STEM-, SPRAAK- EN TAALPATHOLOGIE

Vol. 27, 2022, Special Issue, pp. 69-92
Special Issue for Steven Gillis

Gepubliceerd online: 5 november 2022

University of Groningen Press
https://doi.org/10.21827/32.8310/2022-SG-69

F0 dynamics associated with prominence realisation in children with hearing impairment

Jérémy Genette[1] & Jo Verhoeven[1,2]

[1] *Computational Linguistics & Psycholinguistics (CLiPS) Research Center, University of Antwerp*
[2] *Department of Language and Communication Science, City, University of London*

Samenvatting

Prominentie in spraak wordt typisch gerealiseerd door een grotere amplitude, duur en F0 van de klinkers in lettergrepen met zinsaccent/woordaccent. Het is algemeen bekend dat F0 de belangrijkste dimensie is. Deze studie doet verslag van de analyse van F0 in de uitspraak van woorden van twee groepen kinderen die verschillen in hoorstatus. De ene groep bestond uit kinderen met een gehoorbeperking, terwijl de andere groep bestond uit leeftijdsgenoten met een normaal gehoor. De kinderen met een gehoorbeperking hadden een cochleair implantaat of een conventioneel hoortoestel. De kinderen namen deel aan een woordimitatietaak die bestond uit het herhalen van monosyllabische woorden met een van de klinkers van het Belgisch Standaardnederlands. Uit de studie blijkt dat er interessante F0 verschillen tussen de groepen zijn. De kinderen met gehoorbeperking hebben de hoogste gemiddelde F0. Wat betreft de dynamiek van F0 realiseerden alle kinderen prominentie met een onderliggend stijgend-dalend patroon dat zich op fonetisch niveau manifesteerde als een dalende toonhoogtebeweging. Bovendien was de contour bij kinderen met een conventioneel hoortoestel het steilst, terwijl hij bij kinderen met een cochleair implantaat het vlakst was. De toonhoogtecontouren van kinderen met een normaal gehoor lag tussen de twee voorgaande groepen in. De waargenomen verschillen worden toegeschreven aan de aard van het hoortoestel.

Abstract

Prominence in speech is typically realised by means of greater amplitude, duration and F0 of the vowel nucleus in the syllable that carries word/sentence stress. It is well-established that F0 is the more important physical dimension. The present study reports the analysis of F0 in word realisation of two groups of children differing in hearing status. One group consisted of children with hearing impairment, while the other group consisted of age-matched children with normal hearing. The hearing-impaired children had been fitted with either a cochlear implant or a conventional hearing aid. Children had participated in a (non-)word imitation task which consisted of the repetition of

Correspondentieadres:
Jérémy Genette
University of Antwerp
Computational Linguistics and Psycholinguistics (CLiPS) Research Center
Prinsstraat 13, 2000 Antwerpen, Belgium
E-mail: jeremy.genette@uantwerpen.be

monosyllables containing one of the monophthongs of Belgian Standard Dutch. Measurement and analysis of F0 in the vowel nuclei revealed interesting differences between the groups. The children with hearing impairment had the highest overall F0. In terms of the dynamics of F0 associated with prominence, all children correctly realised an underlying prominence-lending rise-fall pattern which at the phonetic level manifested itself as a falling pitch movement. In addition, the contour in children with a conventional hearing aid was steepest, while it was shallowest in children with a cochlear implant. The contour in children with normal hearing was situated between the two previous groups. The observed differences are attributed to differences in device use.
Keywords: prominence, fundamental frequency, prosody, language acquisition, hearing impairment.

Introduction

Prosody in speech is very important to convey linguistic and paralinguistic information and to generally support efficient communication. Prosody is an umbrella term covering sentence-level phenomena such as intonation, pausing, rhythm, speech rate, voice quality and prominence. One of the functional roles of prosody in speech is to signal prominence. Terken & Hermes (2000) define prominence as a property of linguistic units which makes them stand out from their local context. Adequate use of prominence makes some parts of utterances more salient, typically those parts of utterances which are considered to convey new information in a communicative context. The phonetic characteristics of the Dutch intonation system have been particularly well described in ‘t Hart et al. (1990). This model of Dutch intonation describes intonation contours as sequences of perceptually relevant pitch movements which are defined phonetically in terms of a small number of acoustic parameters, i.e. their direction, timing, size and the rate of F0 change: some of these pitch movements are prominence-lending, while others are not. The specific acoustic characteristics of these pitch movements according ‘t Hart et al. (1990) are summarised in Table 1.

Prominence, in this model, is marked by a perceptually relevant pitch movement or a combination of pitch movements on the syllable which is required to stand out. The model incorporates two prominence-lending pitch movements: a rising and a falling one. The rising movement (1) is located very early in the vowel nucleus, it is abrupt and has a short duration of 100 msec. The falling movement (A) is located late in the vowel nucleus, it is abrupt and it has a short duration of 75 msec. Both pitch movements can be combined into 1&A pitch contours which also signal prominence. In other words, syllable prominence can be realised as a prominence-lending rising (1), falling (A) or as a combination of a rising and a falling (1&A) pitch movement within the same syllable nucleus. However, it should be noted that, in the phonological grammar of Dutch intonation, an A fall cannot occur without a preceding rise (’t Hart & Collier, 1979).

Table 1: Pitch movements taxonomy from 't Hart et al. (1990) and standard values for perceptually relevant pitch movements in Dutch (adapted from van Geel, 1983)[*: depends on accent interval; /: undefined].

Labels		1	2	3	4	5	A	B	C	D	E
Direction	Rise	X	X	X	X	X					
	Fall						X	X	X	X	X
Timing	Early	X				X		X			X
	Late			X			X				
	Very late		X						X		
Rate of change	Fast	X	X	X		X	X	X	X		X
	Slow				X					X	
Size	Full	X	X	X		X	X	X	X	X	
	Half				X						X
Duration (in ms)		100	100	150	*	50	75	75	20-50	*	38
Excursion (in ST)		+4	+4	+6	*	+2	-5	-5	/	*	-2.5
Slope (in ST/s)		40	40	40	40	40	66.666	66.666	/	*	65.789

Prosody and hearing impairment

According to the Universal Neonatal Hearing Screening (UNHS), the prevalence of hearing impairment among new-borns is 3 in 1000. The locus and severity of the impairment determine whether the rehabilitation should be carried out via a cochlear implant (henceforth: CI) or a conventional hearing aid (henceforth: HA) (Korver et al., 2017). Neither of these devices provide users with standard hearing but they improve hearing ability sufficiently to make communication possible.

Although it is now well established that the implantation of a CI in congenitally deaf children improves general language acquisition (e.g., Dettman et al., 2016; Levine et al., 2016), CI devices do not transmit all the intensity, temporal and spectral information present in the original acoustic signal (Moore, 2003; Green et al., 2004; Fu & Nogaki, 2005; O'Halpin, 2010). Furthermore, the access to auditory input for CI children is delayed up to the time of activation of the device. Similarly, the range of frequencies amplified by conventional hearing aids is typically limited to the range from 125 Hz to 8 kHz (Metcalfe, 2017). Although the distortions are less significant than those caused by CI, the provided input to HA is more restricted than in normally-hearing children. In other words, both HA and CI provide distorted input, but the extent of the distortions is higher for CI than HA.

Because of the imperfect access to the acoustic signal provided by these devices, hearing-impaired children's speech production differs from that of their NH peers both at the segmental level (cfr. Baudonck et al., 2010) and at the prosodic level (cfr. Lenden & Flipsen, 2007; Vanormelingen et al., 2016).

