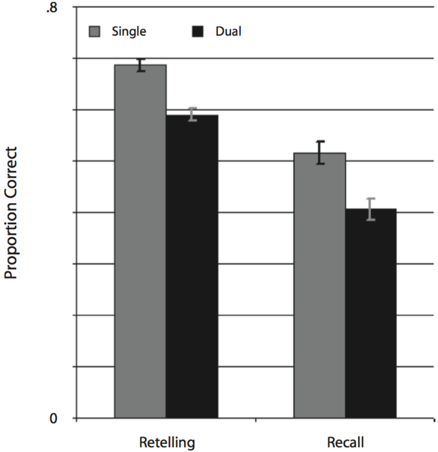
- drivers don't remember as well what they were told



Language comprehension also impaired by driving

When driving, both language comprehension and production are negatively affected. (Becic et al., 2010)

- drivers don't remember as well what they were told
- more language production problems during driving



Language comprehension also impaired by driving

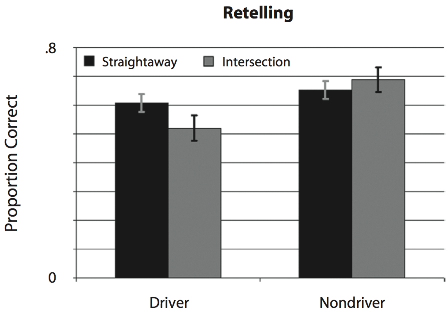
When driving, both language comprehension and production are negatively affected. (Becic et al., 2010)

- drivers don't remember as well what they were told
- more language production problems during driving
- this was true for both younger and older adults.
  (Older adults performed worse in general, but no interaction.)

Language comprehension also impaired by driving

When driving, both language comprehension and production are negatively affected. (Becic et al., 2010)

- drivers don't remember as well what they were told
- more language production problems during driving
- this was true for both younger and older adults. (Older adults performed worse in general, but no interaction.)
- more difficult driving lead to stronger effects.



More on adaptation: Elderspeak

Another example of adaptation to conversational partner:

Elderspeak: shorter and less complex utterances, more filler phrases, more fragments, fewer cohesive cues, slower speech rate and longer pauses

- **beneficial to comprehension:** reduced syntactic complexity; semantic elaborations
- **insulting:** more fragments; more fillers; slow speech

Table of Contents

Adapting to a user

Along which relevant dimensions can users differ?

- familiarity with and knowledge about a task
- dual-tasking abilities / cognitive control
- working memory capacity
- ...

Adapting to a user

Along which relevant dimensions can users differ?

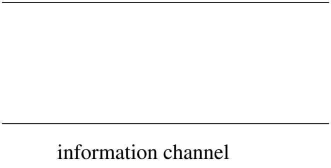
- familiarity with and knowledge about a task
- dual-tasking abilities / cognitive control
- working memory capacity
- ...

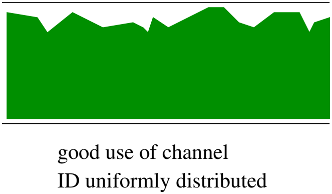
Have seen Surprisal as linguistic model for processing difficulty.

We'll now take the perspective of the brain as an information processing channel.

Information-theoretic perspective

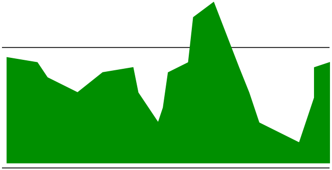
Channel capacity is optimally used when Information Density (ID) is uniformily distributed.



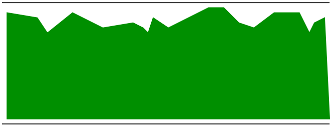


Channel Use

Information channel is not optimally used when Information Density (ID) is very variable.



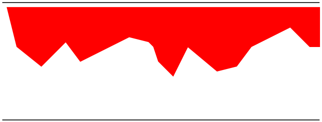
bad use of channel
ID very variable



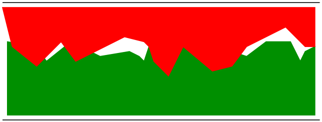
good use of channel
ID uniformly distributed

What happens during dual-tasking?

In dual-tasking, we can think of the channel as partially being taken up by other information.

