Using SimpleDIVA to Model Individual Differences in Speakers' Responses to Auditory Feedback Perturbation

Frits van Brenk[1], Hayo Terband[2], & Ben Maassen[3]

[1]Department of Communicative Disorders and Sciences, University at Buffalo; [2]Utrecht Institute of Linguistics – OTS, Utrecht University;
[3]Center for Language and Cognition Groningen (CLCG), University of Groningen

INTRODUCTION

- Previous research showed stronger *compensatory* and *adaptive* responses to sustained auditory perturbation of formants in 4-to-9 year-old children compared to young [1]
- Additionally, considerable between-speaker variation was noted; a significant proportion of both groups did *not show* a response or showed a response that *followed* the perturbation
- Possible explanations:
  - The formant shifts have caused a "target drift", i.e., the sensory motor system interprets the formant shifts as adjustments of the intended auditory outcome [2]. Speakers might employ a strategy not aimed at neutralizing, but rather to match the perceived formant error
  - Developmental differences in reliance on somatosensory feedback versus auditory feedback [3]

PURPOSE

Evaluate age-related differences in feedback and feedforward control in response to auditory perturbation using the SimpleDIVA application

METHODS

PARTICIPANTS

- 50 adults (**ADULTS**):
  32 female, 18 male; age 19-29 years, *M* = 22.3y
- 18 typically developing children (**CTD**) :
  8 female, 10 male; age 4;0 - 7;1 y;m, *M* = 5;2 y;m
- 6 children with speech sound disorders (**CSSD**):
  3 female, 3 male, age 4;8-6;7 y;m, *M* = 5;5 y;m
- All participants were native speakers of Dutch

SimpleDIVA MODELLING

- SimpleDIVA is a 3-free-parameter computational model that estimates contributions of feedback and feedforward control mechanisms and is based on the Directions Into Velocities of Articulators (DIVA) model [4,5,6]
- SimpleDIVA models three subsystems in speech motor control:
  - auditory feedback control (α_A)
  - somatosensory feedback control (α_S)
  - feedforward control/learning rate (λ_{FF}).
- F1 in a trial (n) is the sum of a feedforward command and sensory feedback-based correction:

$$F1_{produced}(n) = F1_{FF}(n) + \Delta F1_{FB}(n)$$

- The feedback based correction is based on both auditory and somatosensory errors detected at the beginning of the production (before feedback control mechanisms contribute; with $F1_T$ as, in principle, invariable target), scaled by the gains of the auditory and somatosensory feedback subsystems α_A and α_S:

$$\Delta F1_{FB}(n) = \alpha_A * (F1_T - F1_{AF}(n)) + \alpha_S * (F1_T - F1_{SF}(n))$$

- The feedforward mechanism for the next trial is updated by adding a fraction (characterized by learning rate parameter λ_{FF}) of the feedback-based corrective command from the current trial:

$$F1_{FF}(n+1) = F1_{FF}(n) + \lambda_{FF} * \Delta F1_{FB}(n)$$

METHODS CONTINUED

MEASURES AND ANALYSES

- The Audapter software module [6] was used for auditory feedback perturbation of target vowel /ɪː/ in CVC words
- F1 was raised 25% and F2 was lowered 12.5%
- Trials involved a baseline – ramp – hold - end paradigm
- Total of 102 trials for adult speakers and 66 trials for children
- Comparison of the 3 model parameters (α_A, α_S, λ_{FF}) in three ways:
  - Whole group analysis: all participants included, regardless of response direction or strength (all groups)
  - Subgroup analysis: participants showing a significant *compensatory*, *following*, or *neutral* response in both F1 and F2 (Adults and CTD)
  - By modelling the speakers' behaviors individually (Adults and CTD)

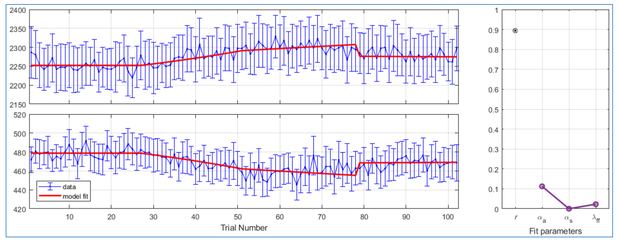
RESULTS - GROUP COMPARISONS

Group	Model fit (Pearson's *r*)			Auditory feedback control gain α_A			Somatosensory feedback control gain α_S			Feedforward learning gain λ_{FF}		
N	Adults	CTD	CSSD	Adults	CTD	CSSD	Adults	CTD	CSSD	Adults	CTD	CSSD
50A - 18CTD - 6CSSD	.74	.73	.01	.040	.008	.001	.000	.008	.000	.024	.999	.999

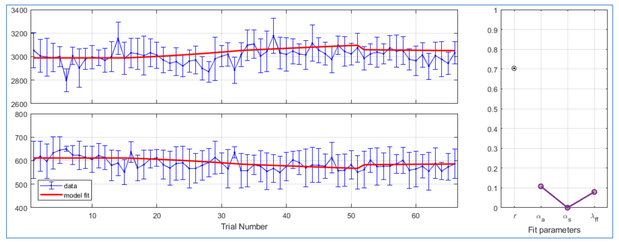
- Reasonable model fits were found for the groups of adults and typically developing children. Data for the group of children with SSD could not be fitted well, possibly due to their weak compensatory responses
- Adults corrected 4% of the detected auditory error, but displayed no somatosensory error correction; typically developing children corrected for less than 1% detected auditory and somatosensory errors; children with SSD did not show corrections
- Adult speakers showed a feedforward learning gain of approximately 2%: the proportion of the correction added from one trial to the feedforward command for the next trial. Both groups of children showed a learning rate approximating 100%, possibly due to their very low correction rates

