A comparative study of kinematic and acoustic age-related variability in speech

Schötz, Susanne; Frid, Johan; Löfqvist, Anders

Published in:
Proceedings from FONETIK 2012

2012

A comparative study of kinematic and acoustic age-related variability in speech

Susanne Schötz[1], Johan Frid[1], Anders Löfqvist[2]
[1] Humanities Lab, Centre for Languages and Literature, Lund University
[2] Department of Logopedics, Phoniatrics and Audiology, Lund University

Abstract

In this study, we compared age-related lip movement variability with acoustic variability across repetitions using Functional Data Analysis (FDA). Lip movement and acoustic data of 15-20 repetitions of a short Swedish phrase from 50 Swedish speakers were collected. Results showed weak to moderate negative correlations for age and variability in lip movements, phrase duration and sound pressure level, which confirm previous studies of decreasing variability in speech with age. Variability in F_0 *did not vary with age, and this may be because we tend to vary our intonation in speech both unconciously when we are learning to speak, and conciously as adult speakers. This will be explored further in a follow-up study.*

Introduction

A number of studies have examined variability in speech production using acoustic measurements and articulatory movements at both the segment and utterance levels (see Smith, 2010, for a review). All of these studies have shown that variability decreases with age up until at least adolescence. A different view is presented by Stathopoulos (1995), who reports similar variability for children and adults for a number of acoustic and aerodynamic measures using syllable trains. It is also well known that the fundamental frequency decreases during childhood and that the rate of change tends to be similar for boys and girls until puberty (e.g., Lee et al. 1999). Stathopoulos and Sapienza (1997) show that children tend to use higher subglottal pressure and have higher glottal resistance to airflow than adults, but do not have a higher glottal airflow. They also report that children use a larger percentage of their vital capacity during speech than adults. Childrens voices also seem to be less efficient than those of adults (Tang and Stathopoulos, 1995). Swedish children usually master the prosody of their native language at the age of five years (cf., Samuelsson and Löfqvist, 2006).

In previous studies, some (Goffman and Smith, 1999; Sadagopan and Smith, 2008) have used the spatiotemporal index (STI, Smith et al., 1995), which provides a single metric of variability incorporating both amplitude and phase. Others (Koenig et al., 2008; Lucero and Löfqvist, 2005) have used functional data analysis (FDA, Ramsay et al., 1996), where amplitude and phase variability are calculated separately. In an earlier study (Frid et al., 2011a, b), we found that the amplitude index of the FDA showed a larger age-related lip movement variability than the phase index of the FDA or the STI.

The purpose of the present study was to extend our earlier findings of decreased repetition variability of lip movements with age (Frid et al. 2011 a, b) to utterance duration, sound pressure level and fundamental frequency.

Experiment and analysis

Lip movements were recorded along with a microphone signal using the Carstens Articulograph AG500. To obtain as large lip movements as possible, we selected the Swedish phrase ”Mamma pappa barn” (Mummy daddy children) as speech material. This phrase is also short and can be spoken on a single breath. 15-20 repetitions from 50 typically developed Swedish children and adults (28 females, 22 males, aged 5-31 years) were collected. Sensors were placed on the upper and lower lip. We corrected for head movements by attaching additional sensors on the nose ridge and behind the right ear. Figure 1 shows the experimental set-up and the positions of the four sensors.

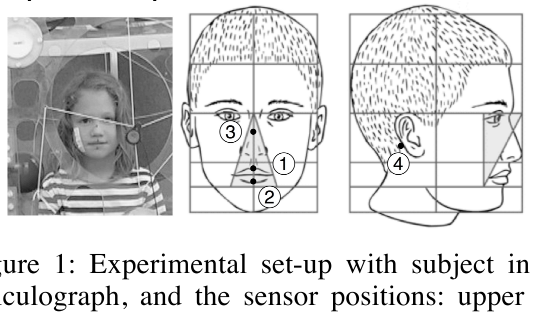

Figure 1: Experimental set-up with subject in the articulograph, and the sensor positions: upper and lower lip midsaggital on the vermilion border (1, 2), reference sensors on the nose bridge and behind the right ear (3, 4).

Kinematic landmark registration

Euclidean distances between the upper and lower lip sensors in three dimensions were calculated from the lip movement data, low-pass filtered at 25 Hz and used in the landmark registration. We delimited each token at consistent kinematic events using the first derivative of the distance function and located two points. To obtain four full cycles of opening-closing gestures of the lips, we set the onset point to the maximum velocity of the distance function in the opening phase during the transition from the first *m* to the first *a* in the word *Mamma.* For the offset point we used the same transition from the *b* to the *a* in the word *barn.* An example of the kinematic landmark registration procedure environment is shown in Figure 2. Tokens with measurement errors or artefacts were excluded from further analysis.

