What are some of the challenges in dynamic cocktail party listening?

Moritz Wächtler[1], Fabian Wenzel[1], Josef Kessler[2], Martin Walger[1,3], Hartmut Meister[1]

[1]Jean-Uhrmacher-Institute for Clinical ENT-Research, University of Cologne, Germany
[2]Department of Neurology, University Hospital Cologne, Germany
[3]Clinic of Otorhinolaryngology, Head and Neck Surgery, University of Cologne, Germany

Background

In a ***cocktail party situation*** multiple talkers speak simultaneously. Those situations can be:

- *static:* the target talker always remains the same
- *dynamic:* the target talker changes in an potentially unpredictable manner

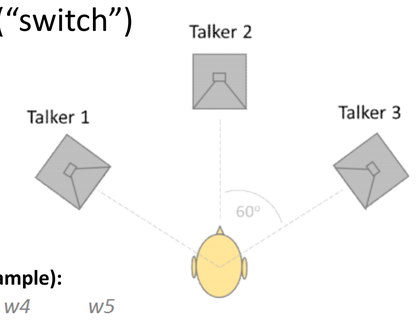
Previous studies have shown that, in dynamic situations, the listener's speech intelligibility often decreases after a switch from one target talker to the other [1].

Research question

Since the moments after a switch appear to be the most challenging in a dynamic cocktail party situation, the question arises: Which factors contribute to the drop in speech intelligibility after a switch?

Methods

- Three competing talkers (high and deep female voice, male voice)
- The target talker is indicated by the word *Stephen* and sometimes changes spatial position ("switch")



Switching type A
Talkers remain at fixed positions but target changes

Talker 1	Talker 2	Talker 3
female high	male medium	female deep
female high	male medium	female deep
female high	male medium	female deep

Switching type B
Talkers change position but target voice remains

Talker 1	Talker 2	Talker 3
female high	male medium	female deep
female deep	female high	male medium
male medium	female deep	female high

Matrix sentences (example):

w1	*w2*	*w3*	*w4*	*w5*
Stephen	buys	eight	wet	cars
Thomas	wins	two	old	stones
Doris	draws	five	green	cups

Switching probabilities:
0% (static), 20%, 100% (dynamic)

Subjects:

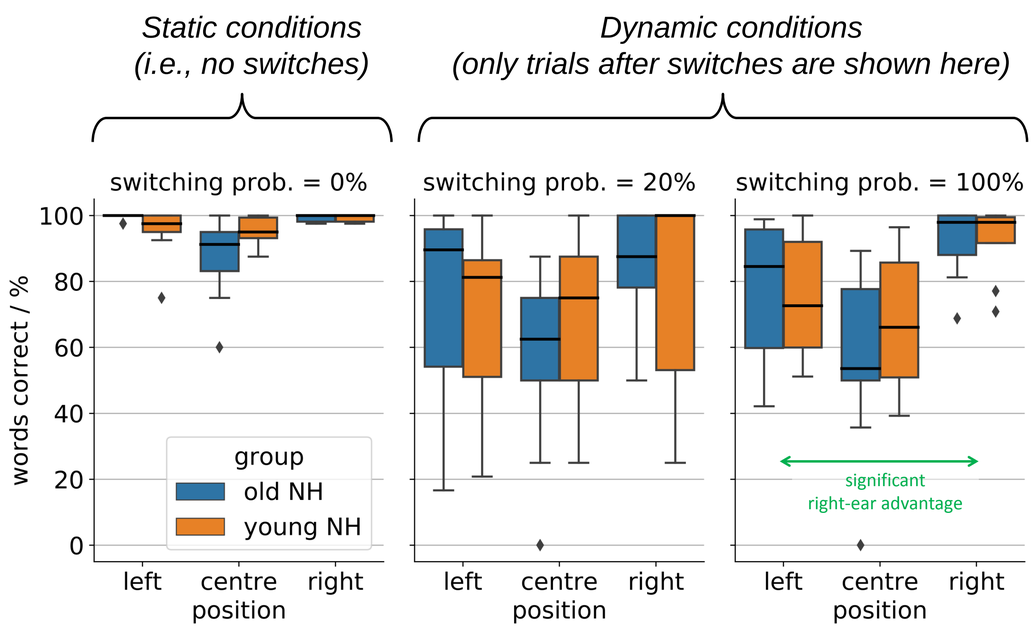
- 14 young normal-hearing (18-33 years)
- 18 old nearly normal-hearing (62-78 years, $BEHL_{0.5\text{-}4kHz} \leq 25$ dB HL)