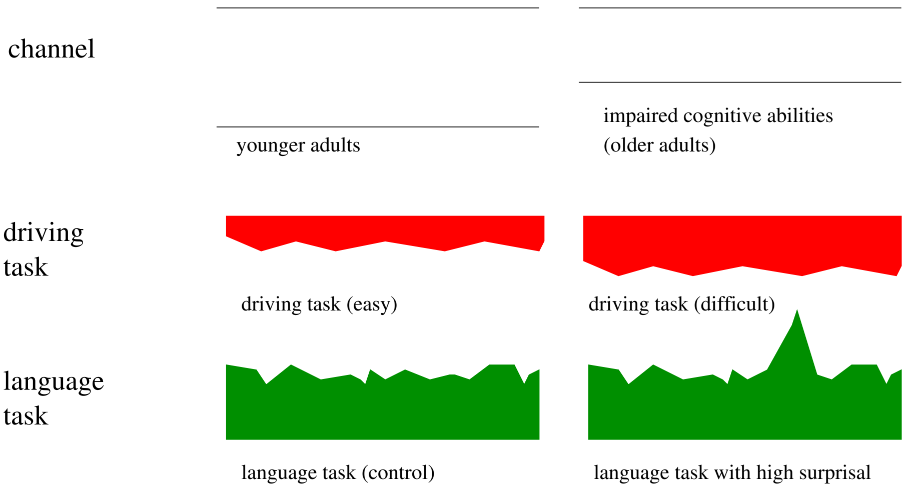


part of the capacity used by other task



if channel is too small, performance on one or both tasks will suffer

Method for testing the effect of adaptation in changing situations and different user groups



Background on Aging

Why elderly people as an example of individual differences?

What happens during aging?

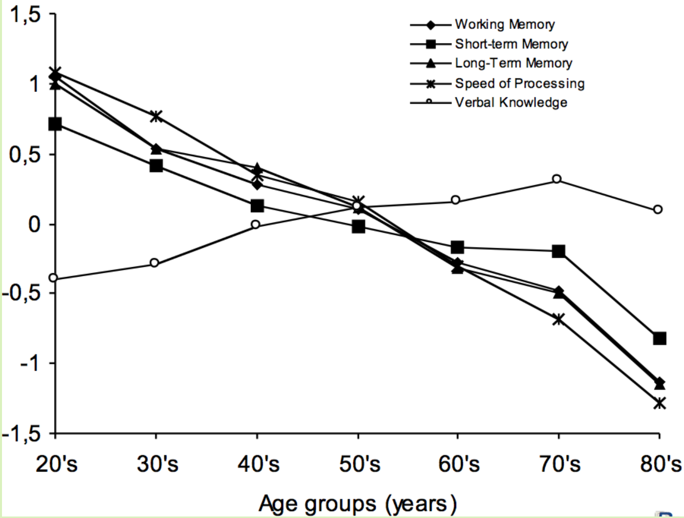
How is language comprehension affected by aging?

How is dual-tasking affected by aging?

Aging and Cognition

- In the cognitive domain, age-related decline can be observed in a variety of cognitive abilities, such as reasoning, working memory, and speed of processing (Kray and Lindenberger, 2007).

Aging and Cognition



(from Park et al., 1996)

Aging and Cognition

- In the cognitive domain, age-related decline can be observed in a variety of cognitive abilities, such as reasoning, working memory, and speed of processing (Kray and Lindenberger, 2007).
- Vocabulary even increases (Ramscar et al., 2014)

Aging and Cognition

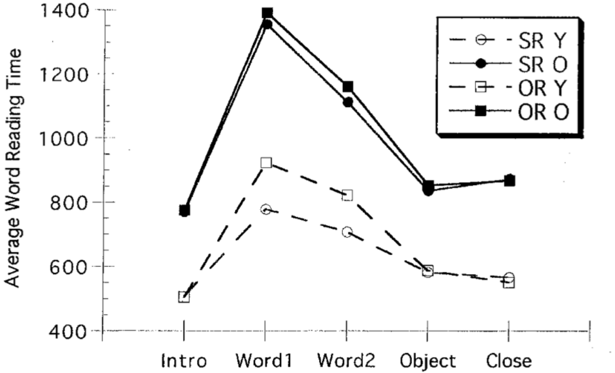
Results on Vocabulary size (Ramscar et al., 2014)

- Ramscar et al., 2014 claim that age decline effects are spurious
- claim: slower reaction times (e.g. on word recognition tasks) just due to larger vocabulary
- idea: brain doesn't work "worse", it just has more data to work with when older, so slower.
- show that older adults actually do have larger vocabulary and knowledge base
- show that larger vocabulary can be predicted to slow down word recognition tasks.

