

Movement of Formants of Sound *a* in Lithuanian Language

D. Balbonas, G. Daunys

Department of Electronics, Siauliai University,

Vilniaus str. 141, LT-76353 Siauliai, Lithuania, tel.: +370 41 595830; e-mail: dainius@tf.su.lt

Introduction

In natural speech there is a substantial variation in vowel realizations, even when spoken by a single person. The changes are generally divided into two groups: coarticulation and reduction.

Coarticulation causes individual vowel realizations to become more similar to their neighboring phonemes in the utterance.

Vowel reduction causes realizations of different vowels to become more like the neutral (schwa) vowel.

Phoneme is a basic, theoretical unit of sound that can distinguish words (that is, changing a phoneme in a word, produces another word. Each phoneme may have one or more corresponding allophones. An allophone is an acoustic manifestation of a phoneme. A particular phoneme may have many allophones, each sounding slightly different due to the position of the phoneme in a word or variant pronunciations in a language of the same letter set [1]. Also the phoneme can be assimilated by other phoneme.

A formant is a concentration of acoustic energy around a particular frequency in the speech wave. The formants well represent vowel sounds. But it is well established that formant frequencies changes during length of syllable. The changes could be tried to explain because of reduction.

L. Kaukiene [2] analysed the spectra of unstressed vowels in Standard Lithuanian, their acoustic and articulations characteristics. The unstressed vowels were compared with the stressed ones. She found that low rise vowels [a], [e] have the biggest qualitative reduction.

Regressive influence of palatalised consonants on the spectrum of the short vowels in Standard Lithuanian was investigated by R Kliukienė [3]. She found that palatalised consonants increase frequency of second formant.

Authors tried to find out how to explain formant movement of Lithuanian sound *a*. The first idea was that so happens because of reduction. But our investigation revealed that for first two formants most influence have stress position in the word.

Method

The authors recorded the speech of three speakers; two males and one female. The speech was recorded in quiet environment using simple computer microphone and sound card. Each speaker's speech consists of the same one hundred words with sound *a*. The words selected for speech was elected from 1200 frequently used words in Lithuanian language. The speaker's speech file was divided into separate sound files and then using Praat software sound *a* and transition between sound *a* and other sounds were marked [4 Regressive influence of palatalised consonants on the spectrum of the short vowels in Standard Lithuanian]. Marking was made manually. Information about marking was saved in TextGrid files (Praat extension). Each word has it's own TextGrid file that contains information about location of syllables and empty intervals in the word. Also file contains information about the names of syllables, the length of word and all marked intervals in the word.

When using specialized Praat script the points of F0, F1, F2, F3 F4 F5 formants at every 10ms were extracted from wave file to Excel data sheet, for each word. When using the Matlab program, Excel data sheet and TextGrid files the point of formants of sound *a* for each word was extracted manually. When using points of F0, F1, F2, F3, F4, F5 formants the third degree polynomial for each formant was produced. Overall authors got 158 syllables for each speaker. The polynomials were corrected when wrong points of formants were removed. Wrong points were removed using standard deviation and 3 sigma rule.

All polynomials of sound *a* were classified by type of syllable. Possible types of syllable described by syllable structure made by P.Kasparaitis [5].

Description of the model: all syllables in Lithuanian language conform to STRARTSK structure, where: S {s,š,z,ž}, T {b,d,g,k,p,t,c,č,dz,dž,c,h,h,f}, R {j,l,m,n,r,v}, K {k,t}, A – any vowel.

From 100 words, that took part in experiment, authors got 158 syllables, classified them into 23 different types : AT, ATS, SA, RA, TRA, STA, RAR, RAT, RAS, SAR, SAT, SAS, TAR, TAT, TAS, TRAS, SRAS, STAR, STAT, TARS, TART, RATS. For each type of syllable we

got a group of phone “a” polynomials. For example the syllable group SA consists of two different syllables “sa” and “ša”. For each syllable we got one word and one word was pronounced by three different speakers. In final we got six polynomials of each formant in this group of syllables.

The coefficients of polynomials were used to compare sound *a* in different syllables of different speakers.