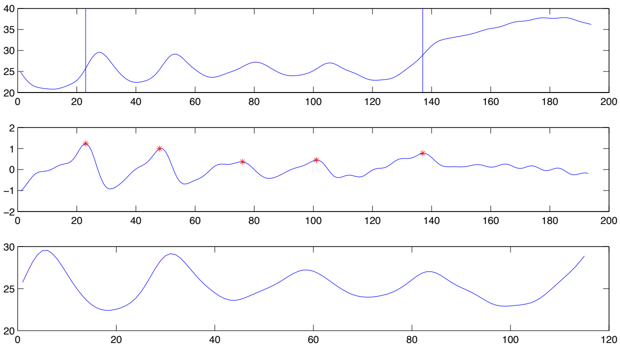

Figure 2: Lip distance function (top), its first derivative with marked velocity peaks (middle) and resulting trimmed portion (bottom) of a token during kinematic landmark registration. The vertical lines show the positions of the start and end boundaries.

Acoustic feature extraction

Phrase duration, sound pressure level (SPL) and fundamental frequency (F_0) were extracted from each repetition using Praat (Boersma and Weenink, 2012). Phrase durations were calculated based on the positions of the corresponding kinematic landmarks. SPL contours were obtained by full wave rectification, which converts the whole of the waveform to one of constant polarity, and then measuring the amplitude in 0.005 s. windows. F_0 contours were also extracted in 0.005 s. windows, and interpolated to eliminate the effect of the discontiuous contour due to the two voiceless instances of /p/. Figure 3 shows an example of duration, SPL and F_0 of one repetition.

Functional data analysis (FDA)

The landmark delimited Euclidean distance functions and the acoustic feature functions for SPL and F_0 were used as input to the FDA, a technique for time-warping and aligning a set of signals to examine differences between them. FDA techniques and applications to speech analysis were first introduced by Ramsay et al. (1996), and further developed by Lucero et al. (1997), and Lucero and Löfqvist (2005). The procedure involves the following steps: (1) temporal normalisation of the signals from a number of tokens, (2) calculation of the mean signal, (3) alignment of individual signals to the mean signal using nonlinear time-warping, and (4) computation of one index of amplitude variability and one of temporal variability (phase). Each token was amplitude normalised by subtracting its mean and dividing by its standard deviation (see Koenig et al., 2008).

We analysed the relationships between age and standard deviation of phrase duration as well as the FDA variability indices for SPL, F_0, and lip movement through correlations, scatterplots and linear regression models using the R statistical environment (R Development Core Team, 2012).

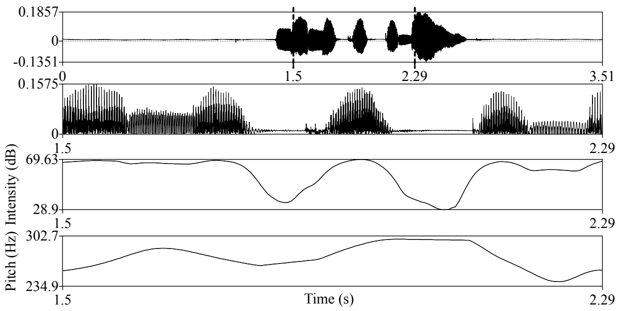

Figure 3: Example repetition with phrase duration and rectified wave based on the kinematic landmarks (upper two panes), along with SPL and F_0 contours (lower two panes).

Results

FDA amplitude and phase indices for lip movements, SPL and F_0 are plotted as a function of age in Figure 4. Table 1 shows the statistical results of the correlation and linear regression analyses, including correlation coefficients, marginal effects (per 10 years (sd(age) ~ 8)), significance levels and coefficients of determination (R^2). For phrase duration, the standard deviation was selected as the measure of variability, and these results are shown in Figure 5 and Table 2. Age significantly predicted amplitude variability for lip movements and SPL, and also explained a significant proportion of variance in amplitude variability, while the relationship for phase variability and age was weaker. The results for F_0 were not significant.