Aging and Cognition

- In the cognitive domain, age-related decline can be observed in a variety of cognitive abilities, such as reasoning, working memory, and speed of processing (Kray and Lindenberger, 2007).
- Vocabulary even increases (Ramscar et al., 2014)
- Culturally-mediated cognitive domains – verbal knowledge and language skills remain relatively stable across the adult lifespan (Kray and Lindenberger, 2007).

Research on aging and language comprehension



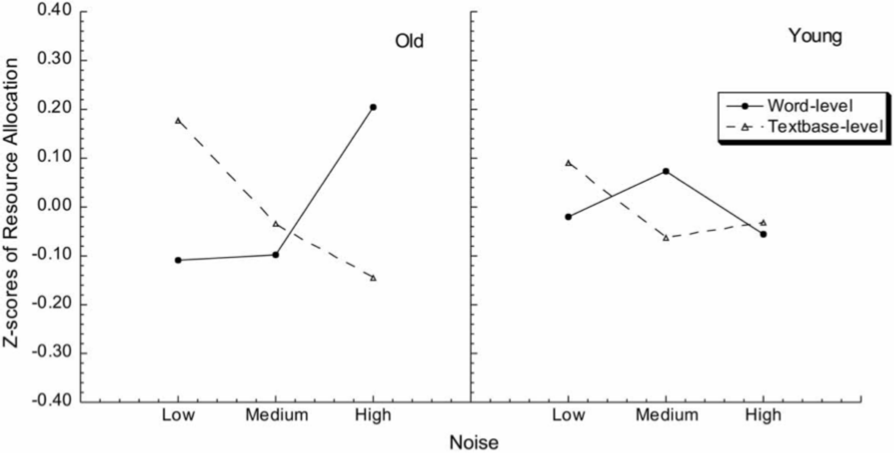
SUBJ REL: The pilot that admired the / nurse / dominated / the conversation / in the bar that evening.
OBJ REL: The pilot that the nurse / admired / dominated / the conversation / in the bar that evening.

- older adults did not slow down at ORC compared to SRC;
- olders more often didn't understand ORCs.

→ elderly not able to allocate necessary resources (Stine-Morrow et al., 2000)

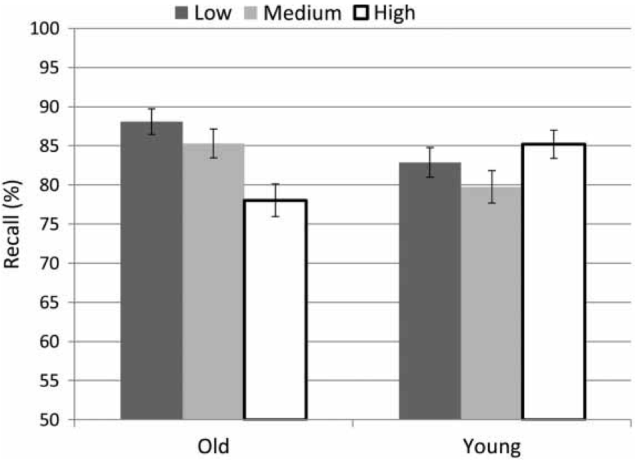
Research on aging and language comprehension

visual clutter leads to poor understanding in old but not in young readers
→ older adults could not allocate enough resources to deeper syntactic / semantic processing (Gao et al., 2012)



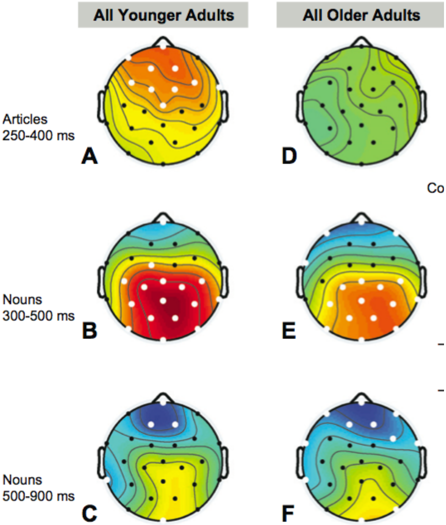
Research on aging and language comprehension

visual clutter leads to poor understanding in old but not in young readers
$\rightarrow$ older adults could not allocate enough resources to deeper syntactic / semantic processing (Gao et al., 2012)



Research on aging and language comprehension

N400 effects in older adults also for predictable items
$\rightarrow$ older adults do not use context cues as effectively
(DeLong et al., 2012)



Age-related decline

Age-related deterioration worst for task-switching and multi-tasking.
(Frensch et al., 1999; Kray and Lindenberger, 2000; Kray, 2006)

Why?