The artificial neural networks (ANN) approach was used to find approximation of formants frequency on factors. The all data was divided into three groups for training, validation and testing. So the earlier stopping was achieved.

The attempts were done to find smallest set of factors which best describe formants frequencies.

Results and discussion

The analysis of formants frequencies distribution of sound *a* for three subjects was done. The results are presented in the plot in Fig. 1 for first formant, in Fig. 2 for second formant, and in Fig. 3 for third formant. As one could see the variation of frequency is big range for the same subject. For female subject there is shift to higher frequencies.

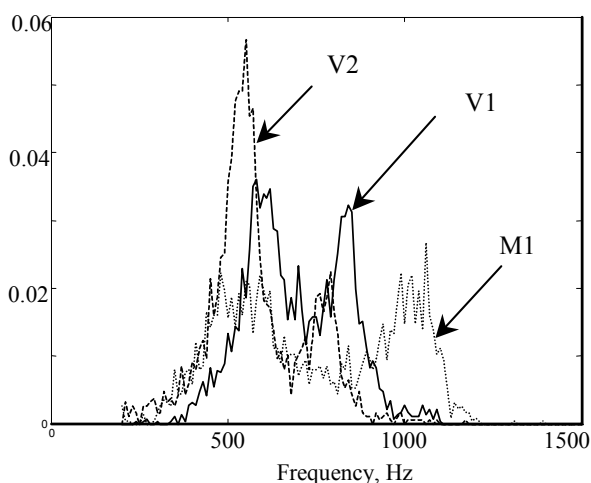


Fig. 1. Distribution of formant F1 frequency for three subjects

The results of qualitative analysis of formants movement are presented in Table 1 (formants F1 and F2) and Table 2 (formants F3 and F4). The dominating style is oscillations about horizontal line. The second result depend on formant.

For the first formant next dominant case is curvature with maximum in the middle of the sound. For the second often occurs case, formant frequency goes up or opposite – down. For the third formant we could see dependence on the subject.

There was not found relation between sound *a* surrounding and formant movement. So attempts to describe dependence upon other parameters was done. The best results were obtained, when polynomial coefficients of intensity regression versus time were used together with sound duration and time moment of sample.

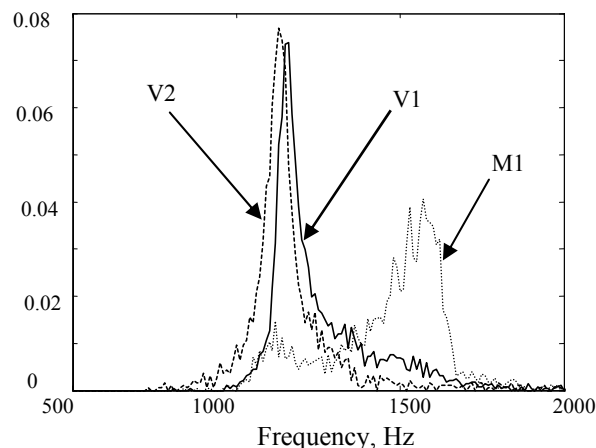


Fig. 2. Distribution of formant F2 frequency for three subjects

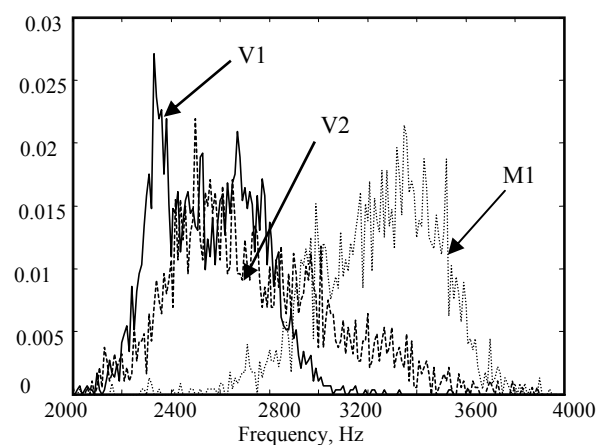


Fig. 3. Distribution of formant F3 frequency for three subjects

Table 1. Style of formant F1 and F2 movement

Formant	F1			F2		
	V1	V2	M1	V1	V2	M1
Constant	47	47	41	62	70	47
Maximum	39	29	37	3	3	9
Minimum	0	0	0	4	4	5
Down	12	16	17	16	9	16
Up	2	8	6	16	15	22
	100	100	100	100	100	100