At the segmental level, the speech of hearing-impaired children is characterised by the distortion of both vowels and consonants. Common problems in the articulation of consonants involve voicing errors (voiceless sounds become voiced and vice versa) and place of articulation substitution errors typically associated with sounds that are articulated posteriorly in the oral cavity where articulatory gestures are less visible. In addition, consonant omission errors have been documented: in some studies word-initial consonant omission appears most frequently (Hudgins & Numbers, 1942), while in others consonant deletion was predominantly word-final (Nober, 1967; Markides, 1970; Smith, 1975). Furthermore, errors pertaining to consonant clusters have been noted: these often have to do with the omission of one of the consonants in clusters or by the insertion of schwas (e.g., Baudonck et al., 2010). The articulation of vowels also seems impaired, be it altogether less frequently than that of consonants. In children with a conventional aid several types of errors have been documented. Vowel substitutions are common and the findings suggest that back vowels are produced more correctly than front vowels and open vowels are more often correct than vowels with a closer degree of stricture (Smith, 1975; Geffner, 1980; Ozbič & Kogovšek, 2008, 2010). Nevertheless, the fronting of back vowels has also been reported (Stein, 1980). Another frequent error involves the neutralization of the peripheral vowels, i.e. the reduction of vowels to a more schwa-like quality (Markides, 1970; Smith, 1975). Furthermore, there have been reports of inappropriate vowel nasalization (Stevens et al., 1976) and the diphthongization of monophthongs (Markides, 1970; Smith, 1975).

At the prosodic level, inappropriate production of prosodic features has been described in hearing-aided speakers, but the findings are quite diverse. As far as the production of prominence is concerned, some studies describe deviant F0 contours (e.g. O'Halpin, 1997), while others mainly find durational differences (e.g. Nickerson, 1975). Sussman & Hernandes (1979) found that in the production of prominent syllables rising pitch contours were replaced by increased intensity. The absence of rising pitch movements on stressed syllables was also found by Gold (1980), Maassen (1986), and Osberger and McGarr (1982). In general, these studies concluded that hearing-aided speakers either produced excessive stress on prominent syllables or used flat and monotone pitch patterns throughout their utterances.

The main objective of the present paper is to study the precise characteristics of the F0 associated with word-level prominence in the speech of children with hearing impairment, as compared to their normally-hearing peers. More specifically, the purpose of this research is twofold. First, it provides a quantitative and empirical comparison of F0 production in children with and without hearing impairment by addressing the average F0 over the entire length of the vowel. Second, it focuses on the dynamics of F0 within the vowel itself.

Method

The data for this study were derived from a corpus of child speech which consists of recordings of Belgian Dutch children imitating Dutch (non-)words that were aurally presented to them (Verhoeven et al., 2016). The children were 6-year-olds differing in hearing status: one group consisted of children with normal hearing, while the other group consisted of children with hearing impairment. The latter contained children with a conventional hearing aid as well as children with a cochlear implant.

Stimuli and data selection

The database contains recordings of 36 monosyllabic (non-)words which consisted of a vowel nucleus with one of the 12 monophthongs of Belgian Standard Dutch, i.e. /i, ʏ, ɪ, ɛ, ɑ, ɔ, u, yː, eː, ø, aː, oː/ (Verhoeven, 2005). Specifically, 16 stimuli were existing Dutch words. The 20 other stimuli were non-words which respect the phonotactic structure of Dutch words. Each vowel occurred in three consonantal contexts: (i) /p_t/, (ii) /l_t/ and (iii) /t_r/. The overall structure of the non-words was compatible with the requirements of the Dutch phonological system.

These (non-)words had first been read aloud by a trained female phonetician and native speaker of Standard Belgian Dutch. The pitch contours of those vocalic productions are illustrated in Figure 1. It is clear that all the example words were realized phonetically with a fall and that there is no indication of list intonation in the examples. The recordings of these stimuli were played to the participating children who were asked to repeat the stimuli one by one. As such, the task is supposed to invoke new information status. All the recordings were made by means of a TASCAM DAT recorder and a head-mounted MicroMic II in a quiet room. The audio files were formatted to WAV-files by means of a TASCAM US 428 Digital Control Surface. The recording sessions with the children yielded a total of 7,985 children's imitations.

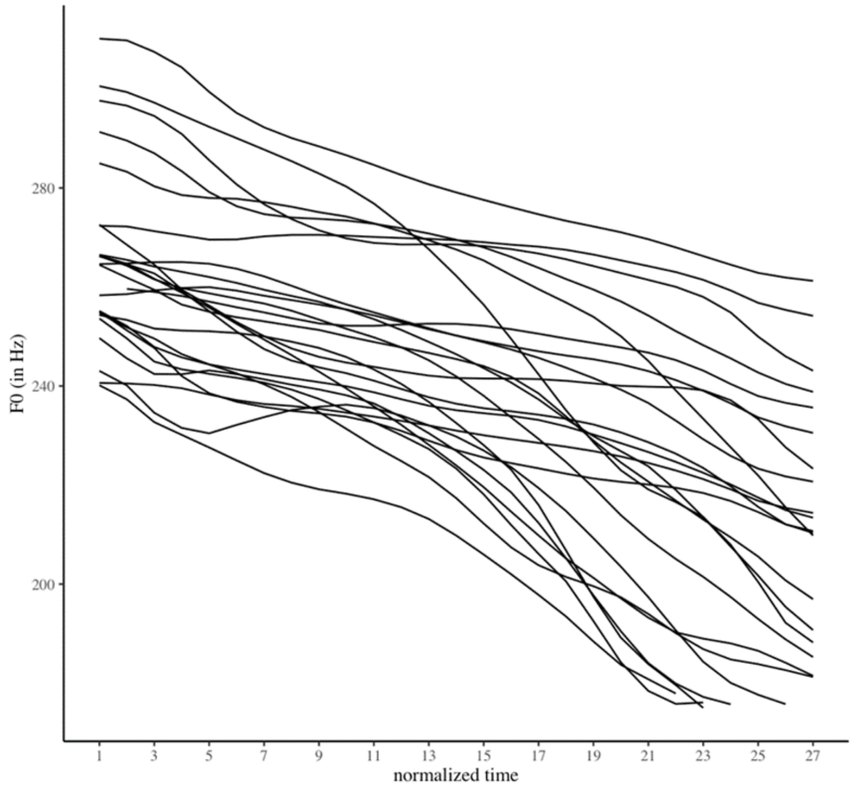


Figure 1: F0 contours in Hz of the stimuli presented aurally to the children.

In the first instance, the children's imitations were perceptually assessed by 6 trained phoneticians who identified the files in which the vowels were correct imitations of the target vowels. The perceptual assessment consisted in a multiple force choice task that was implemented in PRAAT (Boersma & Weenink, 2022). The listeners were aware of the hearing status (CI, HA or NH) and regional background of the children. They worked independently from each other. The judges were instructed to focus specifically on the vowel quality. They had to label the correctness of the vowel imitations as "correct", "incorOzbič & Kogovšekrect", or as "impossible to judge". Inter-rater agreement amounted to 97%. 7,261 vowels from 105 children were labelled as correct, i.e. 90.93% of the files. From these files, only those children were chosen who produced at least 2 repetitions of the corner vowels /i, aː, u/ and at least one repetition of all the vowels in any context. This amounted to a total of 5,557 files produced by 57 children. This selection was further reduced by only selecting the words with vowels following the voiceless stops in order to ensure that the F0 dynamics should be similar for every vowel in the corpus. The pitch tracking procedure did not provide any values for 9 vowels, probably due to the shortness of the window length, and these were also excluded from the analysis. The final selection consisted of 3,695 (non-)words.

Participants

The participants reflected in the corpus were 57 Belgian Dutch children born from native speakers of Belgian Standard Dutch. One group of 47 children had normal hearing on the basis of informal reports from parents and teachers. The mean chronological age of these children was 6 years, with a minimum of 5 and a maximum of 7 years. They all attended their first year of primary school and they had always lived in the region where they were born up to the point of data collection.

Besides children with normal hearing, there was a group of 5 children with a cochlear implant and 5 children with a conventional hearing aid. The hearing characteristics of the two groups of children with hearing impairment are summarised in Table 2 and Table 3:

Table 2: Characteristics of the selected CI children [CI: cochlear implant; HA: hearing aid; HL: hearing loss].

ID	Un-aided HL (in db)	Age at HA (in months)	HL with HA (in db)	Age at CI (in months)	Age at CI fitting (in months)	HL with CI (in db)	Device experience (in months)
CI1	120	1	120	7	8	27	66
CI2	115	2	113	10	12	25	62
CI3	93	5	47	17	18	35	56
CI4	117	4	107	5	6	17	67
CI5	112	2	58	19	21	30	53

Table 3: Characteristics of the selected HA children (HA: hearing aid; HL: hearing loss).