Age-related decline

Age-related deterioration worst for task-switching and multi-tasking.
(Frensch et al., 1999; Kray and Lindenberger, 2000; Kray, 2006)

Why?

- problems due to worse ability to maintain and bias context information or task-relevant information (Braver et al., 2001)
- This ability has been strongly linked to neural efficiency of the dorsolateral prefrontal cortex, a brain region that shows earlier age-related deterioration as compared with other regions of the brain (Raz et al., 2005).

What does this mean for language comprehension?

What does this mean for language comprehension?

- aging $\rightarrow$ decline in cognitive functioning
- ability to maintain task-relevant information deteriorates
- efficiency of language system associated with efficiency of cognitive control processes

What does this mean for language comprehension?

What does this mean for language comprehension?

- aging $\rightarrow$ decline in cognitive functioning
- ability to maintain task-relevant information deteriorates
- efficiency of language system associated with efficiency of cognitive control processes
- **puzzle:** language skills remain relatively stable during aging

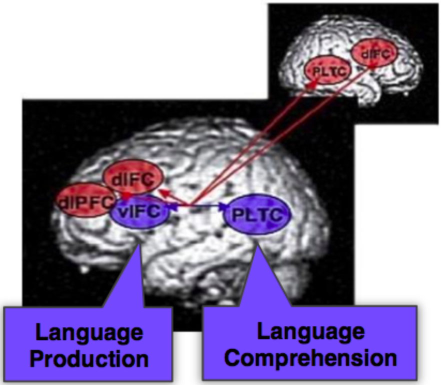
What does this mean for language comprehension?

What does this mean for language comprehension?

- aging $\rightarrow$ decline in cognitive functioning
- ability to maintain task-relevant information deteriorates
- efficiency of language system associated with efficiency of cognitive control processes
- **puzzle:** language skills remain relatively stable during aging
- **key insight:** elderly good comprehenders seem to recruit additional brain regions (cortical regions supporting cognitive control processes) to compensate for age-related decline in maintaining context information (Wingfield and Grossman, 2006)
  $\rightarrow$ these are those areas that are also needed in multi-tasking

Compensation

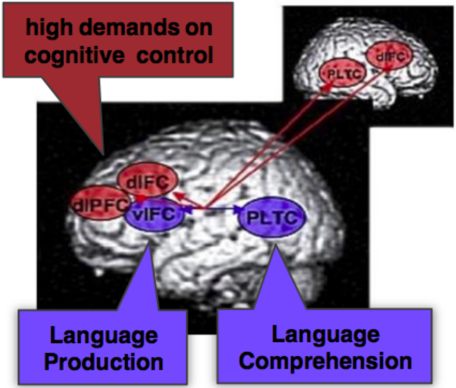
- **Two-component model of sentences comprehension**
  (1) Core sentence network (**Broca's & Wernicke's area**)



Wingfield & Crossman, 2006

Compensation

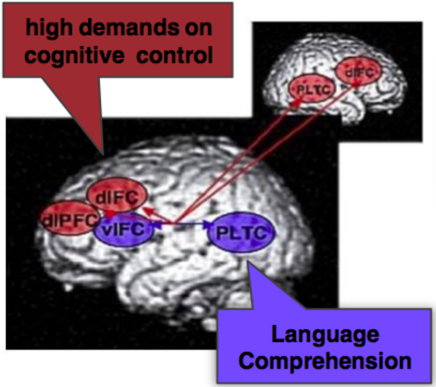
- **Two-component model of sentences comprehension**
  - (1) Core sentence network (**Broca's & Wernicke's area**)
  - (2) Large-scale network (**frontal** & subcortical regions)



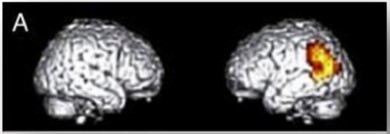
Wingfield & Crossman, 2006

Compensation

- **Two-component model of sentences comprehension**
  - (1) Core sentence network (**Broca's & Wernicke's area**)
  - (2) Large-scale network (**frontal** & subcortical regions)



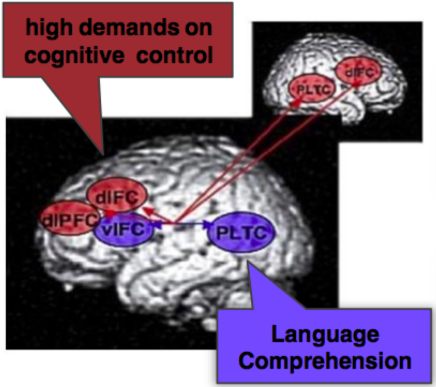
Younger > Older (good performer)

