Age-related lip movement repetition variability in two phrase positions

Frid, Johan; Schötz, Susanne; Löfqvist, Anders

Published in:
Fonetik 2011, Speech, Music and Hearing, KTH, Stockholm, TMH-QPSR

2011

Age-related lip movement repetition variability in two phrase positions

Johan Frid[1], Susanne Schötz[1], Anders Löfqvist[2]

[1]Humanities Lab, Centre for Languages and Literature, Lund University
[2]Department of Logopedics, Phoniatrics and Audiology, Lund University

Abstract

This study examined the relationship between age and lip movement variability across repetitions of an utterance. We applied functional data analysis (FDA) to lip movement data of 15-20 repetitions of a short Swedish phrase from 37 Swedish speakers (19 females, 18 males, 5-31 years) collected with three-dimensional articulography. From each utterance, three different sub-units were extracted semi-automatically by locating consistent kinetic events in the lip movement functions. Results generally showed moderate negative correlations between age and amplitude variability. The longest possible sub-unit, given consistent kinematic events, showed the strongest correlation.

Introduction

This study compared age-related lip movement variability of Swedish speakers in three kinetically delimited slices of an utterance. A number of studies using acoustic analysis (e.g., Kent, 1976; Kent and Forner, 1980; Smith, 1978) and movement recordings (e.g., Sharkey and Folkins, 1985; Smith, 1995; Smith and McLean-Muse, 1986), have shown that lip movement variability across utterance repetitions decreases with age until adolescence. Some previous studies (Goffman and Smith, 1999; Sadagopan and Smith, 2008; Smith and Goffman, 1998) have used the spatiotemporal index (STI, Smith et al., 1995), which only provides a single metric of variability (cf., Lucero et al., 1997), 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 earlier studies (Frid et al., accepted; Schötz et al., submitted), we have demonstrated that using kinematic landmarks to identify the speech segments to be analysed for variability showed more evident trends than using acoustic landmarks. We also found that the amplitude index of the FDA showed a higher age-related lip movement variability than the phase index of the FDA or the STI.

The purpose of the present study was to apply FDA to lip movements, and to compare different slices of the utterance. Our aim was to extend earlier findings of decreasing variability with age to see if utterance position affects speech movement variability. The long-term objective is to examine if children with atypical language development differ from typically developing children in terms of articulatory variability.

Method

To obtain as large lip movements as possible, we recorded the Swedish phrase *Mamma pappa barn* 'Mummy daddy children', which is short and can be spoken on a single breath. Lip movement data of 15-20 repetitions from 37 typically developed Swedish children and adults (19 females, 18 males, aged 5-31 years) were obtained along with a microphone signal using the Carstens Articulograph AG500. Sensors were placed on the upper and lower lip, and to correct for head movements also on the nose bridge and behind the right ear. Figure 1 shows the experimental set-up.

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

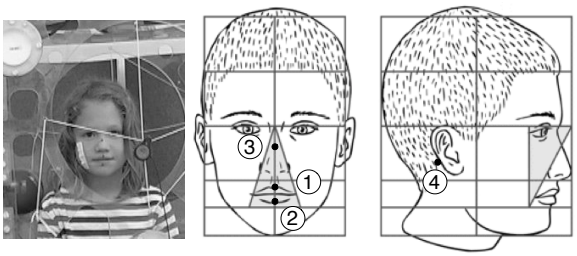

Figure 1: Experimental set-up with subject in the articulograph, and the sensor positions: upper and lower lip midsaggital on the vermilion border (1, 2), reference sensors on the nose bridge and behind the right ear (3, 4).

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.

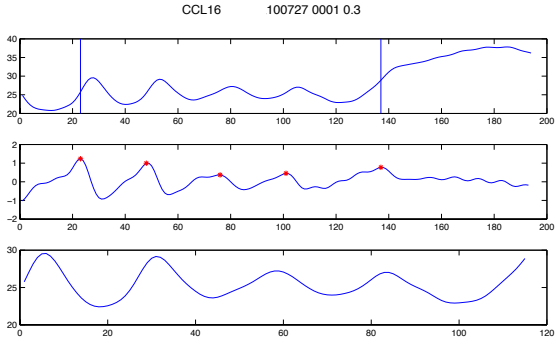

Figure 2: Lip distance function (top), its first derivative with marked velocity peaks (middle) and resulting trimmed and zoomed portion (bottom) of a token during kinematic landmark registration. The vertical lines indicate the positions of the onset and offset points described in the text above.

