5-minute Formant Adaptation Task in Dutch Children

Rebernik, Teja; Tienkamp, Thomas; Polsterer, Katharina; Hukker, Vera; Medvedeva, Masha; van der Ploeg, Mara; Schepers, Iris; Sekeres, Hedwig G.; de Vries, Wietse; Abur, Defne

Published in:
Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS 2023)

5-MINUTE FORMANT ADAPTATION TASK IN DUTCH CHILDREN

Teja Rebernik[1], Thomas B. Tienkamp[1], Katharina M. Polsterer[1], Vera Hukker[1], Masha Medvedeva[2], Mara van der Ploeg[1], Iris Schepers[1], Hedwig Sekeres[1], Wietse de Vries[1], Defne Abur[1], Roel Jonkers[1], Aude Noiray[3,4], Martijn Wieling[1,4]

[1]University of Groningen, [2]University of Leiden, [3]Université PSL, [4]Haskins Laboratories
t.rebernik@rug.nl, t.b.tienkamp@rug.nl, k.m.polsterer@student.rug.nl, raoul.buurke@rug.nl, v.hukker@rug.nl, m.medvedeva@law.leidenuniv.nl, a.m.van.der.ploeg@rug.nl, i.schepers@rug.nl, h.g.sekeres@rug.nl, wietse.de.vries@rug.nl, d.abur@rug.nl, r.jonkers@rug.nl, aude.noiray@ens.psl.eu, m.b.wieling@rug.nl

ABSTRACT

The study investigates whether a short formant perturbation experiment elicits an adaptive response under less controlled experimental circumstances. 30 Dutch children were recruited and tested at a festival. They were asked to produce four target words containing an open-mid front rounded vowel /ɛ/ while we manipulated their feedback so that they would hear /ɪ/ for a period of 16 trials. Despite the short adaptation paradigm, our results show that children significantly changed their vowel productions in response to the perturbation. This suggests that long and monotonous experimental paradigms might not always be necessary, especially with populations that have a shorter attention span.

Keywords: formant adaptation, child speech, speech motor control

1. INTRODUCTION

When speaking, we rely on both feedback and feedforward speech motor control mechanisms, producing sounds in line with a known target and issuing immediate corrections if the produced sound does not correspond to this target (see e.g., [1] for one formal description of the speech motor control system). In the auditory domain, we know what a certain target is supposed to sound like, and we will correct for any mismatches between what we expected to hear and what we actually heard. The auditory *feedback* control subsystem plays a crucial role in child language acquisition, while the auditory *feedforward* control subsystem is more important for adult speakers. Despite this, it is currently unclear how precisely the use of auditory feedback changes throughout childhood.

Formant shifting paradigms (pioneered by Houde and Jordan [2]) can tell us more about articulatory-motor control in children and how they use auditory feedback during speech. Prior studies have shown that children do compensate as a response to formant manipulations but do not do so before the age of 2.5 years [3]. Especially younger children tend to be more variable in their productions compared to adult speakers [4, 3], and at least one study has found a differential effect in F_1 and F_2, with a greater compensatory change in F_1 for younger children, but in F_2 for older children [5]. Additionally, a study with a large sample of 244 either Dutch- or English-speaking children [6] showed stronger compensation in literate compared to pre-literate children.

A recent review by Coughler and colleagues [7] found that the majority of formant perturbation studies in children focus on English native speakers, and rarely examine a broad range of age groups. Furthermore, the experiments are rather long (with a median of 100 trials). This longer duration may not be ideal for children who often have a relatively short attention span.

The goal of our study was therefore to determine whether a short, five-minute, experiment including several target words (instead of the frequently used single target word) is also able to elicit vowel formant adaptation, even when carried out under less-than-ideal experimental circumstances with many potential distractors (see below). We additionally wished to investigate whether age plays a role in how much adaptation is observed. Based on prior studies, we expected the children to exhibit adult-like adaptive responses when exposed to perturbed F_1 and F_2 (e.g., [3]).

2. METHOD

Prior to conducting the study, ethical approval was obtained from the Research Ethics Review Committee of the Faculty of Arts, University of Groningen (CETO; approval number 82182577).

Before participating, parents and children read the information letter, signed the consent form, and filled out a demographics questionnaire together. In the case of younger children (under 12), only the parents did this.