RESULTS - SUBGROUP ANALYSIS BY RESPONSE

Adults (compensatory subgroup)



Children TD (compensatory subgroup)

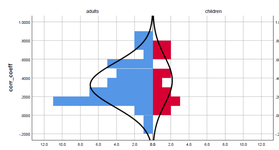


Subgroup		Model fit (Pearson's *r*)		Auditory feedback control gain α_A		Somatosensory feedback control gain α_S		Feedforward learning gain λ_{FF}	
Response	N	Adults	CTD	Adults	CTD	Adults	CTD	Adults	CTD
Compensatory	15A – 4CTD	.89	.70	.113	.108	.000	.000	.023	0.080
Neutral	2A – 1CTD	.09	.25	.010	.046	.320	.000	.999	0.199
Following	5A – 0CTD	NaN	-	.000	-	.450	-	.508	-

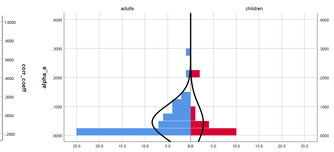
- The group of CSSD could not be further divided in subgroups
- Reasonable to good model fits for subgroups showing an unambiguous compensatory response to perturbation
- Correction for auditory errors for both adults and children is 10%; no somatosensory error correction took place. The feedforward learning rates were higher in children
- As expected, behaviors of subgroups showing neutral or following responses to perturbations could not be modeled successfully

DISTRIBUTION OF INDIVIDUAL SPEAKERS

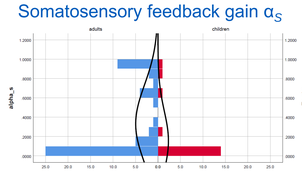
Model fit (Pearson's *r*)



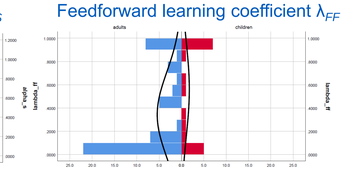
Auditory feedback gain α_A



Somatosensory feedback gain α_S



Feedforward learning coefficient λ_{FF}



- Model fits tend to be better for children compared to adults.
- Trend for differing distributions of auditory and somatosensory feedback between the two age groups
- Feedforward learning rates on average higher in children

SUMMARY AND CONCLUSION

- Without participant preselection, notable but difficult to interpret differences in parameter estimations were found across age groups – possibly due to ambiguous corrective responses
- When including only speakers displaying a significant compensatory effect :
  - Comparable auditory feedback control gain and near-zero somatosensory feedback control gain for both adults and TD children, indicating that auditory-motor integration is similar between these two groups
  - Higher learning rate in children, suggesting a larger speech motor learning plasticity which enables them to rapidly acquire and adapt auditory-articulatory mappings
- Feedforward learning gains are inflated when both auditory and somatosensory feedback control gains are approaching zero.
- SimpleDIVA is not (*yet?*) able to account for responses other than corrective; however the proportion *and* strength of all responses might be a defining and distinguishing characteristic of disordered populations

Results indicate that SimpleDIVA successfully models faster learning rates during speech development, however, both the ***strength of compensatory responses*** and the ***distribution of all responses*** to auditory feedback perturbation might be a defining characteristic in disordered populations, and should be able to be accounted for in models of speech motor control

REFERENCES

[1] H. Terband and F. van Brenk, "Compensatory and adaptive responses to real-time formant shifts in adults and children.," in 18th International Congress of Phonetic Sciences (ICPhS 2015), Glasgow, UK, 2015.
[2] H. Terband, F. van Brenk, and A. van Doornik-van der Zee, "Auditory feedback perturbation in children with developmental speech disorders," Journal of Communication Disorders, vol. 51, pp. 64-77, 2014.
[3] S. Katseff, J. Houde, and K. Johnson, "Partial Compensation for Altered Auditory Feedback: A Tradeoff with Somatosensory Feedback?," Language and Speech, vol. 55, no. 2, pp. 295-308, 2012.
[4] F. H. Guenther, A. Nieto-Castañón, R. Falsini, E. Kearney, and H. R. Weerathunge, "SimpleDIVA (Version 1.1)", 2019.
[5] Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., … & Scott, T. L. (2020). A Simple 3-Parameter Model for Examining Adaptation in Speech and Voice Production. Frontiers in Psychology, 10, 2995.
[6] Kearney, E., Nieto-Castañón, A., Daliri, Y., Guenther, F. H. (2019). SimpleDIVA: A 3-Parameter Model for Examining Adaptation in Speech and Voice Production. Poster presented at 11th Annual Meeting of the Society for the Neurobiology of Language, Helsinki, Finland.
[7] S. Cai, et al., "Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong /iau/ and its pattern of generalization," The Journal of the Acoustical Society of America, 128, 2033-2048, 2010.

Acknowledgements: This study was funded by the Netherlands Organization for Scientific Research (NWO-VENI grant 275-89-016 awarded to the second author). The authors gratefully thank the participants and their parents/caretakers for their time and effort. Margoke Nijssen, Anniek van Doornik-van der Zee, and Noora Somersalo for their help in data collection, and Shanqing Cai for his help with the Audapter software.

brenk@buffalo.edu
H.R.Terband@uu.nl
b.a.m.maassen@rug.nl

@fvbrenk

Department of Communicative Disorders and Sciences
Motor Speech Disorders Laboratory
ubwp.buffalo.edu/msdlab/