Results

Challenge 1: Spatial position of target talker

Intelligibility for trials after switches depends on the spatial position of the target talker:

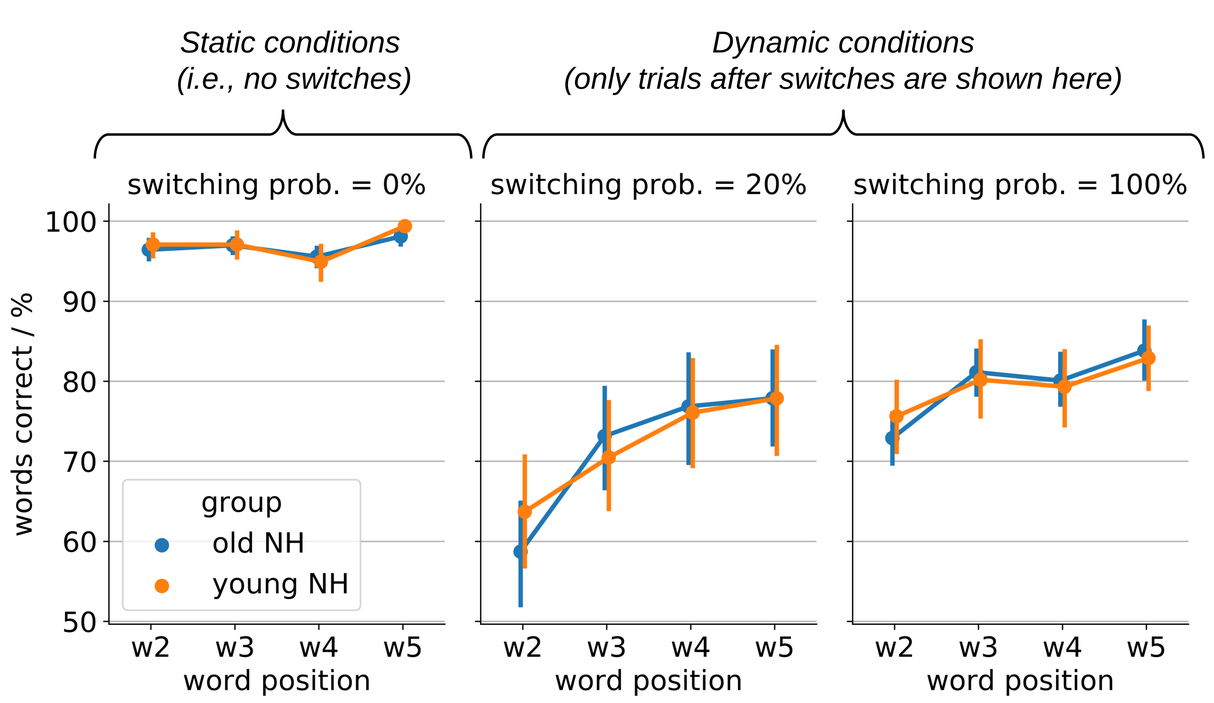
- The centre position is the most challenging.
- In the dynamic conditions (= high cognitive load), a right-ear advantage can be observed.



Only the data of one of the three talkers are shown here to avoid influence of talker voice!