2.1. Participants

Participants were recruited at *Blagenparadijs* - a children's area at one of the biggest music festivals in the Netherlands, *Zwarte Cross*. During three consecutive days, 37 children were recruited, out of which 30 were included in the study. Unfortunately, five children had to be excluded as their consent forms appeared to be incomplete. Additionally, two children were excluded as the data pre-processing stage revealed that their vowel productions were too short to take reliable measurements. The age range of all included children was between 6 and 14 (16 male, and 14 female; see Table 1). All children were native Dutch speakers (1 Dutch-Frisian bilingual). Most children (19 out of 33) were from the province of Gelderland, as that was also where the festival was taking place. None reported any hearing problems, but two reported being dyslectic and an additional two reported to experience stuttering.

Age (in years)	N (F)	N (M)
6	1	-
7	1	1
8	2	5
9	4	4
10	3	3
11	1	2
12	1	-
13	1	-
14	-	1

Table 1: Participants included in the study, separated by sex (M = male, F = female)

2.2. Study design

2.2.1. Procedure and equipment

Data collection took place at the festival grounds, inside a mobile laboratory with a sound-dampened booth (dampened to -40dB[1]). Younger children first completed a short (offline) practice round during which they were asked to practice reading target words. This ensured familiarity with the target words before the experiment began. Children were then positioned in front of a monitor, wearing Sennheiser HD280 Pro headphones to which a Shure WH20XLR headset microphone was attached. Parents generally stayed outside of the van, but were allowed to accompany the child if the child indicated that this would significantly increase their comfort. The microphone and headphones were attached to a Focusrite Scarlett Solo soundcard, which was connected to a HP Z2 Mini G5 Workstation PC.

It is worthwhile to note that the experimental setting itself was controlled — i.e., we tested in a sound booth with high-quality equipment —, but the experimental conditions were not as ideal. This was largely due to the many distractors related to the festival itself: while the children were eager to participate, their attention span was short due to other activities at Blagenparadijs. However, since they were recruited on the spot, this also meant that we were able to include children who otherwise would not have had an opportunity to participate in an experiment at all.

2.2.2. Formant shifting and target words

Formants were shifted in real time using the Audapter software (v2.1.012 [8, 9]). We employed a very short experimental paradigm: the participants began with eight trials of veridical auditory feedback ('START' phase), followed by eight trials during which the formants were gradually shifted until reaching the maximum perturbation ('RAMP' phase). The participants then produced 16 trials at the maximum perturbation of a 20% decrease in F_1 and a 15% increase in F_2 ('STAY' phase), followed by a sudden return to veridical feedback ('END') phase for the final eight trials. The experiment thus consisted of only 40 trials (see Figure 1), which is one of the shortest reported (but see Daliri and colleagues [10] who successfully carried out a 30-trial experiment).

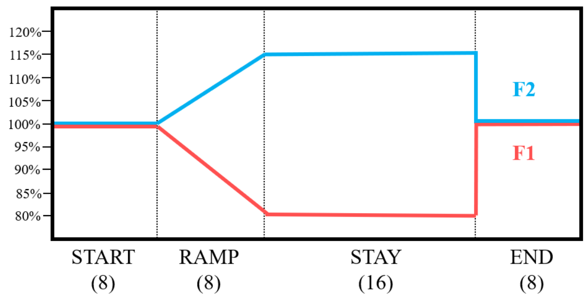

Figure 1: Formant shifting experiment consisting of 40 trials.

Target words in use were four Dutch words with the open-mid front unrounded vowel /ɛ/, namely *pet*, [pɛt] ('cap'); *bed* [bɛt] ('bed'); *dek*, [dɛk] ('deck'); and *bek*, [bɛk] ('beak'). In all cases, the formant shift resulted in another semantically meaningful word: *pit*, [pɪt] ('(fruit) kernel'), bid, [bɪt] ('pray'), dik, [dɪk] ('fat') and bic, [bɪk] (a well-known brand of pens). The frequency of the first three target words and their shifted counterparts is approximately the same (as determined by the SUBTLEX corpus [11]). The fourth target word ('bek') is much more frequent than its shifted counterpart ('bic'). However, also note that while these are high-frequency words for adults, they might not beas well-known by (especially younger) children.