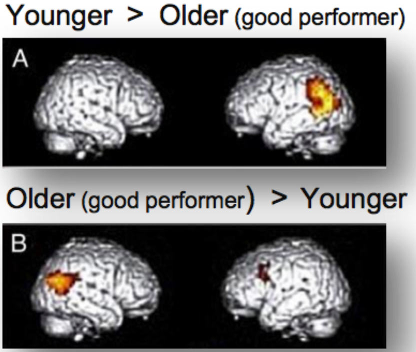


Wingfield & Crossman, 2006

Compensation

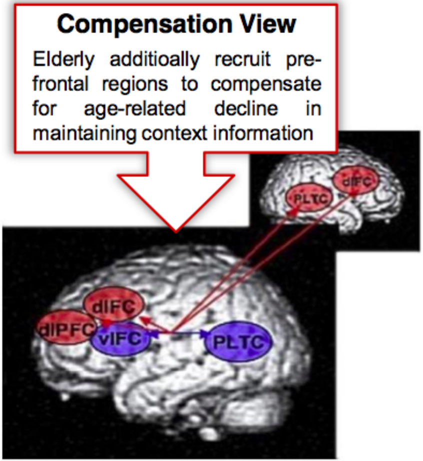
- **Two-component model of sentences comprehension**
  (1) Core sentence network (**Broca's & Wernicke's area**)
  (2) Large-scale network (**frontal** & subcortical regions)

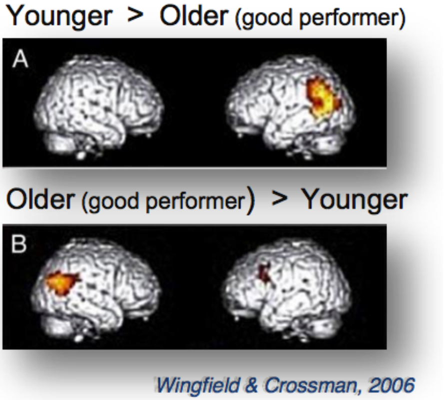




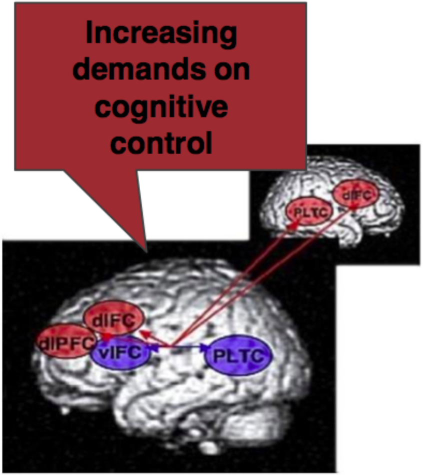
Wingfield & Crossman, 2006

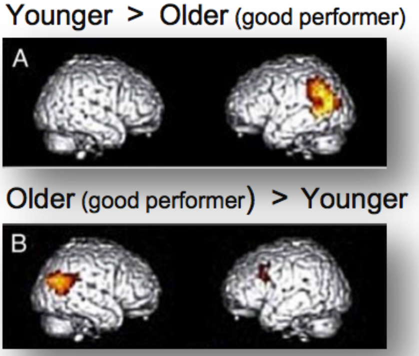
Compensation





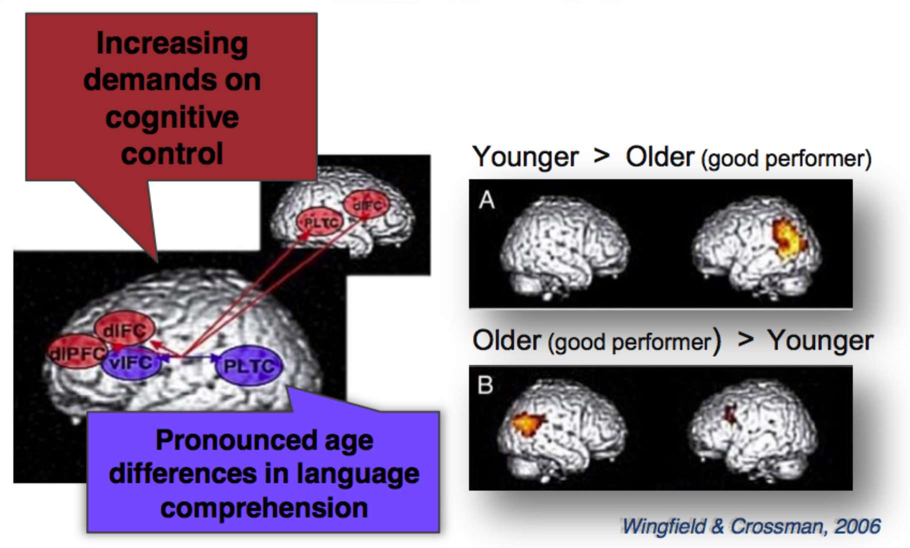
Compensation





Wingfield & Crossman, 2006

Compensation



Conclusions so far

So what does all of this mean for us?

