

Cross-linguistic Comparison of Two-year-old Children's Acoustic Vowel Spaces: Contrasting Hungarian with Dutch

Krisztina Zajdó¹, Jeannette M. van der Stelt², Ton G. Wempe² & Louis C.W. Pols²

¹ Division of Communication Disorders
University of Wyoming, Laramie, WY, USA
Kzajdo@uwyo.edu

² Institute of Phonetic Sciences/ACLC
University of Amsterdam, Amsterdam, The Netherlands
Jeannet.van.der.Stelt@, Ton.Wempe@, Louis.Pols@uva.nl

Abstract

Traditional hand-edited formant measurements may result in biased assessment of vowel formants in children's speech. Therefore, vowel spaces that are constructed by hand-edited formant measures may be unreliable. The recent development of an automated frequency domain analysis method allows for more reliable measurements. Thus, a valid comparison of the size and positioning of young children's vowel spaces across languages can be achieved. Contrasting the extension of the vowel space utilized by young children acquiring Hungarian and Dutch can provide information pertaining to a) children's abilities to explore the vowel space and b) potential cross-linguistic differences in exploiting the potentially available acoustic vowel space. Since a unified theory of vowel acquisition has never been developed, it is hoped that the new method will contribute to the creation of such a theory by comparing and contrasting results from diverse languages. Results suggest that two-year-old Hungarian- and Dutch-speaking children utilize the vowel space language-specifically, by exploiting different regions within the potentially available acoustic space.

1. Introduction

In the last decade, an increasing number of studies have aimed at exploring vowel acquisition processes in young children [1], [2], [3], [4]. One major motivation for these studies is to develop a thorough understanding of children's vowel production processes, to formulate a universal vowel acquisition theory. The generation of such a theory has been clearly motivated by clinical need, since currently there are no well established therapeutic approaches to remediate disordered vowel production in children and adults.

Unquestionably, a universal (non-language-specific) vowel acquisition theory can only be generated by studying vowel acquisition processes in children and adults acquiring diverse languages. However, the cross-linguistic comparison of vowel production patterns is an unexplored area of research. The aim of this project was to investigate, by using a new automated frequency domain analysis method, the size and positioning of the acoustic vowel spaces in Hungarian- and Dutch-speaking two-year-old children, to examine

whether children utilize the potentially available acoustic vowel space differently in the two languages at this age.

Thus, the current study constitutes a much-needed step towards establishing cross-linguistic vowel acquisition research as a new area of investigation that contributes to our understanding of vowel acquisition and production processes in children. It is hoped that this area of research will provide results that will benefit clinical thinking pertaining to vowel disorders and vowel production difficulties both in children and adults.

2. Data sets

This study utilized two pre-existing data sets for the purposes of cross-linguistic comparison of Hungarian- and Dutch-speaking children's vowel production in the two languages. The adult vowel inventories of the two languages are shown in Figures 1 & 2.

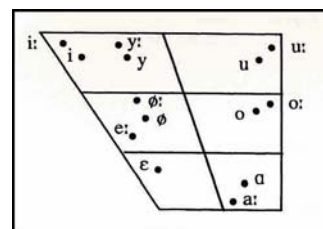


Figure 1: The adult vowel inventory of the standard dialect of Hungarian. This dialect lacks diphthongs [5.]

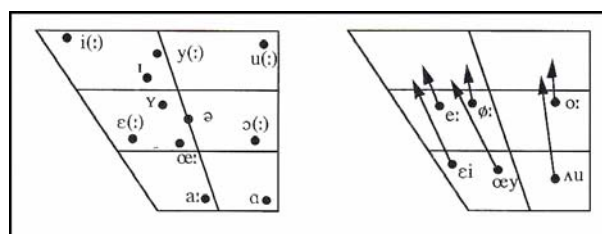


Figure 2: The adult vowel inventory of the standard dialect of Dutch. Left figure: Dutch monophthongs. Right figure: Dutch diphthongs and diphthongized vowels [6.]