2.3. Analysis

2.3.1. Pre-processing

Raw F_1 and F_2 values were extracted after being automatically detected by Audapter. Mean F_1 and F_2 in a 50ms window (from 30ms after the start of the vowel until 80 ms after the start of the vowel) were calculated for each trial per participant. Average F_1 and F_2 values per trial were then normalized per participant by subtracting the mean of all START trials from each individual trial and dividing by the standard deviation of all START trials. This allowed us to assess how much each individual trial produced by a participant differed from their (average) trials produced during the START phase. In our analysis, we treated trials as a continuum from 1 to 40, and did not separate them into phases.

2.3.2. Statistical analysis

To assess whether children changed their vowel productions during our short formant perturbation paradigm, we used generalized additive mixed modelling (GAMM; [12]). This approach is well-suited for analyzing formant perturbation data, as it does not assume a linear relationship between two variables.

The statistical analysis was done in R version 4.2.2. [13]. We fit GAMMs using the `bam` function of the *mgcv* package version 1.8-41 [12], following the model-fitting procedure outlined by [14]. We interpreted and visualized the results using the *itsadug* package version 2.4.1 [15]. As the data was not normally distributed and showed heavy tails, we fitted a model using the scaled-t distribution (`family = "scat"`), in line with [14].

For our hypothesis-testing model (in which we aimed to assess whether the children adjusted their vowel productions based on the altered feedback), we used the normalized formant value as the dependent variable and trial number as independent variable. Specifically, we included a (non-linear) smooth over trial, separated by formant type (F_1 and F_2). We included the maximal random effects structure supported by the data, which included a factor smooth over trial (per formant type) for each participant and a factor smooth over trial (per formant type) for each target word, accounting for variation in the non-linear patterns across trials per participant and target word, respectively. In our exploratory analysis, we additionally assessed the possible effect of age and gender.

3. RESULTS

The hypothesis-testing model showed a significant non-linear effect of trial (p<.001) for both formants. Children therefore adapted in the expected manner, by increasing their F_1 in response to the 20% F_2 decrease and by decreasing their F_2 in response to the 15% F_1 increase. The pattern of F_1 and F_2 can be seen in Figure 2.

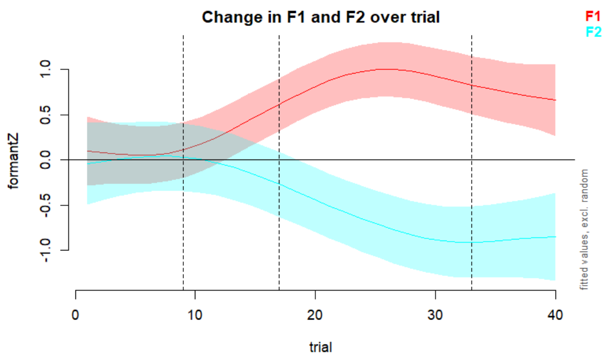

Figure 2: Change in F_1 and F_2 over trial. Shaded areas correspond to 95% confidence bands.

We continued with the exploratory analysis, evaluating the effect of gender and age, as indicated above. We first assessed the effect of gender, separated by formant type (F_1 and F_2). The results indicated no significant difference between male and female participants (p>0.1 for both F_1 and F_2). We then assessed whether or not age showed a significant (potentially non-linear) effect per formant type. This revealed no effect of age (p>0.1 for both F_1 and F_2). Finally, we assessed whether age and trial interacted by including a tensor product interaction for trial and age, separated by formant type. This likewise showed no significant effect of age (p>0.1 for both formants).

4. DISCUSSION

Our study investigated adaptive responses to formant perturbations in a group of 30 Dutch children, aged between 6 and 14. Our goal was to determine whether children adapt to a short experimental paradigm of five minutes when tested at festival grounds. In addition, we investigated whether age is a predictor in how much they adapt (per formant). Our analysis, using generalized additive mixed modeling, showed a significant change in F_1 and F_2 depending on trial ($p > .001$). There was further no effect of neither gender nor of age.

In line with prior studies (e.g., [6], we would have expected some differences in age, especially between the younger participants (aged 6 or younger) and the older participants (aged 9 or older). While plots for F_2 per year did seem to show a trend towards a larger decrease in F_2 for older participants (i.e., more adaptation), this was not statistically significant ($p > 0.5$). Unfortunately, there was a large imbalance in our participants' ages (we only included a single 6-year-old and a single 14-year-old, for example), making any age comparison unreliable.