Challenge 2: Sluggish refocusing of attention – Intelligibility for each word position

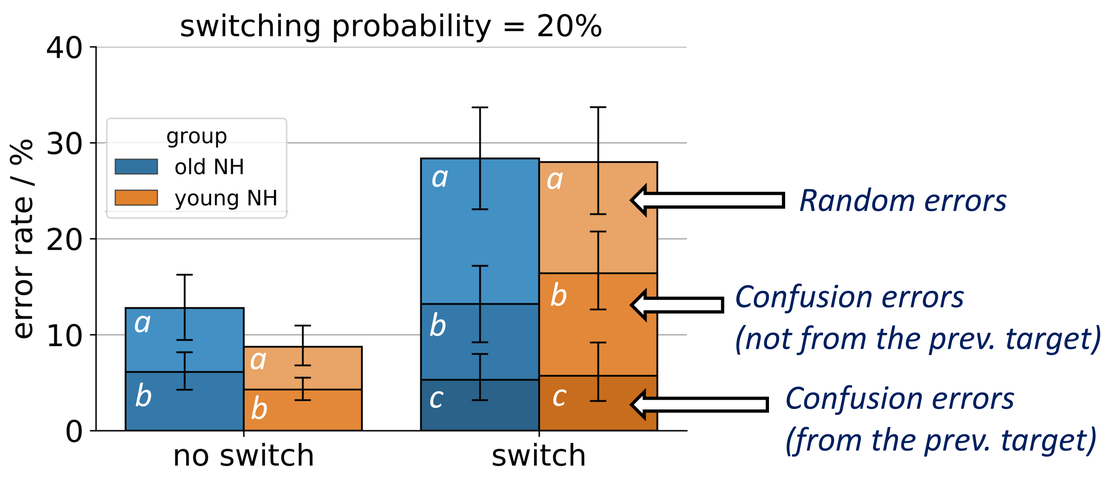
For the static conditions, there is hardly any effect of word position on speech intelligibility.
In the dynamic conditions, the intelligibility is lower for words at the beginning of the sentence. The effect of word position is especially pronounced for a switching probability of 20%.



Challenge 3: Misdirected attention or lost focus? – Error types

Error types were analysed for trials with and without a preceding switch and were categorised as follows:

a) *Random errors:* The subject repeats back a word that is contained neither in the target nor in the masker sentence; or the subject does not repeat back the word at all.
b) *Confusion errors (not from the previous target):* The repeated word is part of the masker and was **not** uttered by the previous target talker.
c) *Confusion errors (from the previous target):* The repeated word was uttered by the previous target talker.



- Ratio of confusion error rate (for "no switch": *b,* for "switch": *b+c*) and random error rate (*a*) shows only slight differences between trials with and without preceding switch.
- Confusions with the previous target talker (*c*) are slightly less frequent than confusions with a previous non-target (*b*).

Discussion and Conclusions

1. **Spatial position**:
   - Lower intelligibility for centre position than for left and right position (perhaps due to lower target-to-masker ratios), especially for trials after a switch. This implies an interaction effect between cognitive load and target position on intelligibility.
   - Right-ear advantage in the cognitively demanding dynamic conditions
2. **Intelligibility for each word position:**
   - Redirecting attention from one target talker to the other does not happen instantaneously and appears to be more difficult when switches are less predictable (switching probability 20% vs 100%).
3. **Error types:**
   - Directing attention to the wrong talker (error type *b* or *b+c*) and losing focus (error type *a*) contribute equally to the error rates (regardless of whether a switch happened).
   - Missing a switch (error type *c*) and thus focusing on the wrong talker accounts for slightly less than half of the confusion errors.

- Hardly any differences between young and old subjects with normal hearing

References

[1] Brungart, D.S., Simpson, B.D., 2007. Cocktail party listening in a dynamic multitalker environment. Percept. Psychophys. Jan;69(1):79-91

Acknowledgements

Supported by Deutsche Forschungsgemeinschaft DFG (ME2751/3-1)

Contact

moritz.waechtler@uhrmacherinstitut.de

Jean-Uhrmacher-Stiftung

UNIKLINIK KÖLN