Even though the two vowel inventories are fairly similar in the two languages, and the corner vowels /i(:)/ and /a:/ are fairly comparable perceptually, the articulation of the high back corner vowel /u(:)/ differs in the two languages. In Hungarian, /u(:)/ is produced with considerable lip rounding, even in running speech. In Dutch, this vowel is formulated without much rounding.

2.1. The Hungarian data set

Participants included 8 monolingual Hungarian-speaking boys who were ± 2 weeks from their second birthday. The children lived in the Budapest area. Speech samples were recorded in a sound treated room. The Sound Forge acoustic software was utilized for recording with recording attributes set to 32 kHz, 16 bit, mono. The children participated in a naturalistic interaction with their caregiver(s). While playing with puppets that had CV(:)CV(:) structured pre-assigned names (e.g., /pipi/, /tu:tu:/) containing each vowel from the Hungarian vowel inventory in two different kinds of consonantal environments (e.g., the vowel /a:/ was embedded in tokens /pa:pa:/ and /ka:ka:/), the caregivers were instructed to elicit the production of the tokens by generating everyday conversations about the puppets (see [7] and [8] for details.)

2.2. The Dutch data set

Participants included 5 monolingual Dutch-speaking two-year-old boys. Researchers recorded the children in a home setting, by using an analog Sony recorder. The children conversed with their mother. The data was digitized with recording attributes set to 48 kHz, 16 bit, mono. For the technical details, see [8.]

2.3. Limitations of direct comparison

Even though there are differences in how the data were obtained from children speaking the two languages (e.g., Hungarian children's speech occasionally included tokens that were produced in a citation format), it is argued that the two data sets are suitable for a comparison of acoustical vowel spaces across the two languages, especially because of the properties of speech in two-year-old children (e.g., relatively slow speech rate, limited syntax with $MLU \approx 2.0$.)

3. Methodology

Until recently, a major obstacle to researching vowel production in young children was that the hand edited formant measurement method yielded questionable results.

3.1. Limitations of traditional hand edited formant measurement approaches

Data from the literature clearly suggest that traditional hand-edited formant measurements can easily be biased by human perception [9], [10], [11.] An additional obstacle to establishing valid formant measures is the high fundamental frequency of young children's voice that results in an undersampling of the spectral envelope. Further, an inability to separate the source and filter functions also impedes the

reliable measurement of formants. Taken together, these problems do not allow for valid formant measures to be determined through the traditional (hand edited) formant measurement method.

3.2. Advantages of new automated frequency domain analysis method

A newly developed pitch-related band-filter analysis method allows for an acoustically unbiased approach to establishing formant measurements (for details, see [8], [12].) The automated analysis method estimates the spectral envelope representation of the utterance.

3.3. Selection criteria

Individual utterances from all children from the two language communities were subjected to analysis as unlabeled tokens. That is, results are not biased by human perception since the phonemic attributes of the tokens are undetermined. Specified intensity criteria were introduced to avoid clipped utterances and low intensity fragments. Only utterance parts with an F_0 below 425 Hz were selected (see next section.)

3.4. Analysis method

To be able to bandfilter the selected part, the F_0 period nearest to the selection is recirculated to get a sufficiently long signal. The filtering of this signal is carried out with a Gauss bandfilter, tuned from 0 to 7000 Hz in 40 steps of 175 Hz. Its bandwidth was set to 1.1×425 Hz, which was a result of a compromise between frequency resolution and F_0 -ripple. The spectral envelope thus obtained contains spectral intensities of 40 adjacent frequency bins of 175 Hz width. Each measurement can be regarded as a point in a 40-dimensional space. Principal Component Analysis transforms this space into a new space (defined by its eigenvectors), whereby the first dimension represents the largest variance, the next dimension the second largest variance, and so on. Using the first two dimensions, a plane can be drawn which displays projections of all measurements onto this plane. When the first two dimensions explain a substantial part of the total variance, this projection can be regarded as a valid representation of the "spectral space" of the selected set [8.]

All selection and analysis procedures were realized with scripts in the "Praat" program [13.]

4. Results

Results of the automated frequency analysis method are displayed on Figures 3-5 and Table 1.