All tokens were further divided into two subsegments at the maximum velocity of the distance function in the opening phase during the transition from the first *p* to the first *p* in the word *pappa*. In the middle pane of Figure 2, this corresponds to the location of the third peak from the left. The sub-units contained two opening-closing cycles each. We will refer to the first sub-unit as **word1**, to the second as **word2**, and to the unit containing both words as **phrase**. Although these labels denote linguistics units to which the kinematic units are not aligned completely, we will still use them as a matter of convenience.

Functional data analysis (FDA)

The landmark delimited Euclidean distance functions 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).

Results

In a previous study with the same data (Schötz et al., submitted), we found that amplitude variability showed a stronger correlation with age than phase variability. Therefore, we will only report the results for amplitude variability here. We analysed the relationships between the three speech units, the FDA amplitude index and age through correlations, scatterplots and linear regression models using the R statistical environment (R Development Core Team, 2011). FDA amplitude indices as a function of age for the three speech units are plotted in Figure 3, while Table 1 shows the statistical results of the correlation and linear regression analyses, including correlation coefficient, slope (β), significance level, coefficient of determination (R^2) and number of samples. The results show that age significantly predicted amplitude variability, and also explained a significant proportion of variance in amplitude variability in all the speech units.

	word1	**word2**	**phrase**
Correlation (*r*)	-0.49	-0.65	-0.66
β	-0.165	-0.279	-0.326
p	0.00219	< .001	< .001
R^2	0.24	0.42	0.44
n	37	37	37

Table 1: Results of correlation and linear regression analysis between age and the FDA amplitude variability index for the three speech unit conditions.

Paired-samples t-tests were conducted to compare the amplitude variability indices in the different

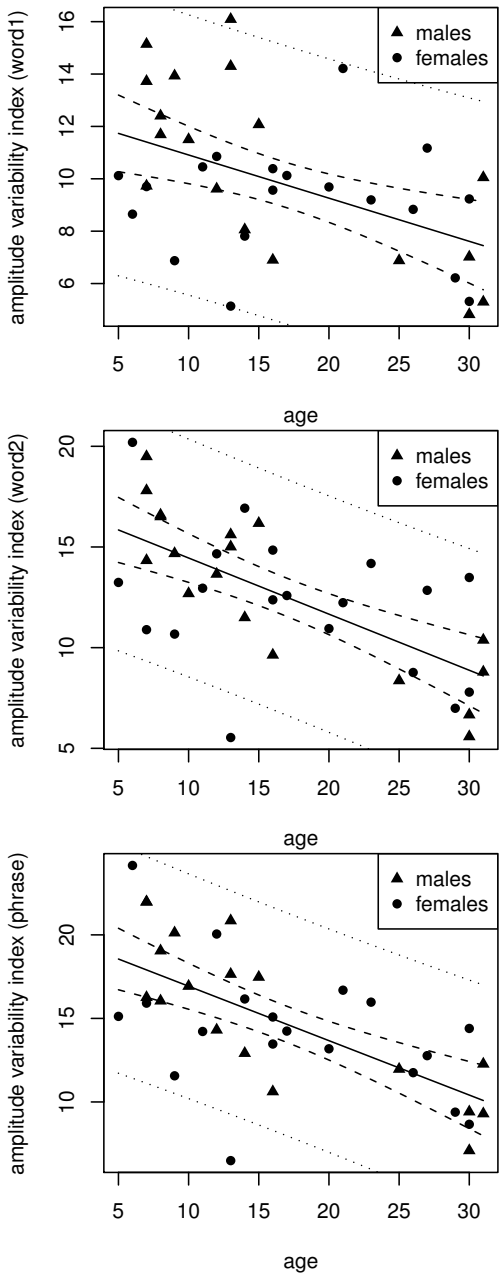

Figure 3: Amplitude variability as a function of age (solid lines), prediction intervals (dotted lines), and confidence intervals (dashed lines) for each of the three speech units conditions .

speech unit conditions. Means and standard deviations (SD) are given in Table 2. There was a significant difference in the scores for **word1** and **word2** conditions, t(36) = 6.49, $p < .001$. Furthermore, there were significant differences in the scores for **word1** and **phrase**, t(36) = 11.63, $p < .001$ and for **word2** and **phrase**, t(36) = 7.70, $p < .001$.