ID	Un-aided HL (in db)	Age at HA (in months)	HL with HA (in db)	Device experience (in months)
HA1	73	26	40	46
HA2	72	10	40	62
HA3	70	9	35	63
HA4	68	32	35	40
HA5	40	9	25	63

Acoustic analysis

The F0 of all the vowels was measured by means of a Python script. The calculation of F0 was done using the Parselmouth API (Jadoul et al., 2018) of PRAAT (Boersma & Weenink, 2022) via its standard auto-correlation algorithm. The maximum number of candidates was set to 15, the silence threshold to 0.03, the voicing threshold to 0.45, the octave cost to 0.01, the octave-jump cost to 0.35, the voiced/unvoiced cost to 0.14. In order to minimise the number of potential octave jumps, the pitch floor and ceiling were set at 175 Hz and 425 Hz respectively after visual inspection of the pitch contours with PRAAT's standard parameters. PRAAT's "Kill octave jumps" function was also applied. The F0 measurements were carried out in Hz, but a Lobanov-normalization procedure (Lobanov, 1971) was used to reduce potential effects of anatomical differences between children. However, the F0 contours in this paper are presented graphically on the Hz scale for ease of inspection.

All the vowels had been segmented by hand. Each vowel was then chunked into 10 frames, each frame representing one tenth of the total duration of the vowel. In addition, there was a frame overlap equal to two-thirds of a frame's duration, which means that each vowel was in effect subdivided into 27 frames. This resulted in 27 F0 measurements which made it possible to draw pitch contours and analyse their dynamics. The F0 value for each frame was taken as the mean of all the F0 measurements in that frame.

Statistical analysis

This study is interested in differences in the mean F0 associated with prominent syllable nuclei. In the first part of the analysis, the mean F0 over the entire vowel between the three groups of children was explored. Given the intrinsically hierarchical nature of these data, multi-level modelling was used as statistical technique.

Models of increasing complexity were built step-by-step by including random and fixed effects one after the other. After the inclusion of each of them, a likelihood ratio test was performed to establish whether the inclusion of that effect significantly improves the fit between the predicted and observed values. A model was preferred over another if the significance level of the model comparison is $p<0.05$. In the case of equally good fit with the data, the most parsimonious model was preferred. This method was based on Bates et al. (2015) and Baayen (2008). More details about the model comparison can be found in Appendix 2. The statistical analysis was carried out in R (R Core Team, 2021) and the R package *lme4* (Bates et al., 2015). *lmerTest* (Kuznetsova et al., 2015) was used to obtain *p*-values.

The final model is presented in Appendix 1. It included Consonant1 (C1), vowel identity, vowel duration as well as the interaction between Consonant1 and vowel and that between vowel and vowel duration. The inclusion of the children's hearing status did not significantly improve the model and was, therefore, excluded. The dependent variable was the mean F0 calculated over all frames of a vowel. As random effects, the basic model included random intercepts per child and per item nested within child.

In the second part of the analysis, this study aimed to investigate the dynamic F0 patterns within the vowels: this was achieved by means of Generalised Additive Mixed Modelling (henceforth: GAMM). This modelling technique allows more flexibility in dealing with non-linear patterns (Wieling, 2018) than polynomial regressions for instance. GAMMs are particularly well suited because they can model variation in height (e.g. vowels produced with different pitch heights) and in shape of trajectory (e.g. vowels produced with different F0 contours). Parametric terms help capture variation in height, and the smooth terms model the variation in shape. The statistical analysis was carried out in R (R Core Team, 2021) by means of the *mgcv* package (Wood, 2015). The *start_event* function of the *itsadug* package (van Rij et al., 2015) was used to order the dataset, and an autoregressive parameter of lag 1 was included which was obtained with a model without the autoregressive parameter, i.e. a ρ value of 0.945. The frame variable was centred so that the parametric terms are representative of the vowel at its centre. The outcome variable of the model are the F0 values normalised through a Lobanov-transformation (Lobanov, 1971). The model comparison consisted of χ^2 tests between a baseline and a more complex model via the *compareML* function of the *itsadug* package (van Rij et al., 2015). The model selection used is similar to the one described in Kirkham et al. (2019). Details can be found in Appendix 4.

The models involved in the model comparison were fitted with the maximum likelihood estimation. The visualization of the residual distribution showed two tails, which are indicative of a *t*-distribution. Consequently, the final model was fitted with a scaled-*t* distribution. This model was fitted with the fast restricted maximum likelihood estimation.

The final GAMM is presented in Appendix 3. It includes a single smooth of time and of vowel duration in order to model the main effect of time and duration on the F0 values. A tensor product interaction was included to capture the nonlinear interaction between the duration of the vocalic segment and time. A single smooth for the effect of child is included. A random reference smooth has been included to model the potential non-linear difference over time according to the time pattern of each individual children. By-children linear random slopes for the different prevocalic consonants (C1) have been included in order to allow the effect of C1 to vary between children. A random difference smooth was added, that is, a curve fit to the non-linear difference over time according to the duration of vowel with a by-effect of (ordered) hearing status and another one with a by-effect of C1. The duration of the vowel is included as a factor in the analysis to capture its effect.

Results

In the present study, a total number of 3,695 vowel productions were analysed acoustically for F0. In terms of average F0, the results of the multi-level modelling procedure in Table 4 show that there are F0 differences between the three speaker groups: the HA children have a somewhat higher F0 (284 Hz; SD=45) than the CI children (277 Hz; SD=38) children and children with normal hearing (270 Hz; SD=34). However, these differences are not significant.

Nevertheless, the statistical analysis revealed a significant fixed effect of both the identity of the vowels and of the identity of the prevocalic consonant (C1). This means that the average F0 varies significantly according to the C1 and the vowel (see Table 4). In terms of the effect of the prevocalic consonant on vowel F0, it is clear that the mean vowel F0 after /p/ is higher than after /t/ (277 Hz vs. 266 Hz respectively). As far as the mean F0 on the different vowels is concerned, the F0 differences are such that F0 is higher on the high vowels as compared to the low vowels: the vowel /aː/ has a mean F0 of 259 Hz, while /i/ and /u/ have higher mean F0 values, i.e. 290 Hz and 284 Hz respectively.[1]

Table 4: Results of the fixed effects of the best-fitting model [C1/p/ & vowel /ɑ/ = reference category]. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1. The results of the interaction effects are not reproduced here.

	Estimate	SE	df	*t*-value	*p*-value	
(Intercept)	0.27773	0.15959	2530.47648	1.74	0.08193	.
vowel duration	-1.42869	1.14618	2833.20236	-1.246	0.21269	
vowel /aː/	-0.87891	0.27992	2624.44372	-3.14	0.00171	**
vowel /ɛ/	-0.22077	0.22242	2606.15682	-0.993	0.321	
vowel /eː/	-0.27817	0.29811	2983.3732	-0.933	0.35084	
vowel /øː/	-0.57127	0.28486	2803.34279	-2.005	0.04501	*
vowel /ɪ/	0.12165	0.21401	2709.28656	0.568	0.5698	
vowel /i/	0.4443	0.20114	2511.09478	2.209	0.02727	*
vowel /ɔ/	0.06885	0.22285	2495.9368	0.309	0.75739	
vowel /u/	0.95945	0.20628	2586.82055	4.651	<0.001	***
vowel /oː/	-0.21534	0.28466	3075.27342	-0.756	0.44942	
vowel /ʏ/	0.26715	0.21378	2500.01094	1.25	0.21156	
vowel /yː/	0.6035	0.20624	2509.0167	2.926	0.00346	**
C1 /t/	-0.39664	0.09913	1436.90168	-4.001	<0.001	***

The second point of interest in this study has to do with the dynamics of the pitch contours in the three speaker groups. The non-normalised F0 contours in the three groups of children are illustrated in Figure 2.

[1] To obtain those values in Hz, the final model was refitted with the unnormalized F0 measurements in Hz.