Conclusions so far

So what does all of this mean for us?

- users want to perform well both on driving and language task
- driving can also have a negative impact on language comprehension
- with increasing age, the dual task performance gets worse
- this is expected even for language tasks, as impossible to compensate
- $\rightarrow$ adaptation even more crucial for older adults (or younger ones with similar properties in terms of working memory etc.)

Table of Contents

Some examples for dialog system domains

What kind of dialog systems are we talking about?

- Information presentation
  - Restaurant reservation
  - Flight / train booking
  - Music recommendation
  - Touristic applications (sights)
- Email
- Calendar

Some examples for dialog system domains

What kind of dialog systems are we talking about?

- Information presentation
  - Restaurant reservation
  - Flight / train booking
  - Music recommendation
  - Touristic applications (sights)
- Email
- Calendar

Objectives

Goals in information presentation:

- **Efficient** interaction
- **Effective** presentation (users find options that work better for them)
- Providing a **good overview** of available options
- Making the user feel **confident** in the system

Approaches to presenting multiple options

1. Listing all options until the user makes a choice
2. Filtering the options to only present a subset of options
3. Structuring options to allow the user to navigate through the option space and narrow in on a smaller set of options

Example: Listing all options

S: *I found 28 round trips from Edinburgh to Brussels on October 31st and returning Edinburgh on November 6th.*
For option 1, I have an outbound flight with KLM departing at 7 am and arriving at 8:49 am non-stop. A return flight is with KLM departing at 10:30 am and arriving at 11:28 am with 1 stopover. Ticket price is 165 Euros. Please say "next option", or, "flight details", or, "I'll take it".

U: Next option

S: *For option 2, I have an outbound flight with KLM departing at 3:30 pm and arriving at 5:19 pm nonstop. A return flight is with KLM departing at 8:25 pm and arriving at 12:28 pm with 1 stopover. Ticket price is 165 Euros please say "next option", or, "flight details", or, "I'll take it."*

U: Next option

S: ...

Figure : Typical Information Presentation Phase of a Communicator Dialogue

Filtering options and user modelling

User modelling approach:

- Use some criterion to select a subset of options to be presented
- User modelling: select those options that are most interesting to user
- Elicit user model by asking users for their preferences explicitly or collecting behaviorial data
- Additional advantage: tailor information presentation

(Carenini & Moore, 2000; Moore et al., 2004; Walker et al. 2004)

Adapting content by filtering options

USER: I want to travel from Edinburgh to Brussels, arriving by 5 pm.

For a starving student:

SYS: *There's a direct flight on BMI with a good price. It arrives at four ten p.m. and costs one hundred and twelve pounds. The cheapest flight is on Ryanair. It arrives at twelve forty five p.m. and costs just fifty pounds, but it requires a connection in Dublin.*

For a business traveller:

SYS: *You can fly business class on British Airways, arriving at four twenty p.m., but youd need to connect in Manchester. There's a direct flight on BMI, arriving at four ten p.m., but theres no availability in business class.*

Figure : Content Selection by User Modelling (FLIGHTS system)

Option structuring

"Summarize-and-Refine" approach (Polifroni et al. 2003, Chung 2004)

- Select an attribute with respect to which the option space is least diverse
- Compute clusters of options for this attribute
- Summarize these clusters
- Encourage the user to provide additional constraints

Option structuring: example

SIM:	What restaurants do you have?
SYS:	*I have found 983 restaurants. Most of them are located in Boston and Cambridge. There are 32 choices for cuisine. I also have information about price range.*
SIM:	Okay tell me about the ones in Boston.
SYS:	*I have found 401 restaurants in Boston. There are 29 choices for cuisine.*
SIM:	Do you have any that serve seafood?
SYS:	*I have found 19 seafood restaurants. They are predominantly in Back Bay, the North End, South Boston and the South End.*
SIM:	In the South End.
SYS:	*I have found 3 restaurants. Grill Fish on Columbus Avenue, McCormick & Schmick's on Columbus Avenue and Joseph's on High Street.*

Figure : Content Structuring (system by Polifroni et al.)