4.1. Projecting the vowels in the reference planes

To allow for a cross-linguistic comparison of language-specific acoustic vowel spaces, a common reference plane that is determined by the pc1-pc2 values was created by using unlabeled data from both languages. The construction of the common reference plane required the generation of two comparable samples from the two languages. From both languages, an equal number of 229 vocalizations were selected.

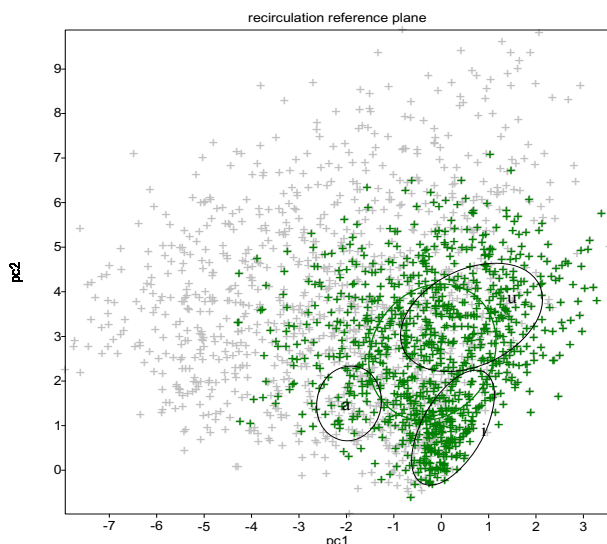


Figure 3: Common reference plane created by the projection of measurement points in $(2 \times 229 =) 458$ utterances for Dutch and Hungarian: 2071 grey pluses (+) represent the spectra of the measurement positions. The green pluses represent the spectra of the measurement positions of the Hungarian data. The green (middle) ellipse represents 1 sd. distance from the mean of the measured variance for the Hungarian vowel measures. Hungarian corner vowel regions are also displayed; the black ellipses represent 1 sd from the mean of the measured variance for each corner vowel.

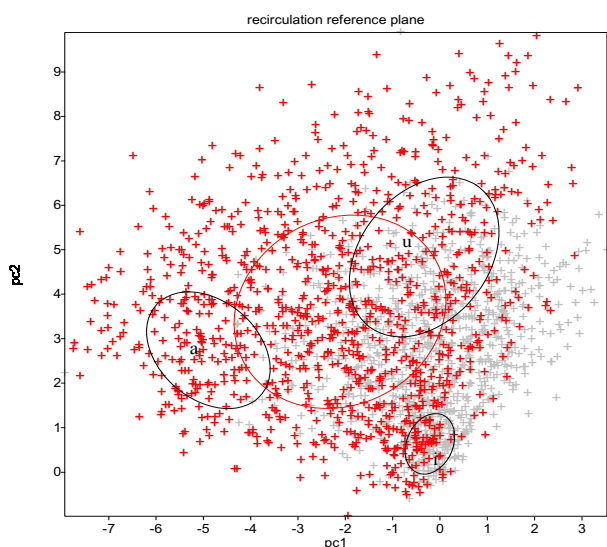


Figure 4: Common reference plane created by the projection of measurement points in $(2 \times 229 =) 458$ utterances for Dutch and Hungarian: 2071 grey pluses (+) represent the spectra of the measurement positions. The red pluses represent the spectra of the measurement positions of the Dutch data. The red (middle) ellipse represents 1 sd. distance from the mean of the measured variance for the Dutch vowel measures. Dutch corner vowel regions are also displayed; the black ellipses represent 1 sd from the mean of the measured variance for each corner vowel.