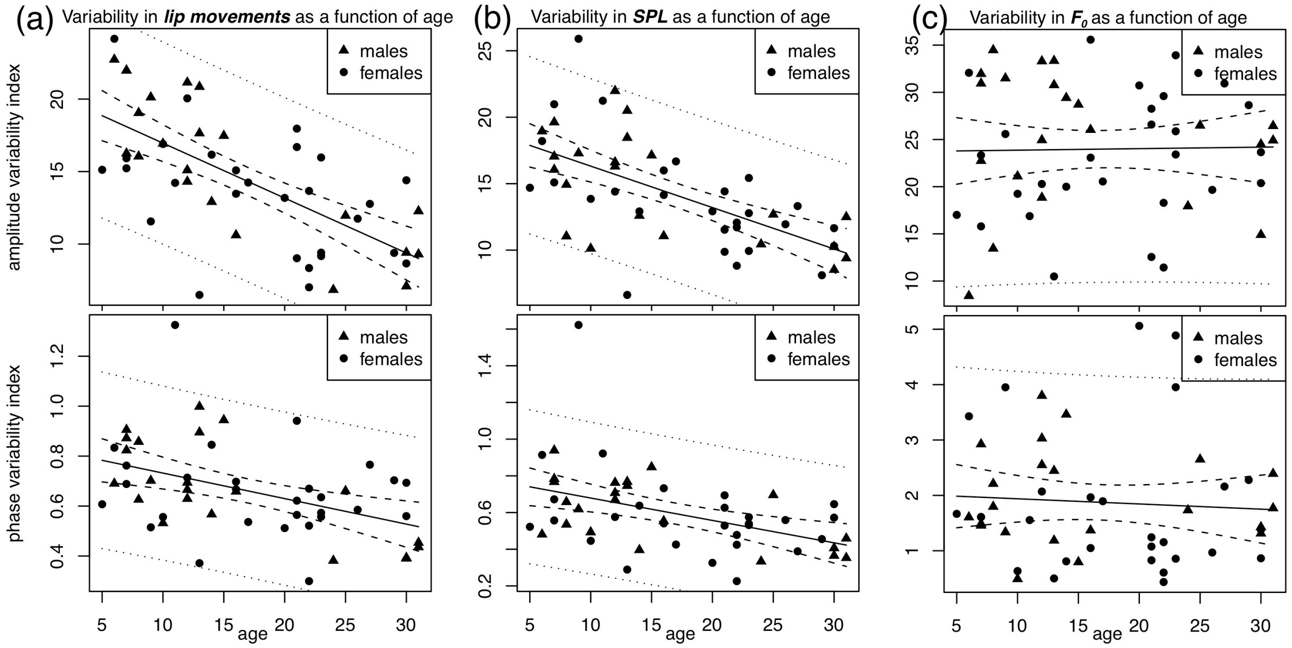

Figure 4: The relationship between amplitude and phase variability and age for a) lip movements, b) sound pressure level (SPL), and c) fundamental frequency (F_0). Scatter plots and lines of best fit (solid line: best fit, dotted lines: prediction intervals, dashed lines: confidence intervals) (n = 50).

Table 1: Statistical results of lip movements (Lip-Mov.), SPL and F_0 variability as a function of age for amplitude (AmpV) and phase (PhaV) (n = 50).

Lip Mov.	*Correlation (r)*	*Marginal effect*	*Significance*	R^2
AmpV	-0.67	-3.4	***	0.43
PhaV	-0.44	-0.10	**	0.20
SPL	*Correlation (r)*	*Marginal effect*	*Significance*	R^2
AmpV	-0.62	-3.1	***	0.39
PhaV	-0.44	-0.12	***	0.20
F_0	*Correlation (r)*	*Marginal effect*	*Significance*	R^2
AmpV	-0.02	0.16	ns	0.00
PhaV	-0.07	-0.09	ns	0.00

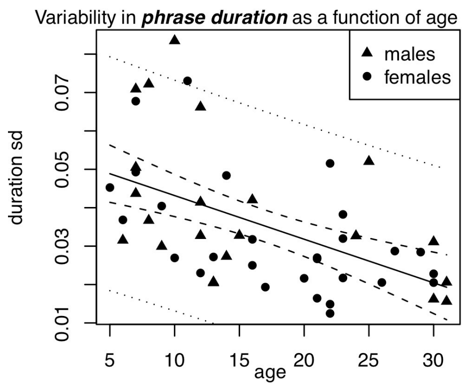

Figure 5: The relationship between phrase duration standard deviation (sd) and age. Scatter plots and lines of best fit (solid line: best fit, dotted lines: prediction intervals, dashed lines: confidence intervals).

Phrase duration variability significantly decreased as a function of age. When comparing the results of amplitude variability for lip movements with acoustic data, lip movements had the strongest negative correlation with age (-0.64) while F_0 had the weakest (-0.02).

Table 2: Numerical results of phrase duration standard deviation (stdev) as a function of age.

	Correlation (r)	*Marginal effect*	*Significance*	R^2
stdev	-0.54	-0.01	***	0.29

Discussion

The results for amplitude variability confirm the results of previous studies, i.e. that variability in kinematic as well as phrase duration decrease with age. We consistently found higher correlations for amplitude than phase in both kinematic and acoustic variability. Koenig et al. (2008) reported the same pattern. The decrease of repetition variability with age is most likely due to a combination of factors. One factor may be cerebral and cerebellar development (Kent, 1976). Another one is practice, which leads to more stable motor performance.