Table 2. Style of formant F3 and F4 movement

Formant	F3			F4		
	V1	V2	M1	V1	V2	M1
Constant	46	34	42	20	51	32
Maximum	3	15	6	6	2	2
Minimum	13	27	7	42	18	17
Down	6	13	6	6	7	15
Up	33	11	38	26	22	34
	100	100	100	100	100	100

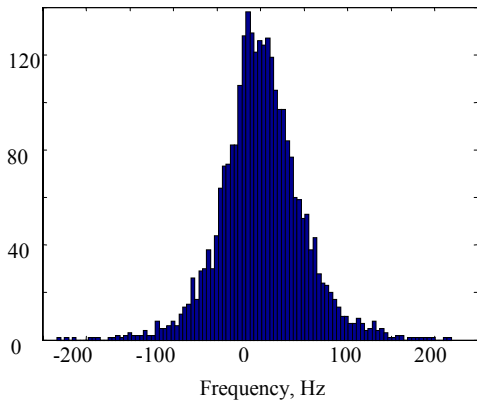


Fig. 4. Distribution of formant F1 frequency errors after approximation with ANN for subject V1

The distribution of remaining errors is plotted in Fig. 4. The distribution is close to Gaussian distribution. The interval, where the errors most frequently occur is ± 100 Hz.

The simulation results of formant movement versus time, when other parameters are constant (they have their mean values of stress group) are shown in Fig. 5. The four cases of stressed syllables are used.

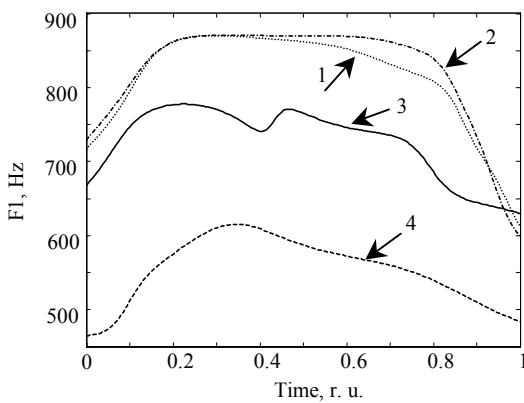


Fig. 5. Formant F1 movement during utterance of sound *a* versus normalized time in case of differently stressed syllable (1 – acute accent, 2 – circumflex, 3- grave, 4 – non-stressed)

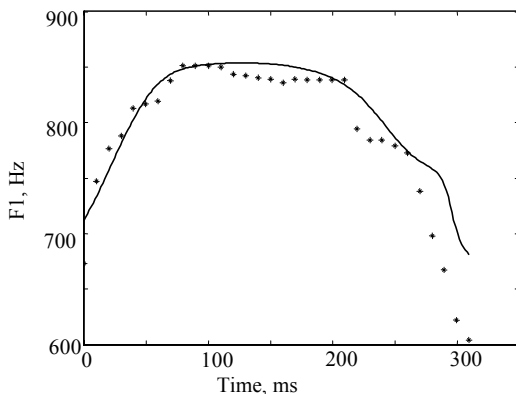


Fig. 6. Formant F1 movement during utterance of first sound *a* in Lithuanian word “baltas”

In the Fig. 6 is presented an example how well the simulation results correspond to experimental values in the case of the first sound *a* in the Lithuanian word “baltas”. In

the Fig. 7 similar plot for first *a* sound in Lithuanian word “kartas”.

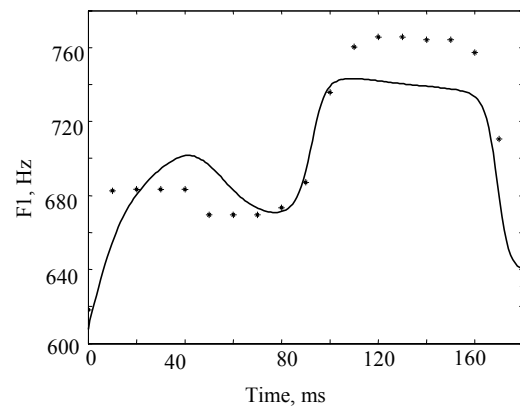


Fig. 7. Formant F1 movement during utterance of first sound *a* in Lithuanian word “kartas”

The distribution of remaining errors after approximation by ANN for formant F2 is plotted in Fig. 8. The distribution even more is concentrated to zero than for first formant. The simulation results of formant F2 movement versus time when other parameters are constant are shown in Fig. 9.