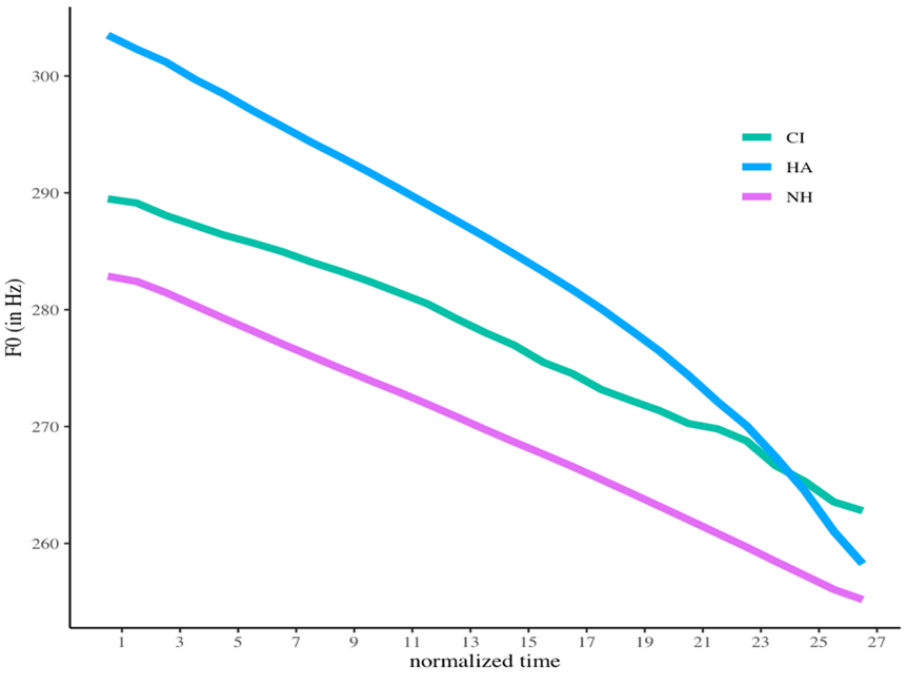


Figure 2: Average F0 contours in Hz for the three groups of children differing in hearing status [CI: cochlear implant; HA: hearing aid; NH: normally hearing].

In terms of the slope of the pitch contours in the speaker groups, there are clear slope differences. For the HA group is the slope of the fall 16.313 ST/sec (*SD*=10.133), while the slope for the CI group amounts to 11.196 ST/sec (*SD*=12.557) and 11.161 ST/sec (*SD*=8.319) for the NH group. Interestingly enough, the F0 slope of the HA group seems to consist of two phases. A first part situated over the first seven-nineth of the vowel has a slope of 13.427 ST/sec (*SD*=9.247) and is followed by an abruptly steeper decrease in F0 which amounts to 22.772 ST/sec (*SD*=20.879). On the contrary, the NH and CI groups display a more constant F0 slope throughout the vowel.

Looking at the variability in the F0 contours between the children, it can be seen in Figure 3 that there is a considerable amount of variability between children. The majority of the NH-children produce a prominence-lending accent realized as a fall similar to the fall produced in the stimuli which were presented to them. It should, however, be signalled that four children produce a rise late in the vowel. This might reasonably be attributed to a question intonation contour through which the child asks whether he or she is correctly performing the task. Visual inspection of the individual contours reveals that two children from the CI group (CI1 and CI5) produce a fall very much similar to the general pattern observed in the other groups and the original stimuli, while the three other children (CI2, CI3 and CI4) produce a rather flat F0 contour.

Note that the F0 contours displayed in Figures 1, 2 and 3 are time normalised. This means that the differences in vowel duration between each item are normalised. However, Figure 4 indicates that there is some variability in vowel durations which needs to be accounted for in the analysis. The CI group produces slightly shorter vowels (*M*= 163 ms; *SD*= 70) than the HA (*M*= 190 ms; *SD*= 85) and NH (*M*=183 ms; *SD*= 70) groups.

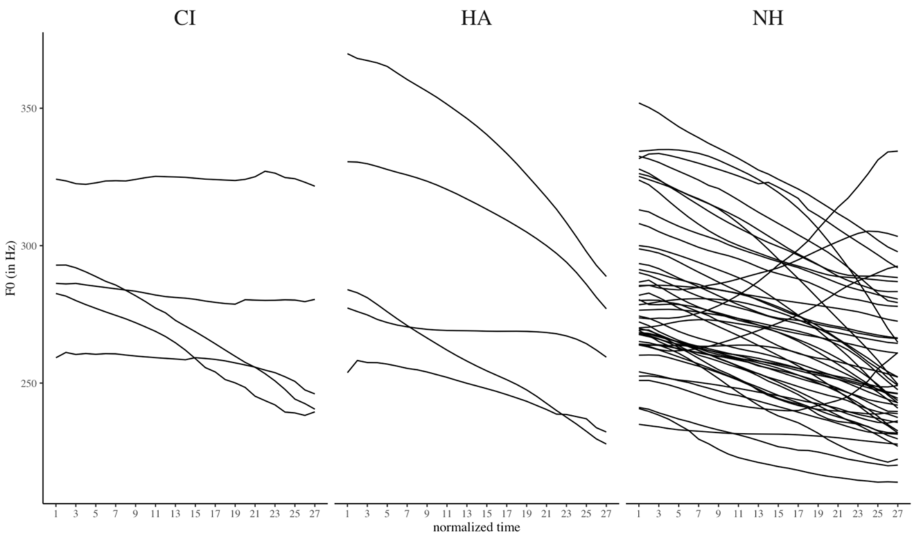


Figure 3: Average F0 contours in Hz per child and hearing status [CI: cochlear implant; HA: hearing aid; NH: normally hearing].

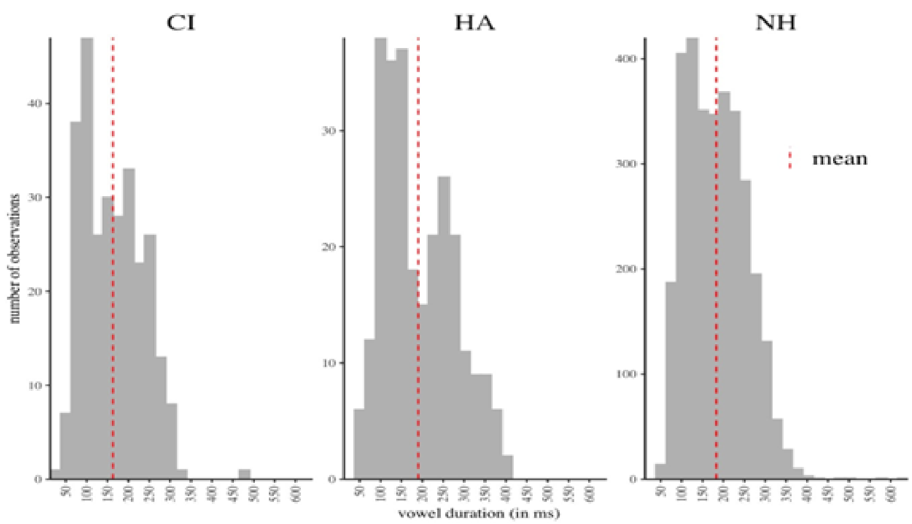


Figure 4: Distribution of vowel duration (in ms) per hearing status [CI: cochlear implant; HA: hearing aid; NH: normally hearing].

As far as the dynamics of the F0 contour are concerned, the results of the GAMM which fits the data best are reported in Table 5. There is a significant parametric effect of C1 (p<0.001).

As expected, the model shows a significant effect of time and of the duration of the vowel ($p<0.001$), their interaction is not significant after refitting the model with a t-distribution. Significant random smooths per subjects and random smooths over C1 are observed. Their effect is significant ($p<0.001$). A random reference smooth per subject over time is shown in the model and its effect is also significant ($p<0.001$). Tensor product interactions between time and duration of the vowel according to C1 are included in the model and their effect are significant for /t/ ($p<0.001$), but not for /p/.

As far as the research question of this paper is concerned, the effect of the interaction between duration of the vowel and time is significantly different between the CI and the HA children ($p<0.01$), but not between the CI and the NH. The non-significance of the latter can reasonably be attributed to the presence of two distinct patterns within the CI group. In fact, if the same model is refitted with the data of only the CI children who do not produce a clear fall, the group difference proves significant ($p<0.001$). It indicates that CI children might differ in the way they modulate F0 to mark prominence, but some of them might produce typical NH-like patterns. That individual pattern is levelled out in the analysis if all CI children are considered as a single group. In other words, the results cannot be generalized to all children, but they show that the modulation of F0 can be affected by CI use.