The correlation between lip movement variability and SPL variability is rather high (r=0.61 for amplitude and r=0.54 for phase). This is expected as a larger lip opening will increase the amplitude of the first formant and thus lift the whole spectrum of the vowel /a/.

Interpolation of the F_0 contours may have affected the FDA analysis. In order to test F_0 without interpolation we would need another phrase consisting of only voiced segments. The rather weak correlation for F_0 should be interpreted with some caution; the task at hand was not specifically designed to examine F_0 variability. Subjects did not receive explicit instructions to repeat the same intonation pattern throughout the 15-20 repetitions. Focus patterns, degree of accentuation, even list intonation and different regional backgrounds may therefore all add to the overall variability, thereby obscuring any effect of age. Another possible explanation is that children learning to speak are likely to vary their intonation, but the F_0 of adults is also conciously varied for a number of linguistic or paralinguistic reasons Therefore, variability may be a sign of developing speech motor control as well as of expert flexibility in speech. These two aspects of variability will be examined further in follow-up studies with adults and children with both typical and atypical speech development.

Acknowledgements

The authors gratefully acknowledge support from the Linnaeus environment Thinking in Time: Cognition, Communication and Learning, financed by the Swedish Research Council, grant no. 349-2007-8695. We are also grateful to J. Lucero for the use of his FDA MATLAB toolkit.

References

Boersma, P., Weenink, D. (2012) Doing phonetics by computer [Computer program]. Version 5.3.11, retrieved 3 April 2012 from http://www.praat.org/.

Frid, J., Schötz, S., Löfqvist, A. (2011a) Age-related lip movement repetition variability in two phrase positions. Proceedings of Fonetik 2011, Speech, Music and Hearing, KTH, Stockholm, TMH-QPSR, Vol. 51. pp. 21–24.

Frid J., Schötz S. and Löfqvist L. (2011b) Functional data analysis of lip movements: repetition variability as a function of age. In Proc. of ICPhS XVII, Hong Kong.

Goffman L. and Smith A. (1999) Development and phonetic differentiation of speech movement patterns. J Exp Psychol: Hum Percept Perform, 649–660.

Kent R. (1976) Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. J Speech Hear Res, 421–427.

Koenig L., Lucero J. and Perlman E. (2008) Speech production variability in fricatives of children and adults: Results of functional data analysis. J Acoust Soc Am, 3158–3170.

Lee, S., Potamianos, A, and Narayanan, S. (1999) Acoustics of children's speech: Developmental changes of temporal and spectral parameters. J Acoust Soc Am 105, 1455-1468.

Lucero J. and Löfqvist A. (2005) Measures of articulatory variability in vcv sequence. Acoust Res Lett Online, 80–84.

R Development Core Team (2012) R: A language and environment for statistical computing. Webpage: http://www.R-project.org, accessed on 17 April 2012.

Ramsay J. O., Munhall K. G., Gracco V. L, Ostry D. J. (1996) Functional data analysis of lip motion. J Acoust Soc Am, 3718–3727.

Sadagopan N. and Smith A. (2008) Developmental changes in the effects of utterance length and complexity on speech movement variability. J Speech Hear Res, 1138–1151.

Samuelsson, C. & Löfqvist, A. (2006) The role of Swedish tonal accents in children with language impairment. Clin Linguistics Phonetics 20, 231–248.

Smith A., Goffman L., Zelaznik H. N., Ying G. and McGillem C. (1995) Spatiotemporal stability and patterning of speech movement sequences. Exp Brain Res, 439–501.

Smith L. and Thelen E. (2003) Development as a dynamical system. Trends Cog Sci, 343–348.

Smith, A. (2010) Development of neural control of orofacial movements for speech. In W. Hardcastle, J. Laver, and F. Gibbon (Eds.) *The Handbook of Phonetic Sciences, 2nd edition,* pp. 251–296. Oxford: Blackwell.

Stathopoulos, E. (1995) Variability revisited: An acoustic, aerodynamic, and respiratory kinematic comparison of children and adults during speech. J Phonetics 23, 67–80

Stathopoulos, E. and Sapienza, C. (1997) Developmental changes in laryngeal and respiratory function with variations in sound pressure level. J Speech Lang Hear Res 40, 565–614.

Tang, J. and Stathopoulos, E, (1995) Vocal efficiency as a function of vocal inrensity: A study of children, women, and men. J Acoust Soc Am 97, 1885–1892.