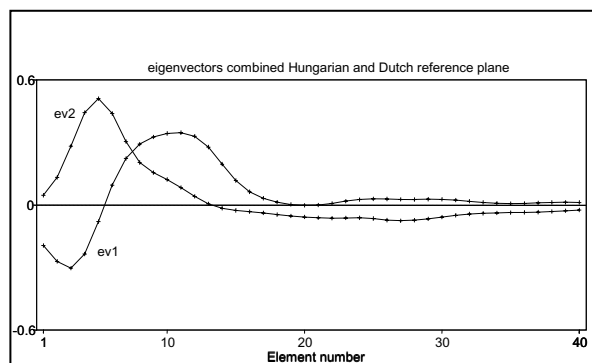


Figure 5: Graphs of the first two eigenvectors of a PCA of the measurement points found for the combined Hungarian and Dutch data sets, explaining 52.1 % of the total variance.

eigenvector 1	28.0%	28.0%
eigenvector 2	24.1%	52.1%
eigenvector 3	12.6%	64.7%
eigenvector 4	10.0%	74.7%

Table 1: Percentage of variance explained by each of the first eigenvectors in the PCA for the analyzable utterances of the two-year-old Hungarian and Dutch children (resulting in 1083 measurement points.) The cumulative percentages are given as well.

4.2. Projecting the corner vowels in the reference planes

To interpret the findings, corner vowels of the two-year-olds' that were perceptually judged to be most adult-like by the first and second authors (native speakers of Hungarian and Dutch, respectively) were projected in the F_1 - F_2 plane/ $pc1$ - $pc2$ plane in both languages. The black ellipses represent 1 sd from the mean of each corner vowel in each language (see Figures 3 & 4; for details, consult [8].)

5. Discussion

Results suggest that two-year-old children speaking the standard dialects of Hungarian and Dutch indeed utilize the potentially available vowel plane differently, by producing some vowels that occupy different regions of the common vowel space.

Recently, Serkhane et al. [14], [15] speculated that young children are still exploring the workings of their vocal tract during the earliest phases of speech acquisition. Potentially, this exploration then would result in the production of vowels characterized by a great diversity. Thus, one might speculate that vowel production in two-year-old children is more biologically and motorically determined than it is constrained by language-specific vowel categories that are present in the environment. This argument would suggest that no language-specific differences may be observed in young children's vowel production.

The PCA-plane can be interpreted as being similar but not necessarily identical to a (rotated) F_1 - F_2 plane [16.] In doing so, it appears that the majority of Hungarian-speaking

children's vowel realizations occupy the region of the common vowel space that corresponds to the high vowels. At the same time, a region of the common vowel plane that corresponds to low and especially low front vowels is utilized much less. In contrast, Dutch-speaking children do produce vowels in the low front region of the F_1 - F_2 plane/pc1-pc2 plane, while not producing vowels that are as high and especially as high back as some of the Hungarian ones. For further evaluation of these acoustic results, they must still be compared with results from articulatory studies.

One possible interpretation of the results is that while the /u(:)/ in Hungarian is characterized by considerable lip rounding that places these vowels to the high back corner of the F_1 - F_2 plane/pc1-pc2 plane, the relatively unrounded Dutch /u(:)/ vowel production corresponds to lower and more frontal regions of the common plane. Thus, lip-rounding differences may explain the different distributions of vowels in the high back region of the plane cross-linguistically.

The interpretation of the language-specific difference in the utilization of the low front region of the F_1 - F_2 plane/pc1-pc2 plane is more problematic. One argument may be that Dutch-speaking children, similarly to adults, produce the low vowel /a:/ with a more frontal tongue position than their Hungarian-speaking peers. Further, the low front region is more crowded by vowels in Dutch than in Hungarian. However, close inspection suggests that it is the whole frontal region of the common plane that is utilized by Dutch-speaking children to a greater extent than by their Hungarian-speaking peers.

In this study, two existing corpora of sounds were used to compare vowel space utilization patterns cross-linguistically. Woefully, the two datasets differ in various ways. The Hungarian dataset can be considered as (deferred) imitated adult speech, whereas the Dutch data were collected from spontaneous mother-child interaction. This difference may explain the rather restricted area Hungarian children utilize from the potentially available vowel space.

Further research will have to reveal the importance of language-specific features in young children's vowel plane utilization patterns.

6. Acknowledgements

The first author gratefully acknowledges funding for this project that was awarded by the Graduate School and the Center for Mind, Brain and Learning of the University of Washington, Seattle, USA.