Overview of basic approaches

- Listing all options until the user makes a choice
  - inefficient
  - relies on user to remember complex trade-offs between options → places large **memory burden** on user
- Filtering the options (based on UM)
  - doesn't **scale up** to domains with many relevant options
  - does not provide **overview** of option space
- Structuring options (without UM)
  - **exploration** of trade-offs difficult
  - mentions irrelevant entities
  - suboptimal choice of attributes for summarization

Combining user modelling and content structuring

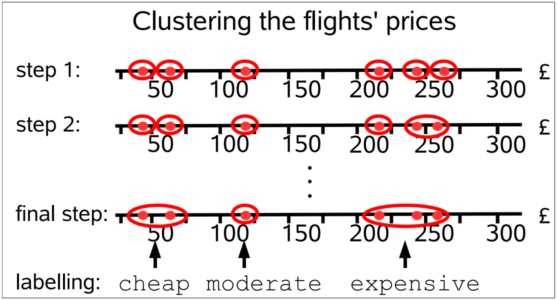
User Modelling with Content Structuring (Demberg et al., 2011)

- Select relevant options
- Structure them based on user's valuations
- Automatically determine tradeoffs
- Tailor summarizations
- Improve overview of options space by briefly summarizing irrelevant options

$\rightarrow$ efficient by information-dense.

Sentence planning

- Talk about options in a way that's relevant to the user
  - continuous values: "cheap", "before 4pm"
  - discrete-valued (few values): "direct"
  - discrete-valued (many values): "none / all of them"



Sentence planning – example

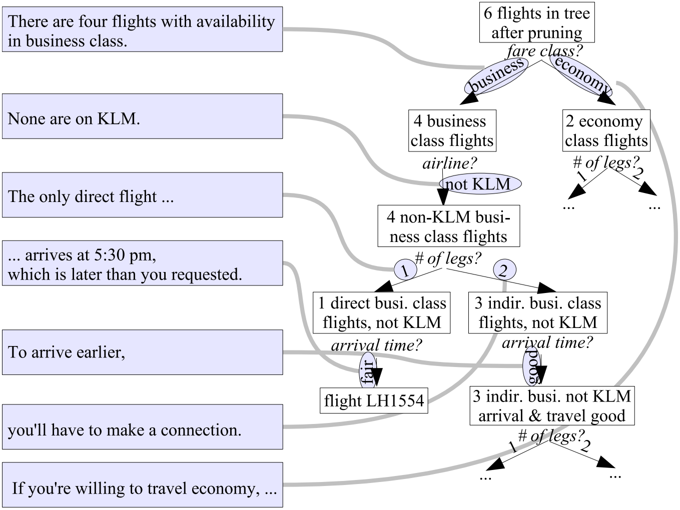


Figure : Diagram showing how the pruned option tree is mapped onto language.

Overview of experiments

- Systems compared: UMSR (user modelling and content structuring) vs. SR system ("summarize and refine", no user modelling)
- 2 experiments
  1. Wizard of Oz experiment: interaction experiment with controlled complexity
  2. dual task experiment: driving a car
- each experiments ca. 40 participants
- 6 dialog pairs (UMSR vs. SR)
- questionnaire with 5 questions after each dialog pair

Evaluation: User Models

the business traveler	He wants, above all, to travel in business class and prefers also KLM.
the student	He cares most about price, everything else being equal
the frequent flier	She collects business miles on KLM and therefore cares most about airline

Table : Example user models used in our experiments.

Questionnaires

1. Did the system give the information in a way that was easy to understand?
   1: very hard to understand
   7: very easy to understand
2. Did the system give X a good overview of the available options?
   1: very poor overview
   7: very good overview
3. Do you think there may be flights that are better options for X that the system did not tell X about?
   1: I think that is very possible
   7: I feel the system gave a good overview of all relevant options.
4. How quickly did the system allow X to find the optimal flight?
   1: slowly
   3: quickly
5. Forced Choice Question:
   Which of these systems would you recommend to a friend?