Generally speaking, it means that the effect of the non-linear interaction between time and duration of the vowel is different according to the hearing status of the child. It means that the effect of time modulated by the duration of the vowel has a different non-linear effect on the F0 values according to the hearing status of the child. GAMMs coefficients are per se difficult to interpret. The visualization of the values predicted by the model are therefore presented in Figure 5. We observe in Figure 5 that the CI group has a relatively flatter F0 contour than the two other groups. On the contrary, the HA and NH groups produce a clearly marked fall. During the first part of the vowel, the HA and NH contours are, after normalization, quite close to each other. However, towards the end of the vowel, the F0 contour becomes more abrupt in the HA group than in the NH group.

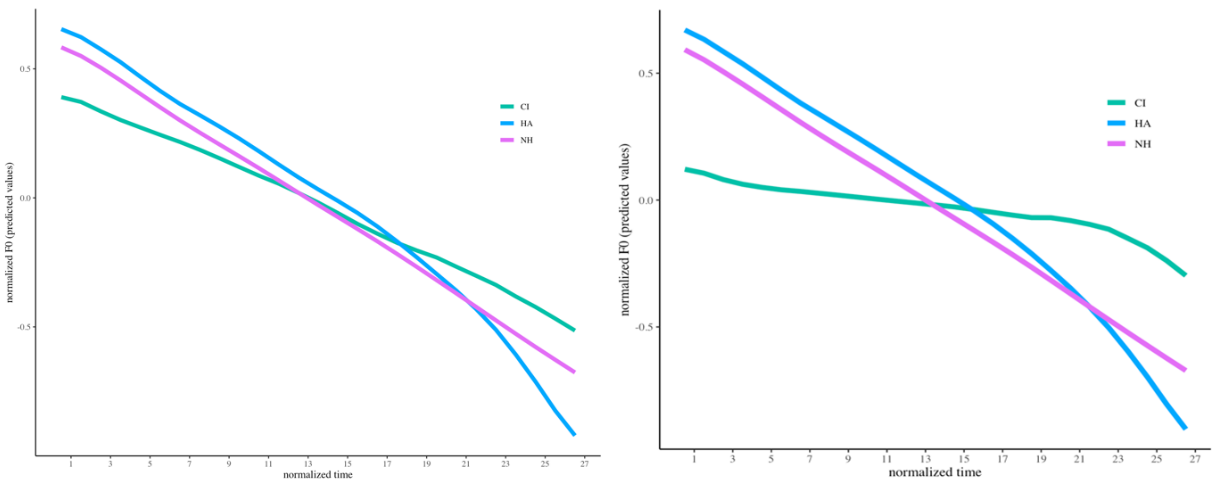
Discussion

The key research question of this study was whether children differing in hearing status differ in the (dynamics of) F0 associated with prominence after voiceless plosives within the nucleus of monosyllabic words. This study was carried out on the speech of 6-year-old hearing-impaired children who had been fitted with either a cochlear implant (N=5) or a conventional hearing aid (N=5) in comparison with a control group which consisted of 47 children with normal hearing. The pitch values were extracted via PRAAT inside time-normalised overlapping frames. The mean F0 was analysed by means of multilevel modelling, while the overall contours for the three groups were modelled via GAMMs. From the results, it appears that there are differences in the mean F0 and the dynamics of the F0 contours between the three groups. However, not all the observed differences are significant.

As far as the mean F0 is concerned, visual inspection of the data reveals that the mean F0 associated with prominence in children with hearing impairment is somewhat higher

Table 5: Summary of the GAMM created to compare the F0 contours of the CI, HA and NH groups. Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1. [C1/p/ = reference category].

A. Parametric coefficients	Estimate	SE	*t*-value	*p*-value
Intercept	0.0843	0.0396	2.129	0.0333*
C1/t/	-0.30483	0.05428	-5.615	<0.001***
B. Smooth terms	edf	Ref.df	F-value	*p*-value
s(total_duration)	4.008	5.006	60.425	<0.001***
s(time_centred)	3.711	4.204	114.866	<0.001***
ti(Time_centred, total_duration)	2.05	2.905	0.242	0.86855
s(name, bs="re")	0.001	56	0	0.16922
s(name, C1, bs="re"):C1/p/	117.9	512	0.993	<0.001***
s(name, C1, bs="re"):C1/t/	289.2	512	1.886	<0.001***
s(Time_centred, name, bs="fs",m=1)	326.3	512	2.616	<0.001***
ti(Time_centred, total_duration):HA	2.599	3.279	3.656	0.00961**
ti(Time_centred, total_duration):NH	7.12	8.897	1.4	0.18368
ti(Time_centred, total_duration):C1/p/	2.286	2.672	1.86	0.12082
ti(Time_centred, total_duration):C1/t/	5.177	6.242	5.626	<0.001***
Rank: 1695/1696				
R-sq. (adj) = 0.387				
Deviance explained = 36.1%				
fREML= 48695				
Scale est. = 1				
n = 99028				



(a) *(b)*

Figure 5: Time course of F0 (in Hz) per hearing status (predicted values) [CI: cochlear implant; HA: hearing aid; NH: normally hearing]. (a) Full corpus; (b) CI1 and CI5 excluded.

than in children with normal hearing. This is consistent with previous research on hearing impairment which indicates that higher F0 values are characteristic of the speech of both hearing-impaired adults (Mora et al., 2010) and children (Angelocci et al., 1964; Higgins et al., 2003; Dehqan and Scherer, 2011). It should however be pointed out that the differences found in the present study are statistically not significant, so these results do not provide strong confirmation of the findings of previous research. It is not clear how to account for the non-significance of the findings in this paper, but this may be due to a combination of very rigorous statistical testing (as compared to other research), the small number of participants in some of the subgroups and the high variability in these groups.

In addition, the statistical analysis revealed a significant effect of both the vowel and the prevocalic consonant on the mean F0. In terms of vowel identity as a factor it was clear that the high vowels have a significantly higher F0 than the low vowels. The difference amounts on average to 23 Hz. This finding is consistent with many studies which have documented such F0 differences between high and low vowels in a large number of languages and in different groups of speakers (among others: Ladefoged 1964; Traunmüller, 1981; Whalen & Levitt, 1994; Van Hoof & Verhoeven, 2011). However, this is the first time that a robust effect of instrinsic vowel F0 has been documented in such a large group of child speakers, both with and without hearing impairment. This may open interesting perspectives for further research, particularly into the question of whether intrinsic F0 can be actively controlled by speakers for communicative purposes such as to perceptually enhance vowel contrasts. However, it should be noted that other sources of variability that the present carefully controlled task is unable to account for might be present in more ecological contexts.

Furthermore, the analysis also showed a significant effect of the prevocalic consonant on the F0 of the vowel nucleus. /p/ as a C1 triggers higher F0 values in the vowel than /t/ i.e. 277 Hz vs. 266 Hz. Although the differential effect of plosive voicing on the F0 of adjacent vowels is well-known and has been documented before (House & Fairbanks, 1953; Lehiste & Peterson, 1961; Xu & Xu, 2021), the finding of a significant difference in place of articulation of the voiceless plosives is new: the data in House & Fairbanks (1953) and Lehiste & Peterson (1961) show F0 differences of around 1 Hz between the two places of articulation.

As far as the pitch configuration is concerned, the modelling procedure indicates that the general pitch configuration associated with prominence in this study is a prominence-lending F0 fall which is the phonetic manifestation of an underlying rise-fall nuclear accent contour. Visual inspection of the F0 contours in the different groups of children in Figure 3 reveals clear differences between the three groups which are found to be significant at a statistical level.

