Assessing the significance of Tallal’s transform

Bjursäter, U.

journal: Proceedings of Fonetik, TMH-QPSR
volume: 44
number: 1
year: 2002
pages: 141-144

http://www.speech.kth.se/qpsr
Assessing the significance of Tallal’s transform

Ulla Bjursäter, Eeva Koponen, Fransisco Lacerda, Ulla Sundberg
Department of Linguistics, Stockholm University

Abstract
The perceptual significance of enhancing amplitude contrasts at the onset of formant transitions in CV-syllables and of reducing the “speaking” tempo was studied with a group of normally developing school children. Natural and synthetic speech stimuli were used in the perception experiments. A total of 83 children, second and third graders, were tested on their ability to discriminate between CV-syllables presented in pairs. The results indicate that the children’s discrimination performance resisted acoustic manipulations of both the natural and synthetic stimuli. Neither spectral nor timing manipulations rendered significant differences in discrimination results.

Introduction
It has been earlier reported (Merzenich, Jenkins, Johnston, Schreiner, Miller & Tallal, 1996; Tallal, Miller, Bedi, Byma, Wang, Nagarajan, Schreiner, Jenkins & Merzenich, 1996), that training methods using temporally and spectrally modified speech can help Specific Language Impaired (SLI) children. While the results obtained by the training program proposed by Tallal and her associates suggest that the method may be beneficial for SLI children, a number of questions concerning the methods rationale nevertheless have to be raised. From a phonetic perspective, for instance, the positive impact of the slowed down formant transitions is not readily compatible with traditional phonetic observations. Indeed, expanding the time scale of a CV-syllable’s formant transitions may be associated with a degradation of the phonetic cues for manner of articulation. This is for instance, particularly evident in the loss of stop-consonant characteristics of a [ba] syllable, which tends to be perceived as [wa] when the duration of the transitions is extended (Lacerda & Lindblom, 1998). Also the impact of the spectral amplitude manipulation may be difficult to integrate in a phonetic perspective. Under ecologically relevant listening conditions, speech sounds are often affected by room acoustics that, although to some extent affecting the detailed spectral amplitude specifications, do not seem to have appreciable impact in their phonetic representation. Nevertheless Tallal’s training with transformed speech has reportedly led to improvements in the language performance of language-impaired children, making it necessary to reexamine the possible role of such temporal and spectral transformations.

The overall aim of this project is therefore to investigate to what extent a training program based on manipulated speech can lead to improved phonological awareness in children with Specific Language Impairment (SLI). As a part of this project a series of discrimination tests using both transformed and non-transformed stimuli, were performed by normally developing children in order to assess the perceptual impact of the manipulations.

Method

Subjects
The subjects participating in the perception tests were a group of 83 normally developed second- and third-grade children from four different classes at two Swedish public schools in the Stockholm area.

Stimuli
The stimuli were natural utterances, [ba] and [da], produced by an adult male speaker of Swedish, and synthetic [ba] and [da] stimuli. The synthetic stimuli were 4-formant utterances generated by a parallel speech synthesizer. The formant transitions were exponential between the static vowel and the stop-consonant loci with durations defined as the time interval to reach 90% of the total transition excursion. All the four formant transitions had the same duration. The voice source was created by a pulse train in with F0 decaying linearly from 135 Hz to 120 Hz.
throughout the stimulus. The vowel’s formants were set at 592 Hz, 1070 Hz, 2400 Hz and 3300 Hz. The F-pattern of the loci were defined as F1=200 Hz, F2=700 Hz for [b] and 1600 [d], while F3 and F4 were identical to the vowel’s F3 and F4.

In addition to the original versions, both the natural and synthetic stimuli were manipulated temporally and/or spectrally as follows. The natural stimuli were manipulated in their transition durations to obtain additional series, with doubled transition durations for [ba] and [da]. The synthetic stimuli consisted of series of [ba] and [da], created with transition durations of 20 ms and 60 ms. Both the natural and synthetic stimuli were amplified according to Tallal’s transform – an increase by 20 dB in the consonantal part (Merzenich et al., 1996). The fourth series of stimuli were created by a combination of temporal and spectral manipulation; the duration of the consonantal segment of both the natural and synthetic stimuli were doubled/elongated to 60 ms in relation to the vowel duration, in combination with amplification at the onset of the formants transitions according to Tallal’s transform (table 1). To increase the difficulty of the task, all stimuli were presented with a background of pink noise with S/N of 3 dB.