Our study faces several limitations. First, as stated above, while we tried to recruit children in a broad age range, most children were aged between 8 and 11. Second, while all children could read, not all felt equally comfortable reading during the experiment (although they did complete a short offline training session beforehand), which might have affected the results. Finally, while a short paradigm does work in eliciting adaptive responses, fewer productions can be problematic if any data needs to be discarded. In our case, we had to exclude two participants because there were not enough trials of sufficient length available.

Despite the variability present in children's speech, our results suggest that future studies may employ a similarly shortened paradigm to investigate responses in large(r) groups of children, including teens (ages 10 and up). While testing at large events is not always ideal, it ensures that children of different ages and backgrounds are more easily recruited.

---

[1] For more information on the mobile laboratory, SPRAAKLAB, please see [16]

5. REFERENCES

[1] J. A. Tourville and F. H. Guenther, "The diva model: A neural theory of speech acquisition and production," *Language and Cognitive Processes*, vol. 26, no. 7, pp. 952–981, 2011.

[2] J. H. Houde and M. I. Jordan, "Sensorimotor adaptation in speech production," *Science*, vol. 279, no. 5354, 1998.

[3] E. MacDonald, E. Johnson, J. Forsythe, P. Plante, and K. Munhall, "Children's development of self-regulation in speech production," *Current Biology*, vol. 22, no. 2, pp. 113–117, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960982211013340

[4] F. van Brenk and H. Terband, "Compensatory and adaptive responses to real-time formant shifts in adults and children," *The Journal of the Acoustical Society of America*, vol. 147, no. 4, pp. 2261–2270, 2020.

[5] S. T. Cheung, K. Thompson, J. L. Chen, Y. Yunusova, and D. S. Beal, "Response patterns to vowel formant perturbations in children," *The Journal of the Acoustical Society of America*, vol. 150, no. 4, pp. 2647–2654, 2021.

[6] M. R. van den Bunt, M. A. Groen, S. Frost, A. Lau, J. L. Preston, V. L. Gracco, K. R. Pugh, and L. T. W. Verhoeven, "Sensorimotor control of speech and children's reading ability," *Scientific Studies of Reading*, vol. 22, no. 6, 2018.

[7] C. Coughler, K. Quinn de Launay, D. Purcell, J. Oram Cardy, and D. Beal, "Pediatric responses to fundamental and formant frequency altered auditory feedback: A scoping review," *Frontiers in Human Neuroscience*, vol. 16, 2022.

[8] J. A. Tourville, S. Cai, and F. H. Guenther, "Exploring auditory-motor interactions in normal and disordered speech," in *Proceedings of the 165th Meeting of the Acoustical Society of America, Montreal, Quebec, Canada*, 2013.

[9] S. Cai, M. Boucek, S. S. Ghosh, F. H. Guenther, and J. S. Perkell, "A system for online dynamic perturbation of formant frequencies and results from perturbation of the mandarin triphthong /iau/." in *Proceedings of the 8th International Seminar on Speech Production, Strasbourg, France,*, 2008, pp. 65–68.

[10] A. Daliri, E. A. Wieland, S. Cai, F. H. Guenther, and S.-E. Chang, "Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter," *Developmental Science*, vol. 21, no. 2, p. e12521, 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/desc.12521

[11] E. Keuleers, M. Brysbaert, and B. New, "Subtlex-nl: A new measure for dutch word frequency based on film subtitles," *Behaviour Research Methods*, vol. 42, no. 3, pp. 643–650, 2010.

[12] S. N. Wood, *Generalized Additive Models: An Introduction with R*. Taylor & Francis, 2017.

[13] R Core Team, *R: A Language and Environment for*

Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2022. [Online]. Available: https://www.R-project.org/

[14] M. Wieling, "Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between l1 and l2 speakers of english," *Journal of Phonetics*, vol. 70, pp. 86–116, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0095447017301377

[15] J. Van Rij, M. Wieling, R. Baayen, and H. van Rijn, "itsadug: Interpreting time series and autocorrelated data using gamms," 2017. [Online]. Available: https://cran.r-project.org/web/packages/itsadug

[16] M. Wieling, T. Rebernik, and J. Jacobi, "Spraaklab: a mobile laboratory for collecting speech production data," in *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS 2023)*, 2023.