Experiment 1: Single Task Interaction

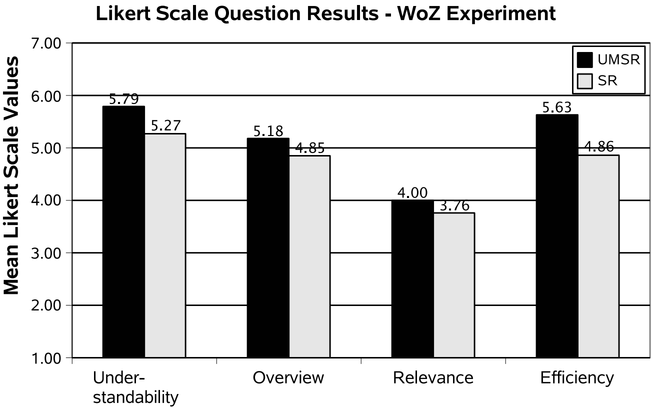
UMSR system allowed users to complete task more quickly.

	SR	UMSR
Turns	14.53**	10.53**
Duration (sec)	391.65**	252.55**

Users more often managed to select the objectively "best" flight with the UMSR system.

	SR	UMSR
Best flights selected	50 (73.53%)*	62 (91.18%)*

Experiment 1: Likert scale questions



Users overall preferred the system using user modelling.

Experiment 2: Dual Task

- Participants driving in a simulator while interacting with the SDS

UMSR system again more efficient than SR system.

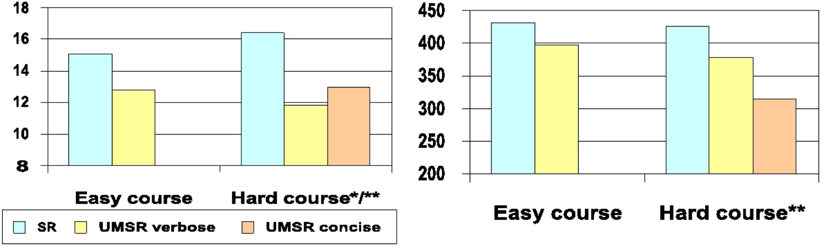
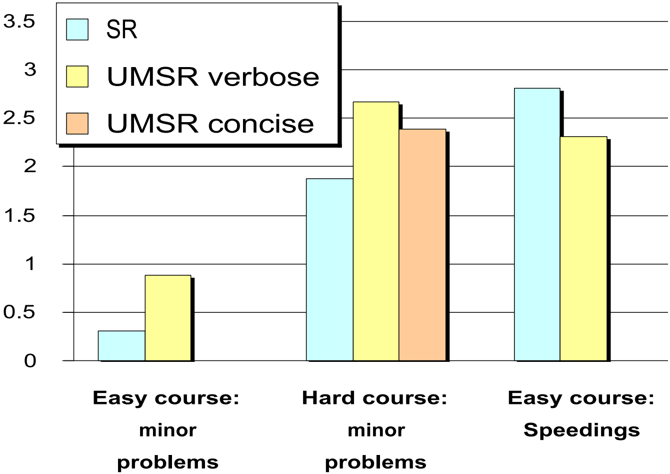


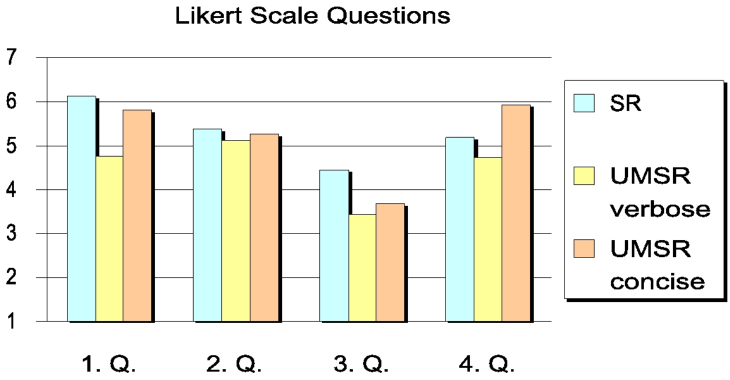
Figure : Task Completion Efficiency: # of dialog turns and dialog duration.

Experiment 2: Dual Task

But the UMSR system led to more driving errors!!



Experiment 2: Dual Task



1. question: understandability
2. question: overview
3. question: relevance
4. question: efficiency

Subjects preferred the SR system in dual task setting!

Evaluation Summary

Overview table for the experiments:

	interation	driving
system pref?	UMSR*	?
understandability	UMSR*	SR
overview	UMSR	SR
relevance	UMSR	SR
efficiency	UMSR*	UMSR
driving errors	NA	same
task success	UMSR*	UMSR*
task duration	UMSR*	UMSR*

Conclusions Information Presentation

Conclusions

- There is no one right way to present information
- User adaptation allows to present relevant content
- Single task: short and efficient is best
- Dual task: user initiative / more redundant is better