From Figure 5, it is clear that the slope of the falls associated with prominence in CI children is significantly shallower than in the other groups. This indicates that there is less modulation of F0 in the CI group than in the HA and NH groups. The smaller modulation of F0 by CI users is consistent with previous research on the speech production of CI users (e.g. Hide, 2014, De Clerck et al., 2018) which describes the speech of CI speakers as more monotonous and this might indicate that the lack of detail in the auditory input provided to children fitted with CI is mirrored by a lack of F0 modulation produced by them (cfr. Moore, 2003; Green et al., 2004; Fu & Nogaki, 2005; O'Halpin, 2010). It should, however, be emphasised that there are differences between individual CI children in this respect. Although most of the CI children in this study produce a relatively flat F0 contour, there are some CI children who produce a fall which is very similar to their NH peers. The reasons for these individual differences are not immediately clear since there are no specific auditory characteristics which can explain these differences in F0 modulation. It can nevertheless be seen from this study that the acquisition of F0 contours similar to that of children with normal hearing is not impossible to achieve by children fitted with a CI. The CI's greater variability in speech production in general is also consistent with the present findings.

When it comes to the HA group, a more similar F0 contour to the NH group is found: in all the participants with HA the F0 drops without exception towards the end of the vowel nucleus at a rate which is very similar to the children with normal hearing. However, towards the end of the vowel, the F0 tends to drop more steeply. The cause of this sudden drop in F0 is not clear but might be attributed to a lower degree of control over phonation towards the end of the vowel. The finely-tuned control of F0 needed for the realisation of prominence might be achieved by HA children throughout the first part of the vowel but might be difficult to maintain until the end. Note also in Figure 4 that the HA group produces, on average, slightly longer vowels than the two other groups.

This difficulty to maintain fine control over phonation throughout the vowel might lead to a change in voice quality, such as creaky voice, towards the end of the vowel and this might be the cause of the observed lower F0 values towards the end of the vowels produced by the HA children. This would be in accordance with what Van Lierde et al. (2005), Valero Garcia et al. (2010) and Baudonck et al. (2011) suggest about the voice quality of hearing-impaired speakers, especially for HA users. In fact, Valero Garcia et al. (2010) suggest that the use of analogue hearing aids, contrary to digital hearing aids, is associated with worse voice quality than that of cochlear implanted children. Similarly, Van Lierde et al. (2005) find that the two groups differ with respect to the norm in terms of jitter, whereby the CI group exhibits lower jitter and the HA group higher jitter. Based on a GRBAS perceptual rating of voice quality Baudonck et al. (2011) found that HA and CI (both unilaterally and bilaterally implanted) exhibit a strained and unstable phonation. They also point out that the HA obtained the worst scores on the GRBAS scale.

The above-mentioned differences between groups are observable as an effect of the hearing status of the children on the non-linear interaction between time and duration. It shows that the effect of time on the normalised pitch values is different between groups, but the effect of time is modulated by the duration of the vowels. As such, the differences in vowel duration between groups are accounted for in the analysis. It means that the different groups do not differ in height of F0 but in the non-linear F0 pattern, i.e. the shape of the contour.

A difference in F0 contour can also be observed between the two hearing-impaired groups. It can reasonably be attributed to device use. Hearing aids produce a more faithful input of F0, which is not the case for CI. The potential sound distortion caused by the amplification of given frequencies by a conventional hearing aid (Metcalfe, 2017) might distort the original input to a certain extent, but the resolution would still be higher than that provided by the limited number of electrodes of a CI (Moore, 2003; Green et al., 2004; Fu & Nogaki, 2005; O'Halpin, 2010). This might consequently make the acquisition of an NH-like pattern easier for HA than for CI children. The data nevertheless show that, despite the low resolution of the signal provided by CI, some children manage to acquire an NH-like pattern by the age of 6. Over time, CI children might learn to infer information about F0 from the higher harmonics because the degraded input might make it difficult to rely mainly on F0 itself. That might be an explanation for the delayed, but not impossible, acquisition of NH-like contours by CI children.

In a nutshell, the three hearing groups produce the F0 modulations needed for the realisation of prominence differently. The children fitted with a cochlear implant produce a less steep fall. This is in agreement with research describing CI users' speech as monotonous as a result of the lack of detail in acoustic input provided by the implant. On the contrary, the HA group produce more naturally modulated F0 patterns which are close to the patterns of their NH peers apart from a sharp decrease in F0 towards the end of the vowels.

Limitations of this study

This research sheds light on the modulation of F0 by three groups of children differing in hearing status, but it may be useful to point out some limitations arising from the methodological choices that were made.

Firstly, the nature of a word and non-word repetition task is anything but ecological. In fact, the design of the experiment aimed to elicit speech in laboratory conditions in order to control for experimental variables. Secondly, to enhance within-children data homogeneity and reduce computational cost, a further selection was applied on the original corpus (Verhoeven et al., 2016). As a result, the data selection significantly reduces the sample size and caution must be taken when interpreting the results of this much more restricted dataset. The structure of the data is also somewhat unbalanced. Thirdly, the parametric settings of the F0 tracking algorithm were set in order to reasonably include the pitch range of all the participating children. It does not entail automatically that those settings are ideal for each child and that the F0 tracking does not make mistakes. Furthermore, this paper focuses on the description of detailed modulations in the speech of normally hearing and hearing-impaired children but does not engage with the larger F0 modulations to be expected in longer and more complex utterances. The extent to which these detailed F0 modulations participate in the modulations of a larger prosodic structure is unclear.

Another limitation is that the speech materials are made up of repetitions of monosyllables and as such word-stress coincides with sentence stress in utterance-final position so that effects of end-of-utterance on F0 can be expected such as the sudden drop in F0 in HA children.

It should also be noted that the GAMM analysis does not engage with vowel effects for computational cost reasons. We nonetheless believe that a potential effect of vowel category should be reflected in a height effect as demonstrated by the multi-level modelling analysis, but it is not expected to have the shape of a trajectory effect. Moreover, the functional relevance of the differences in F0 dynamics described here above are still unclear. In other words, it is not clear whether a less steep nuclear fall hampers the intelligibility of CI users. Given the individual differences within a single group, it would be of interest to collect more data. Longitudinal data might also help understand whether the production of an NH-like pattern by CI users depends on individual learning trajectories.

Conclusions

This paper investigated the dynamics of F0 contours in prominent vowel nuclei after voiceless plosives in two groups of hearing-impaired children and one group of normally hearing children. The modelling procedure indicated clear differences in F0 both in terms of its overall mean in the vowel nucleus as well as in the shape of the contours. The overall mean F0 in the vowel nucleus was somewhat higher in children with hearing impairment than in children with normal hearing. Although this is consistent with previous findings on hearing impaired speech, the difference in this study was not significant. Additionally, this study revealed that most of the NH children realise the prominent syllable in this specific task through a pointed-hat pattern realised phonetically as a fall. On the one hand, HA children produce F0 contours similar to those produced by NH children, even if HA children tend to produce a steeper fall towards the end of the vowel. This might be indicative of a difficulty to maintain fine-tuned control over the total vowel duration. On the other hand, CI children often produce a rather flat F0 contour. This might be a consequence of the lack of spectral details provided by the CI device. However, some CI children show F0 contours similar to those of their NH peers.

Acknowledgements

This work was supported by the Research Foundation in Flanders (FWO) [grant G004321N].

References

Angelocci, A. A., Kopp, G. A., & Holbrook, A. (1964). The Vowel Formants of Deaf and Normal-Hearing Eleven- to Fourteen-Year-Old Boys. *Journal of Speech and Hearing Disorders, 29*(2), 156-170. https://doi.org/10.1044/jshd.2902.156

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge ; New York: Cambridge University Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. http://dx.doi.org/10.18637/jss.v067.i01

Baudonck, N., Dhooge, I., D'haeseleer, E., & Van Lierde, K. (2010). A comparison of the consonant production between Dutch children using cochlear implants and children using hearing aids. *International Journal of Pediatric Otorhinolaryngology, 74*(4), 416-421. http://dx.doi.org/10.1016/j.ijporl.2010.01.017

Baudonck, N., Lierde, K. V., D'haeseleer, E., & Dhooge, I. (2011). A comparison of the perceptual evaluation of speech production between bilaterally implanted children, unilaterally implanted children, children using hearing aids, and normal-hearing children. *International Journal of Audiology, 50*(12), 912-919. https://doi.org/10.3109/14992027.2011.605803

Boersma, P. & Weenink, D. (2022). *Praat: doing phonetics by computer* (version 6.2.06).