Special thanks are due to the Laboratory of Speech Acoustics at the Department of Telecommunications and Telematics of the Budapest University of Technology and Economics for encouragement and support.

Additionally, heartfelt thanks are expressed to the College of Health Sciences and the Division of Communication Disorders at the University of Wyoming for additional funding.

An "International Travel Award" from the International Programs Office of the University of Wyoming is also much appreciated. Without this award, the establishment of this international cooperative project would not have been possible.

7. References

- [1] Lee, S., Potamianos, A., and Narayanan, S. "Analysis of children's speech: duration, pitch and formants", *EuroSpeech '97: 5th European Conf. on Speech Comm. and Technology on Rhodes, Greece, 1997*, p 473-476.
- [2] Robb, M.P., Chen, Y., Gilbert, H.R. "Developmental aspects of formant frequency and bandwidth in infants and toddlers", *Folia Phon. Log.*, Vol. 49, 1997, p 88-95.
- [3] Zajdó, K. and Stoel-Gammon, C. "The acquisition of vowels in Hungarian: developmental data", *Proc. of the 15th Int. Congress of Phonetic Sciences in Barcelona*, Vol. 3., 2003, p 2229-2232.
- [4] Van der Stelt, J.M., Wempe, A.G., and Pols, L.C.W. "Progression in vowel production: comparing deaf and hearing children", *Proc. of the Inst. Phon. Sci. Univ. Amsterdam*, Vol. 25, 2003, p 197-206.
- [5] Szende, T. "Hungarian", *Handbook of the Int. Phon. Assoc.*, p 104-107. Int. Phonetic Assoc., Cambridge, UK, 1999.
- [6] Gussenhoven, C. "Dutch", *Handbook of the Int. Phon. Assoc.: a guide to the use of the Int. Phonetic Alphabet*, p 74-77. Int. Phonetic Assoc., Cambridge, UK, 1999.
- [7] Zajdó, K. "The acquisition of vowels by Hungarian-speaking children aged two to four years: a cross-sectional study", *Unpublished doctoral dissertation, Univ. of Washington (Seattle, USA)*, 2002.
- [8] Van der Stelt, J.M., Zajdó, K., and Wempe, T.G. "Exploring the acoustic vowel space in two-year-old children: results from Dutch and Hungarian", *Speech Communication*, in press.
- [9] Palethorpe, S., Wales, R., Clark, J.E., and Senserrick, T. "Vowel classification in children", *J. Acoust. Soc. Amer.*, Vol. 100, 1996, p 3843-3851.
- [10] Wempe, A.G. "F₀-related formant measurements", *Proc. of the Inst. Phon. Sci. Univ. Amsterdam*, Vol. 24, 2001, p 167-187.
- [11] Vallabha, G.K. and Tuller, B. "Systematic errors in the formant analysis of steady-state vowels", *Speech Comm.*, Vol. 38, 2002, p 141-160.
- [12] Wempe, A.G. and Boersma, P. "The interactive design of an F₀-related spectral analyser", *Proc. of the 15th Int. Congr. Phonetic Sci. in Barcelona*, Vol. 1., 2003, p 343-346.
- [13] Boersma, P. & Weenink, D. "PRAAT: a system for doing phonetics by computer", *Inst. of Phon. Sci. Univ. of Amsterdam*, 1996, Report number 132.
- [14] Serkhane, J., Schwartz, J.L., and Bessières, P. "Simulating vocal imitation in infants, using a growth articulatory model and speech robotics", *Proc. of the 15th Int. Congr. Phonetic Sci. in Barcelona*, Vol. 3., 2003, p 2241-2244.
- [15] Serkhane, J., Schwartz, J.L., Boë, L.J., Davis, B., and Matyear, C. "Motor specifications of a baby robot via the analysis of infants' vocalizations", *Proc. of Int. Conf. on Spoken Lang. Processing in Denver*, 2002, p 45-48.
- [16] Pols, L.C.W. "Spectral analysis and identification of Dutch vowels in monosyllabic words", *Ph.D. dissertation, Free University of Amsterdam*, 1977.