

1 Introduction

- breathing possible in various ways and combinations
 - air flow direction (in- vs exhalation)
 - airway (oral, nasal, simultaneous oral-nasal, alternations beginning with either oral or nasal)
- breath noise categorization by audio relevant for looking at respiration in detail [1-3], or their acoustic analysis

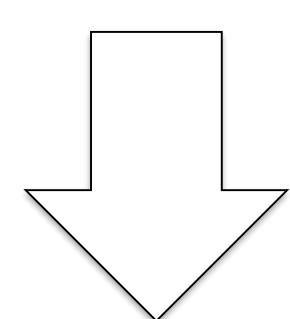
- how reliable is the audio categorization of breath noises?
- does context (+1sec before & after) help?
- are phoneticians better than lay people?
- are there differences by breath noise category?

2 Methods

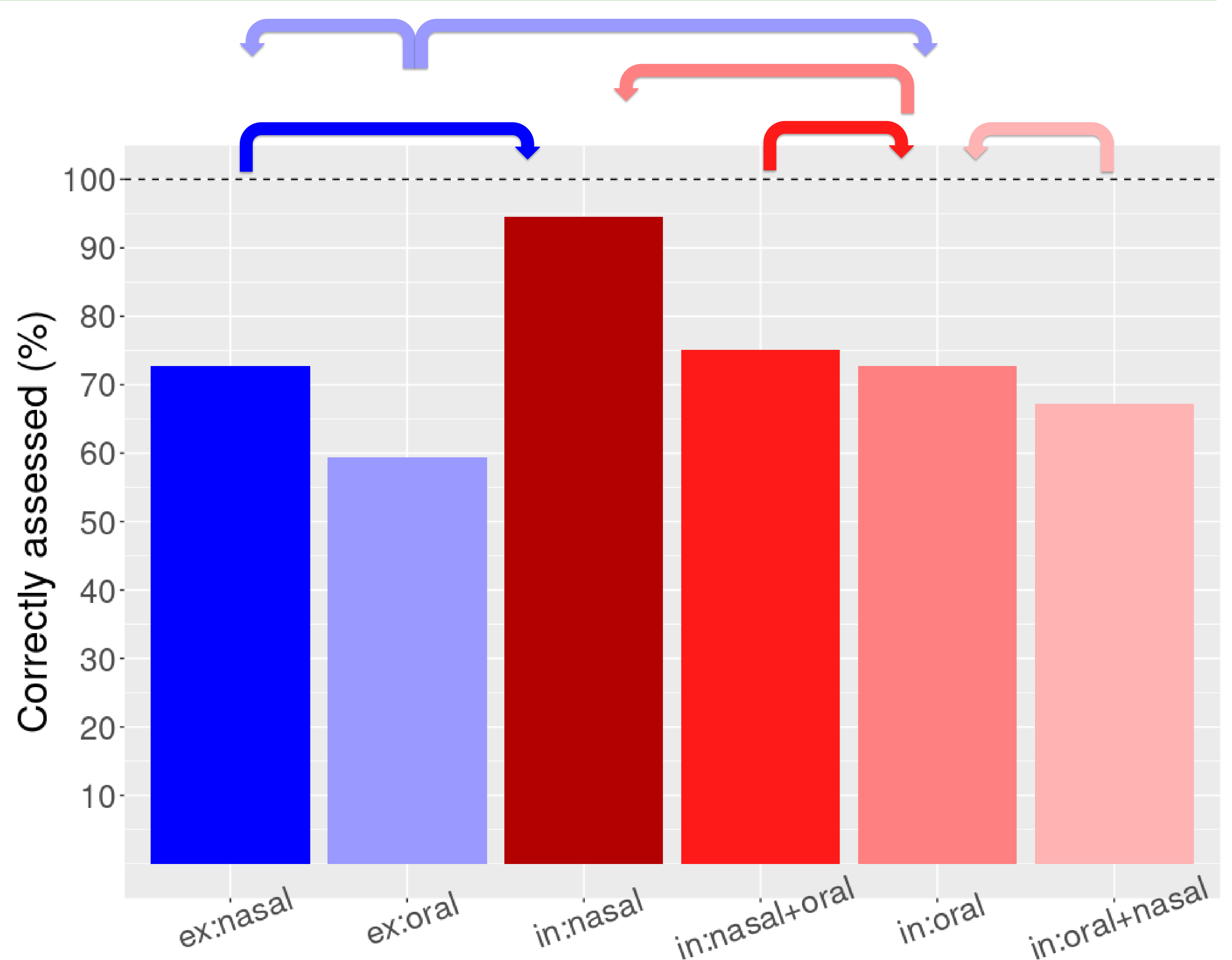
- 20 speakers (10m, 10f) from Dutch audio-visual corpus [4] → mouth opening as cue for oral contribution
- 812 breath noises annotated by 2 raters (inter-rater agreement on 20% subset ≈ 92%, Cohen's $\kappa = .88$)
- 6 frequent types chosen:
 - exhalation: oral, nasal
 - inhalation: oral, nasal, oral+nasal, nasal+oral
- 2 conditions (with/without 1 sec context); randomly selected 4 noises per type & condition
- 48 stimuli assessed by 8 phoneticians & 8 lay people via Labvanced → 768 stimuli in total

3 Results

	correct (%)
overall	73.6
with context	76.8
without context	70.3
phoneticians	74.0
lay people	73.2
ex:nasal	72.7
ex:oral	59.4
in:nasal	94.5
in:nasal+oral	75.0
in:oral	72.7
in:oral+nasal	67.2



- overall ~ 74 %
- with context > without context
- phoneticians ≈ lay people
- no interactions between context & phoneticians
- $in:nasal > in:nasal+oral, in:oral, ex:nasal > in:oral+nasal > ex:oral$



- $in:nasal$ is highest in correctness but also most attractive for other types (biggest migrations from $ex:nasal$ & $in:oral$)
- $ex:oral$ lowest and least attractive for others; loses most towards $ex:nasal$ & $in:oral$
- only little exchange between 'complex' inhalations ($in:nasal+oral$ & $in:oral+nasal$)

4 Discussion & Conclusion

- no difference between experts & lay people
- context may be helpful → on smaller or larger scale?
 - smaller: e.g. nasal inhalations after/before nasal sounds
 - larger: e.g. audible exhalations often appearing outside of fluent speech
- $in:oral$ may be simultaneous oral-nasal inhalations [5]
- studying airway usage difficult
 - reliable ground truth?
 - non-invasive, non-influential measurement?
- overall rate of ~74 % → reliable/usable?

5 References

- [1] Trouvain, J., & Belz, M. (2019). Zur Annotation nicht-verbaler Vokalisierung in Korpora gesprochener Sprache. ESSV 2019, 280-287.
- [2] Kienast, M., & Glitz, F. (2003). Respiratory sounds as an idiosyncratic feature in speaker recognition. ICPhS XV, 1607-1610.
- [3] Scobbie, J. M., Schaeffler, S., & Mennen, I. (2011). Audible aspects of speech preparation. ICPhS XVII, 1782-1785.
- [4] van Son, R. J. J. H et al. (2008). The IFADV corpus: A free dialog video corpus. Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2008, 2(1), 501-508.
- [5] Lester, R. A. & Hoit, J. D. (2014). Nasal and oral inspiration during natural speech breathing. J. Speech, Lang. Hear. Res., vol. 57, no. 3, 734-742.