De Clerck, I., Pettinato, M., Gillis, S., Verhoeven, J., & Gillis, S. (2018). Prosodic modulation in the babble of cochlear implanted and normally hearing infants: A perceptual study using a visual analogue scale. *First Language, 38*(5), 481-502. http://dx.doi.org/10.1177/0142723718773957

Dehqan, A., & Scherer, R. C. (2011). Objective Voice Analysis of Boys with Profound Hearing Loss. *Journal of Voice, 25*(2), e61-e65. https://doi.org/10.1016/j.jvoice.2010.08.006

Dettman, S. J., Dowell, R. C., Choo, D., Arnott, W., Abrahams, Y., Davis, A., Dornan, D., Leigh, J., Constantinescu, G., & Cowan, R. (2016). Long-term communication outcomes for children receiving cochlear implants younger than 12 months: A multicenter study. *Otology & Neurotology, 37*(2), e82-e95. http://dx.doi.org/10.1097/MAO.0000000000000915

Fu, Q.-J. & Nogaki, G. (2005). Noise Susceptibility of Cochlear Implant Users: The Role of Spectral Resolution and Smearing. *Journal of the Association for Research in Otolaryngology, 6*(1), 19-27. http://dx.doi.org/10.1007/s10162-004-5024-3

Geffner, D. (1980). Feature characteristics of spontaneous speech production in young deaf children. *Journal of Communication Disorder, 13*(6), 443-454.

Gold, T. (1980). Speech production in hearing-impaired children. *Journal of Communication Disorders, 13*(6), 397-418. https://doi.org/10.1016/0021-9924(80)90042-8

Green, T., Faulkner, A., & Rosen, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *The Journal of the Acoustical Society of America, 116*(4), 2298-2310. http://dx.doi.org/10.1121/1.1785611

't Hart, J. & Collier, R.P.G. (1979). On the interaction of accentuation and intonation in Dutch. In Fischer-Jørgensen, E., Rischel, J., & Thorsen, N. (Eds.), *Proceedings of the Ninth International Congress of Phonetic Sciences,* Held in Copenhagen, 6-11 August 1979 (pp. 395–402). University of Copenhagen.

´t Hart, J., Collier, R., & Cohen, A. (1990). *A Perceptual Study of Intonation: An Experimental Phonetic Approach to Speech Melody.* Cambridge ; New York: Cambridge University Press.

Hide, O. (2014). *Acoustic Features of Speech by Young Cochlear Implant Users: A Comparison with Normal-Hearing and Hearing-Aided Age Mates* [PhD thesis, University of Antwerp].

Higgins, M. B., McCleary, E. A., Carney, A. E., & Schulte, L. (2003). Longitudinal changes in children's speech and voice physiology after cochlear implantation. *Ear and Hearing, 24*(1), 48-70. https://doi.org/10.1097/01.AUD.0000051846.71105.AF

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America, 25*, 105-113.

Hudgins, C. V., & Numbers, F. C. (1942). An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs, 25*, 289-392.

Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics, 71*, 1-15. https://doi.org/10.1016/j.wocn.2018.07.001

Kirkham, S., Nance, C., Littlewood, B., Lightfoot, K., & Groarke, E. (2019). Dialect variation in formant dynamics: The acoustics of lateral and vowel sequences in Manchester and Liverpool English. *The Journal of the Acoustical Society of America, 145*(2), 784-794. https://doi.org/10.1121/1.5089886

Korver, A. M. H., Smith, R. J. H., Van Camp, G., Schleiss, M. R., Bitner-Glindzicz, M. A. K., Lustig, L. R., Usami, S.-I., & Boudewyns, A. N. (2017). Congenital hearing loss. *Nature Reviews. Disease Primers, 3*:16094. http://dx.doi.org/10.1038/nrdp.2016.94

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. (2015). *Package 'lmertest'*. R package version 2, 734.

Ladefoged, P. (1964). *A phonetic study of West African languages.* Cambridge: Cambridge University Press.

Lehiste, I., & Peterson, G. E. (1961). Some Basic Considerations in the Analysis of Intonation. *The Journal of the Acoustical Society of America, 33*, 419-425. https://doi.org/10.1121/1.1908681

Lenden, J. M. & Flipsen, P. (2007). Prosody and voice characteristics of children with cochlear implants. *Journal of Communication Disorders, 40*(1), 66-81. http://dx.doi.org/10.1016/j.jcomdis.2006.04.004

Levine, D., Strother-Garcia, K., Golinkoff, R. M., & Hirsh-Pasek, K. (2016). Language development in the first year of life: What deaf children might be missing before cochlear implantation. *Otology & Neurotology, 37*(2), e56-e62.

Lobanov, B. M. (1971). Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America, 49*(2B), 606-608. http://dx.doi.org/10/cqm563

Maassen, B. (1986). Marking word boundaries to improve the intelligibility of the speech of the deaf. *Journal of Speech, Language, and Hearing Research, 29*(2), 227-230.

Markides, A. (1970). The speech of deaf and partially-hearing children with special reference to factors affecting intelligibility. *British Journal of Disorders of Communication, 5*(2), 126-140.

Metcalfe, T. (2017). *Perception of Speech, Music and Emotion by Hearing-Impaired Listeners* [PhD thesis, University of Sheffield].

Moore, B. C. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology & neurotology, 24*(2), 243-254. http://dx.doi.org/10.1097/00129492-200303000-00019

Mora, R., Jankowska, B., Guastini, L., Santomauro, V., Dellepiane, M., Crippa, B., & Salami, A. (2010). Computerized Voice Therapy in Hypofunctional Dysphonia. *Journal of Otolaryngology - Head & Neck Surgery, 39*, 615-621. https://doi.org/10.2310/7070.2010.090209

Nickerson, R. S. (1975). Characteristics of the Speech of Deaf Persons. *Volta Review, 77*(6), 342-62.

Nober, E. H. (1967). Articulation of the Deaf. *Exceptional Children, 33*(9), 611–621. https://doi.org/10.1177/001440296703300902

O'Halpin, R. (1997). Contrastive Stress in the Speech of Profoundly Deaf Children: A Preliminary Investigation. *Volta Review, 99*(2), 89-105.

O'Halpin, R. (2010). *The Perception and Production of Stress and Intonation by Children with Cochlear Implants* [PhD thesis, University of London].

Osberger, M. J., & Mcgarr, N. S. (1982). Speech Production Characteristics of the Hearing Impaired. In N. J. Lass (Ed.), *Speech and Language* (Vol. 8, pp. 221–283). https://doi.org/10.1016/B978-0-12-608608-9.50013-9.

Ozbič, M., & Kogovšek, D. (2008). An acoustic comparison of formant frequencies in individuals with normal hearing, profound and severe hearing impairment. *Investigationes Linguisticae, 16*, 150–162. https://doi.org/10.14746/il.2008.16.13

Ozbič, M., & Kogovšek, D. (2010). Vowel formant values in hearing and hearing-impaired children: A discriminant analysis. *Deafness & Education International, 12*(2), 99-128. https://doi.org/10.1179/146431510X12626982043804

R Core Team, (2021). *R: A Language and Environment for Statistical Computing.* Vienna, Austria. Retrieved from https://www.R-project.org/

Smith, C. R. (1975). Residual hearing and speech production in deaf children. *Journal of speech and hearing research, 18*(4), 795-811.

Stein, D. M. (1980). *A study of articulatory characteristics of deaf talkers* [PhD thesis, The University of Iowa].

Stevens, K. N., Nickerson, R. S., Boothroyd, A., & Rollins, A. M. (1976). Assessment of Nasalization in the Speech of Deaf Children. *Journal of Speech and Hearing Research, 19*(2), 393–416. https://doi.org/10.1044/jshr.1902.393

Sussman, H., & Hernandez, M. (1979). A spectrographic analysis of the suprasegmental aspects of the speech of hearing-impaired adolescents. *Audiology and Hearing Education, 5*, 12-16.