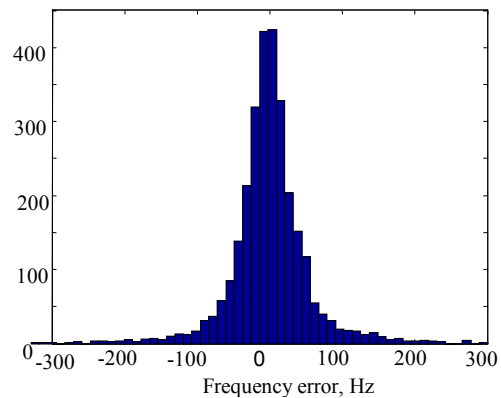


Fig. 8. Distribution of formant F2 frequency errors after approximation with ANN for subject V1

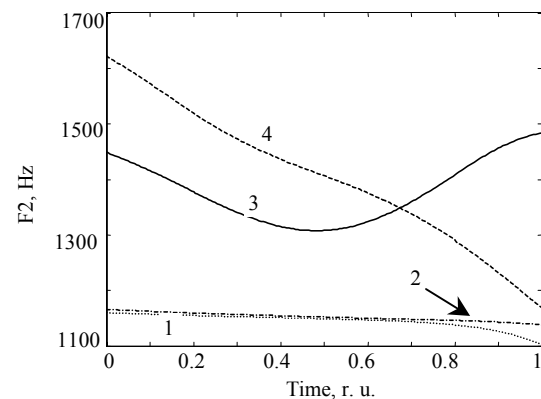


Fig. 9. Formant F2 movement during utterance of sound *a* versus normalized time in case of differently stressed syllable (1 – acute accent, 2 – circumflex, 3- grave, 4 – non-stressed)

Two examples of F2 formant movement of unstressed syllables are plotted in Fig. 10 and Fig. 11. In both cases big formants movements are well simulated with ANN.

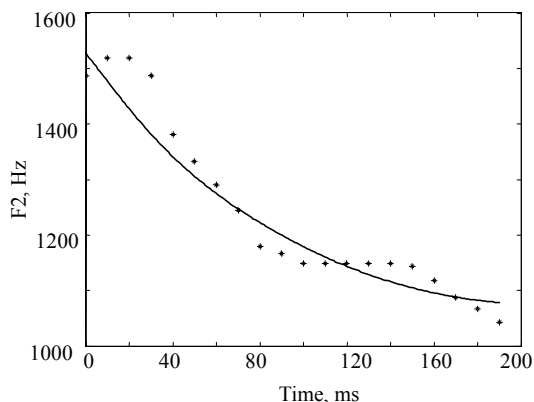


Fig. 10. Formant F2 movement during utterance of first sound /a/ in Lithuanian word “dabar”

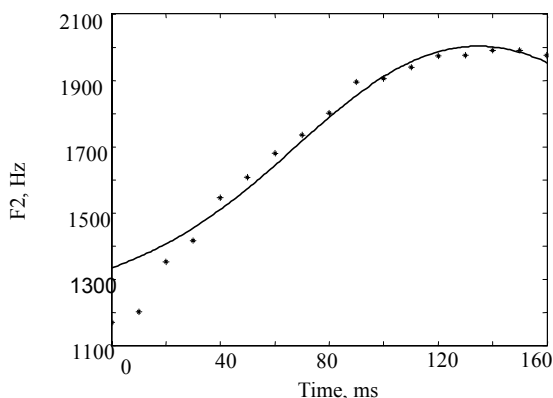


Fig. 11. Formant F2 movement during utterance of sound /a/ in Lithuanian word “bažnyčia”

Conclusions

The formants movement depends upon stress. The good approximation of experimental results of F1 and F2 by ANN was obtained, when the polynomial coefficients of intensity approximation upon time were used as input parameters together with sound duration and current moment time.