Table 1. Type of discrimination pairs used in this experiment.

<table>
<thead>
<tr>
<th>Contrasts:</th>
<th>ba-ba, ba-da, da-ba, da-da</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural stimuli</td>
<td>original</td>
</tr>
<tr>
<td></td>
<td>doubled</td>
</tr>
<tr>
<td></td>
<td>amplified</td>
</tr>
<tr>
<td></td>
<td>doubled+ amplified</td>
</tr>
<tr>
<td>Synthetic stimuli</td>
<td>original</td>
</tr>
<tr>
<td></td>
<td>elongated</td>
</tr>
<tr>
<td></td>
<td>amplified</td>
</tr>
<tr>
<td></td>
<td>elongated+ amplified</td>
</tr>
</tbody>
</table>

Procedure

Five students of Linguistics validated the stimuli to be used in the discrimination tests.

A group of 83 second- and third-grade school children was asked to classify the (2 x 32) pairs of stimuli as “same” or “different”. The stimuli were presented in random order. The subjects listened to the stimuli through headphones, in some cases headphones with Active Noise Reduction (ANR) were used. The children were given a short practice period consisting of 8 pairs of test stimuli. The experimenter presented each pair at the request of the child as a means of individualizing the pace of the test. The children could listen to each stimulus pair as often as they wanted. Two children performed only half of the test.

Results

A test of between-subjects effects indicates that the children treat each stimulus pair in a significantly different way, F(3,61)=6.884, p=0.001. The pooled results from the natural and synthetic stimuli show an overall tendency towards correct responses for the [ba-ba] and the [da-da] pairs by all four classes. The [da-ba] stimuli render slightly less correct answers, especially by the younger second-grade children. As for the [ba-da] combination, the results show a marked decrease of correct answers, below 40%, though non-significant.

The results indicate an asymmetry regarding presentation order, with somewhat higher discrimination scores for initial [da] vs initial [ba]. No overall significant difference was found between the discrimination performances with synthetic or natural stimuli (see Fig. 1 and 2), although the accuracy for the natural [ba-ba] pairs was clearly higher and had lower variance than for the corresponding synthetic sequence.

A closer look at the responses to the natural stimuli (Fig. 1) shows a difference in discrimination between the [ba-ba] and the [ba-da] stimuli, with higher discrimination scores for the first pair relative the second pair. There is little difference between the [da-ba] and the [da-da] discrimination for the third-grade children.
while the second-grade children show decreased discrimination ability for the first pair relative to the second pair. It is noteworthy that performance of the older age group tends to increase in the [da-ba] discrimination in comparison to the [ba-da] sequence.

Turning to the synthetic stimuli (Fig. 2) the [ba-da] sequence seems to be harder to discriminate than all the other sequences. No age effect is detected for the two leftmost sequences and only minor variations for the rightmost sequences.

Results from discrimination of stimuli with or without time manipulation (Fig. 3) reveal no significant difference in any aspect, only a slight overall advantage for the responses to the natural stimuli with doubled transition time.

In addition, the manipulation of the spectral amplitudes has apparently not affected the children’s discrimination scores (Fig. 4). The stimuli manipulated according to Tallal’s transform showed only a slight, non-significant increase in discrimination scores F(1,31)=3.716, p=0.063.

Discussion and conclusions

Overall, the results suggest that the natural sequence [ba-ba] is the easiest pair of stimuli to discriminate and that the natural [ba-da] is the hardest. Spectral and time manipulations don’t seem to affect discrimination at all. The results from this study partly corroborate Lacerda’s (2001), who found that natural CV-stimuli were nearly unaffected by transition speed manipulation. In contrast to the present study, Lacerda’s results suggest an increased sensitivity to synthetic stimuli.

A potential reason for the lack of clear-cut discrimination differences between manipulated and non-manipulated stimuli might be that linguistically normally developing children are rather insensitive to acoustic manipulations. The present results also imply the need to perform perception studies with context-embedded stimuli giving the children an opportunity to process the stimuli a more relevant ecological setting.

Admittedly, children with documented language impairments (SLI) may be more sensitive to acoustic manipulation of CV-syllables than the children in this present study proved to be. This issue is currently under investigation (Segnestam, Johnsson and Lacerda, in preparation).
Acknowledgements

Swedish Research Council for Humanities (HSFR) supports this research.

References