Terken, J., & Hermes, D. (2000). The Perception of Prosodic Prominence. In M. Horne (Ed.) *Prosody: Theory and Experiment, studies presented to Gösta Bruce* (pp. 89–127). Dordrecht ; Boston ; London: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-015-9413-4_5

Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *The Journal of the Acoustical Society of America, 69*(5), 1465-1475. https://doi.org/10/c2vbc9

Valero Garcia, J., Rovira, J. M. V., & Sanvicens, L. G. (2010). The influence of the auditory prosthesis type on deaf children's voice quality. *International Journal of Pediatric Otorhinolaryngology, 74*(8), 843-848. http://dx.doi.org/10.1016/j.ijporl.2010.04.011

van Geel, R. C. (1983). *Pitch inflection in electrolaryngeal speech.* [PhD thesis, University Utrecht].

Van Hoof, S., & Verhoeven, J. (2011). Intrinsic vowel F0, the size of vowel inventories and second language acquisition. *Journal of Phonetics, 39*(2), 168-177. https://doi.org/10/bnth37

Van Lierde, K. M., Vinck, B. M., Baudonck, N., De Vel, E., & Dhooge, I. (2005). Comparison of the overall intelligibility, articulation, resonance, and voice characteristics between children using cochlear implants and those using bilateral hearing aids: A pilot study. *International Journal of Audiology, 44*(8), 452-465. http://dx.doi.org/10.1080/14992020500189146

van Rij, J., Wieling, M., Baayen, R. H., & van Rijn, D. (2015). *Itsadug: Interpreting time series and autocorrelated data using GAMMs.*

Vanormelingen, L., De Maeyer, S., & Gillis, S. (2016). A comparison of maternal and child language in normally-hearing and hearing-impaired children with cochlear implants. *Language, Interaction and Acquisition, 7*(2), 145-179. http://dx.doi.org/10.1075/lia.7.2.01van

Verhoeven, J. (2005). Belgian Standard Dutch. *Journal of the International Phonetic Association, 35*(2), 243-247. http://dx.doi.org/10.1017/S0025100305002173

Verhoeven, J., Hide, O., De Maeyer, S., Gillis, S., & Gillis, S. (2016). Hearing impairment and vowel production. A comparison between normally hearing, hearing-aided and cochlear implanted Dutch children. *Journal of Communication Disorders, 59*, 24-39. http://dx.doi.org/10/f787b7

Whalen, D., & Levitt, A. (1994). The universality of intrinsic F0 of vowels. *Haskins Laboratories Status Report on Speech Research*, SR-117/118: 1-14.

Wieling, M. (2018). 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, 70*, 86-116. http://dx.doi.org/10.1016/j.wocn.2018.03.002

Wood, S. (2015). *Package 'mgcv'.* R package version, 1(29), 729.

Xu, Y., & Xu, A. (2021). Consonantal F0 perturbation in American English involves multiple mechanisms. *The Journal of the Acoustical Society of America, 149*(4), 2877-2895. https://doi.org/10.1121/10.0004239

Appendix

1. Listing - Final multi-level model

```
# Fixed effect of the duration of the vowel on the average
# normalized F0
final_model <- lmer(mean_z_pitch   (total_duration

# Fixed effect of prevocalic consonant (/p/ vs /t/)
+ C1)

# Fixed effect of the vowel category and interaction with C1 and
# duration of the vowel
* vowel

# Random effects of child and item nested within child
+(1|name/C1_V))
```

2. Model comparison - Multi-level modelling analysis

Table 6: Model comparison [*significantly improves data fit or more parsimonious model].

	Added term	Baseline model	Compared model	Npar	AIC	BIC	Loglik	Deviance	χ^2	df	*p*-value	Comparison output	Retained model *
Null Model with random intercept per child(M1)	Duration fixed effect (M2)	M1	M2	4	8190.4	8215.3	-4091.2	8182.4	753.72	1	<0.001	Random intercept per vowel needed	M2
	Duration fixed effect (M3)	M2	M3	5	8052.4	8083.5	-4021.2	8042.4	140	1	<0.001	Duration fixed effect needed	M3
	C1 fixed effect (M4)	M2	M4	6	7956.0	7993.3	-3972.0	7944.0	98.454	1	<0.001	C1 fixed effect needed	M4
	V fixed effect (M5)	M4	M5	17	7449.6	7555.2	-3707.8	7415.6	528.38	11	<0.001	V fixed effect needed	M5
	C1 * V interaction (M6)	M5	M6	28	7414.7	7588.7	-3679.3	7358.7	56.925	11	<0.001	C1*V interaction needed	M6
	C1 * Duration interaction (M7)	M6	M7	29	7410.3	7590.5	-3676.1	7352.3	6.394	1	0.01145	C1 * Duration interaction needed	M7
	Duration * vowel interaction (M8)	M7	M8	51	7429.0	7745.9	-3663.5	7327.0	25.296	22	0.2831	Duration * vowel interaction not needed	M7
	Hearing status fixed effect (M10)	M6	M10	41	7428.7	7683.5	-3673.4	7346.7	0.2745	2	0.8718	Equivalent data fit	M6

3. Listing - Final generalised additive mixed model

```
# Parametric term of C1 (i.e. fixed effect) on normalized pitch:
final_model <- bam(z_pitch   C1

# Single smooth (i.e. random intercept) for the effect of
# time_centred and total_duration
+s(total_duration)
+s(time_centred)

# Tensor product interaction of time_centred and total_duration
#(i.e. interaction between variables)
+ti(Time_centred, total_duration)

# Single smooth (i.e. random intercept) for the effect of child
+s(name, bs="re")

# Difference smooth (i.e. random slope) for the effect of child
# per C1
+s(name, C1, bs="re")

# Random smooth for the non-linear effect of time
# according to general pattern of each speaker
+s(Time_centred, name, bs="fs",m=1)

# Tensor product interaction for the effect of child per C1
# and hearing status
+ti(Time_centred, total_duration, by= C1)
+ti(Time_centred, total_duration, by= Hearing status))
```

Table 7: Model comparison [*significantly improves data fit or more parsimonious model].

	Added term		Removed term	Baseline model	Compared model	Score	Edf	Difference	Df	*p*.value	Comparison output	Retained model*
Null Model (M1)	Model single and random smooths per child (M1B)			M1	M1B	-66425.34	11	195341.267	3	<0.001	Single and random smooths per child needed	M1B
	With C1 predictor (M2A)			M1B	M2A	-70447.85	23	4022.510	12	<0.001	C1 predictor needed	M2A
		With smooth term only M2B)		M2A	M2B	-70447.85	23	14.500	1	<0.001	Parametric term needed	M2A
		With parametric term only (M2C)		M2A	M2C	-70447.85	23	3652.534	13	<0.001	Smooth terms needed	M2A
			Without tensor product interaction (M2A1)	M2A	M2A1	-70447.85	23	56.404	6	<0.001	Tensor product interaction needed	M2A
			Without single smooth (M2A2)	M2A	M2A2	-70449.01	22	-1.162	1	Only small difference in ML	Equivalent data fit	M2A2
			Without random smooth (M2A3)	M2A	M2A3	-70447.85	23	1923.328	4	<0.001	Random smooth needed	M2A
	With hearing status predictor (M3A)			M2A2	M3A	-70596.66	33	147.650	11	<0.001	Hearing status predictor needed	M3A
		With smooth term only (M3B)		M3A	M3B	-70572.56	35	0.458	2	0.633	Equivalent data fit	M3B
		With parametric term only (M3C)		M3A	M3C	-70572.56	35	124.263	11	<0.001	Smooth terms needed	M3A
				M3B	M3C	-70572.1	33	123.806	9	<0.001		M3B
			Without tensor product interaction (M3B1)	M3B	M3B1	-70572.10	33	122.932	6	<0.001	Tensor product interaction needed	M3B
			Without single smooth (M3B2)	M3B	M3B2	-70572.95	32	-0.850	1	Only small difference in ML...	Equivalent data fit	M3B2
			Without random smooth (M3B3)	M3B	M3B3	-70572.11	29	-0.011	4	Only small difference in ML...	Equivalent data fit	M3B3
			Without random and single smooths (M3B4)	M3B4	M3B2	-70572.95	32	0.973	4	0.746	Equivalent data fit	M3B4
				M3B4	M3B3	-70572.11	29	0.134	1	0.604		