The aim for the next research is to search formant movement relation with the same formant intensity versus time dependence coefficients.

References

1. **Suzuki H., Zen H., Nankaku Y., Miyajima C., Tokuda K., Kitamura T.** 'Acoustic modeling in consideration of unknown variation factors at the time of recognition. Technical Report of IEICE, SP2003-139, Dec. 2003
2. **Kaukėnienė L.** The spectra of unstressed vowels in standard Lithuanian // *Kalbotyra*. – 2002. – LI (1).
3. **Kliukiene R.** Regressive influence of palatalised consonants on the spectrum of the short vowels in Standard Lithuanian // *Kalbotyra*. – 2002. – LI(1). – P. 73–78.
4. **Praat project homepage.** Page can be accessed by Internet: <http://www.praat.org/>.
5. **Kasparaitis P.** Computer synthesis of Lithuanian speech. Doctor thesis. – Vilnius: University of Vilnius, 2001 (in Lithuanian).

Submitted for publication 2006 05 25

D. Balbonas, G. Daunys. Movement of Formants of Sound *a* in Lithuanian Language // Electronics and Electrical Engineering. – Kaunas: Technologija, 2006. – No. 8(72). – P. 75–78.

More frequently used words with different sound *a* surrounding were selected for investigation. Praat program was used for words annotation into phonemes. the sound *a* sections were annotated by program Praat. The values of formants also were calculated by Praat program. Plot of formants frequency distribution shows a big spread of histogram. We thought that wide distribution is because of coarticulation. Dividing to groups by surrounding phonemes didn't revealed reasonable results. Also approximation formants by other sound parameters as duration, position from beginning, intensity didn't showed reasonable results. The best approximation was obtained, when to input vector were included polynomial coefficients of signal's intensity regression by reduced time. Ill. 11, bibl. 5 (in English, summaries in English, Russian, Lithuanian).

Д. Бальбонас, Г. Даунис. Движение формант звука *a* в литовском языке // Электроника и электротехника. – Каунас: Технология, 2006. – № 8(72). – С. 75–78.

Были отобраны более часто используемые слова, в которых и различном контексте находится звук *a*. Программа Praat была использована для аннотации звуковых файлов. Значения формант тоже были рассчитаны с помощью программы Praat. Была произведена аппроксимация значений формант искусственными нейронными сетями. Лучшая аппроксимация была получена, когда были использованы полиномиальные коэффициенты разложения во времени. Ил. 11, библи. 5 (на английском языке, рефераты на английском, русском и литовском яз.).

D. Balbonas, G. Daunys. Lietuvių kalbos garso *a* formantų kitimas // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2006. – Nr. 8(72). – P. 75–78.

Tyrimui buvo pasirinkti dažniausiai lietuvių kalboje pasitaikantys žodžiai, kuriuose garso *a* apsuptis yra įvairi. Praat programa žodžiuose buvo pažymėti garso *a* intervalai ir nustatyta, kaip keičiasi keturi pirmieji formantai. Nubraižius formantų dažnių pasiskirstymo diagramas, buvo gauta didelė sklaida. Buvo manyta, kad formantų dažniai keičiasi dėl koartikuliacijos. Tačiau skaidymas pagal skiemenų tipus ar gretimas fonemas jokių apčiuopiamų rezultatų nedavė. Atlikti bandymai, siekiant formantų kitimą aproksimuoti pagal kitus garsų parametrus, tokius kaip trukmė, laikas nuo garso *a* tarimo pradžios, intensyvumas. Tačiau buvo gauti dideli liekamieji nuokrypiai ir formantų kitimas skiemens viduje nebuvo paaiškintas. Geriausi rezultatai buvo gauti, kai į neuronų tinklo įėjimus buvo perduotos ne tik intervalo ilgio ir padėties intervale vertės, bet ir intensyvumo skleidimo pagal laiką polinomo koeficientai. Manoma, kad šitaip automatiškai nustatomas žodžio kirčiavimas. Il. 11, bibl. 5 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).