One of our motivations for splitting up the tokens into smaller units was to compare the results of each sub-unit with the result of the whole utterance. As the variability index scales differed in

	word1	**word2**	**phrase**
Mean	9.80	12.58	14.74
Median	9.70	12.85	14.40
SD	2.91	3.70	4.26

Table 2: Means, medians and standard deviations of the FDA amplitude variability index for the three speech unit conditions.

word1 and **word2**, they were rescaled using the means and standard deviations. We then calculated the correlation between age and the rescaled and combined amplitude variability indices. The two variables were correlated, $r(72) = -0.57$, $p < .001$, but the correlation was smaller than the one we obtained between age and **phrase** (-0.66).

Discussion and Future Work

The results for amplitude variability confirm the results of previous studies, i.e. that lip movement variability decreases with age. In (Frid et al., accepted) and Schötz et al. (submitted) we found higher correlations for amplitude than phase in both acoustic and kinematic landmarks. Koenig et al. (2008) reported the opposite pattern, with more variability for phase than amplitude. Those results were, however, based on records of airflow during fricative production, thus reflecting both articulatory and expiratory factors. The current results are based on articulatory movements alone. Similar developmental changes have been observed in non-speech motor activities such as reaching and finger tapping (Deutsch and Newell, 2003, 2004). 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. It is also likely that a developing and changing system will show increased motor variability during transitions, when a new mode of organisation is replacing an old one (Smith and Thelen, 2003).

In this study, the correlation between age and variability was almost the same for **word2** as for **phrase**, but weaker for **word1**. There are a few possible explanations for this. One is that **word2** has a longer duration than **word1**. The segmental content is also different: **word1** contains something like >amap< (one plosive and one nasal), wheras **word2** consists of the sequence >apab< (two plosives).[1] Another explanation may be that

[1]The 'greater than' and 'less than' are used to symbolise in- and outgoing transitions.

the phrase positions differ: **word1** is initial, while **word2** is medial in the phrase. It is possible that the initial position offers a potential anchoring point for articulation, and therefore obscures any age-related effects. It could also be that variability is revealed better in prominent words or sub-units. The utterance in this study was produced with a broad focus by all subjects, i.e. with the highest prominence on the final word *barn*, of which the *b* is included in **word2**.

Splitting up the utterance into sub-parts increased the sample size (by a factor of 2), but it did not yield a stronger relationship between age and amplitude variability. This is an interesting finding, which we would like to examine more thoroughly in the future. It would also be interesting to compare variability in different prosodic contexts. In further studies, we will also record not only more typically developed children, but also atypically developed children. Future work also includes an examination to see if children with atypical language development differ from typically developing children in terms of articulatory variability. We also want to examine the possible relationship of our results with cerebellar function as assessed by the blink reflex.

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

Deutsch K M and Newell K M (2003). Deterministic and stochastic processes in children's isometric force variability. *Dev Psychobiol*, 335–345.

Deutsch K M and Newell K M (2004). Changes in the structure of children's isometric force variability with practice. *J Exp Child Psychol*, 319—-333.

Frid J, Schötz S and Löfqvist L (accepted). Functional data analysis of lip movements: repetition variability as a function of age. In *Proc. of ICPhS XVII, Hong Kong, August 17-21, 2011*.

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.

Kent R and Forner L (1980). Speech segment duration in sentence recitation by children and adults. *J Phonetics*, 157–168.

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.

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

Lucero J, Munhall K, Gracco V and Ramsay J (1997). On the registration of time and the patterning of speech movements. *J Speech Hear Res*, 1111–1117.

R Development Core Team (2011). R: A language and environment for statistical computing. Webpage: http://www.R-project.org, accessed on 11 Mar 2011.

Ramsay J, KG M, VL G and DJ O (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.

Schötz S, Frid J and Löfqvist L (submitted). Lip movement repetition variability as a function of age: A comparison of two landmark registration strategies. In *Proc. of Interspeech, Florence, Italy, August 27-31*.

Sharkey S G and Folkins J W (1985). Variability in lip and jaw movements in children and adults: Implications for the development for speech motor control. *J Speech Hear Res*, 8–15.

Smith A and Goffman L (1998). Stability and patterning of speech movement sequences in children and adults. *J Speech Hear Res*, 18–30.

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 B (1978). Temporal aspects of english speech production: A developmental perspective. *J Phonetics*, 37–67.

Smith B and McLean-Muse A (1986). Articulatory movement characteristics of labial consonant productions by children and adults. *J Acoust Soc Am*, 1321–1328.

Smith B L (1995). Variability of lip and jaw movements in the speech of children and adults. *Phonetica*, 307–316.

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