ABSTRACT

The present study attempts to investigate the significance of the release burst and the formant transitions in the perception of Russian voiceless plosives by native listeners. The method involved deleting of releases in some consonants, it resulted in worsening recognition of these sounds — 34% for initial plosives, 60% and 70% for intervocalic and final plosives respectively. Thus, it is stated that release segments carry important information bearing on the place of articulation of Russian stops. The results of the study are in agreement with those obtained on the material of English and Hungarian stops and defy the prevailing significance of CV-transitions in voiceless plosives recognition.

INTRODUCTION

Although the history of experimental studies stimulated by the development of the Visible Speech Sound Spectrograph dates back to the 40s, it is hardly possible to say that the distinctive features of phonemes are fully investigated. It is not surprising since the acoustic features of sounds in fluent speech can vary dramatically due to the context, the speaker's peculiarities, the mode of articulation, etc. Besides, the speech signal is highly redundant and possesses a great variety of distinctive features.

Quite a number of works have been dedicated to the distinctive features of plosives. It is a stated fact that information about the place of articulation of a stop can be found in the formant transition of adjacent vowels as well as in the stop burst. The relative significance of release and transition in the stop identification, however, is to be further investigated.

This problem is relevant for the systems of automatic speech recognition and high-quality speech synthesis. Another important problem concerns the search for invariant (i.e. independent of a context) features of the place of articulation.

METHOD

The model of a voiceless plosive is used according to which the four segments can bear information about the place of articulation of an intervocalic plosive: 1) the segment of VC- or final transition; 2) the closure; 3) the release after an abrupt closure break; 4) the segment of a CV- or initial transition. The VOT was taken...
for a release end. In a more detailed model a release is further divided into: a) a starting impulse, b) a transition (cf./I/I). A certain amount of residual noise may add to the voiced beginning of a following vowel/T/I/. Since these peculiarities of the Russian language are not phonemically relevant they are not considered here.

To investigate the relative significance of the transition and release cues for prevocalic, intervocalic and postvocalic plosives some meaningless successions of the CV/CVCV-type were chosen, the consonants being the same in the one case (e.g. /papap/, /totot/, etc) and different in the other (e.g. /gatak/, /kotop/, etc). The vowel has been taken out of the set {a,o,ü,î}, the second syllable of each non-word was stressed.

Non-words were tape recorded by two female speakers. The instruction to the speakers was to utter the stimuli distinctly without changing the quality of vowels. The interval between the stimuli was 5 seconds.

The tape rings were made of the original recordings that underwent segmentation by means of the low-noise electronic separator described elsewhere (/F/).

A release for one of the plosives (initial, central or final) has been deleted in every non-word (CVCV-type) as well as in the /papap/, /totot/, etc. The fragments of non-words with a release deleted were used as test stimuli. Thus the relative significance of the CV-transition, CV-transitions and VCV-transition place identification was studied.

The deletion procedure has been controlled several times with the aid of oscilloscope. The test stimuli have been recorded on the second tape-reproducer. Each fragment was recorded three times on a tape ring. The trials of the identical stimulus were separated by pauses marked by the monophone markers. The presentation rate of the test stimuli which depended on the ring speed and tape speed was about 1.8 seconds.

Non-words with release deleted were mixed with undamaged non-words and thus presented to ten listeners (students, laboratory assistants, etc.) without hearing loss. Most of the listeners were experienced in listening to articulatory tests and syntheses speed patterns. The signal was fed into the head-phones in a quiet room. The listeners could adjust the volume in his head-phones.

The instruction given to the listeners was as follows: "Listen to the triads of the CV/CVCV-type, where C is any of the plosives — /p/, /p'/, /t/, /t'/, /k/, /k'/, etc. Each non-word is repeated 3 times. After listening to a triad you choose the second syllable of the CV/CVCV-type, where C is any of the plosives — /p/, /p'/, /t/, /t'/, /k/, /k'/, etc. In this case you are to write down in ordinary letters ("papap", for example). If you detect a distorted (damaged) consonant please underline it as shown below: /gatak/ or /gutuk/ or /gatak/.

Notice: the soft /k/ may occur in a final position that are not typical for the Russian language.

The instruction was presented by the experimenter orally and then its printed text was distributed among the listeners along with 120 sets of the CV/CVCV-type test stimuli. A sample of such a matrix is given in the table below. The right and wrong judgments for the /p/, /t/, /k/ stimuli (with release deleted) preceding /a/ vowel in the first hearing session (speaker L) are presented in the table.

<table>
<thead>
<tr>
<th>Table</th>
<th>perceived</th>
<th>/p/</th>
<th>/t/</th>
<th>/k/</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The Judgements are summerized for ten speakers in every matrix. The "-"-sign stands for refusal. The left matrix corresponds to the initial position of stops, the second position to the central and the final positions, respectively.

It can be seen that in two presentations of a damaged initial /p/ (non-words with release deleted) preceded by /a/ vowels the right judgments of transition and release in stop recognition is omitted here.

DISCUSSION

The results of the present study agree with the results of other authors. Thus in /I/ the relative significance of transition and release is not as important as the place of articulation of damaged plosives.

It is not quite correct to mention the guessing level since the listeners judgments in this case were not gueesing. Most of these were judgments of /p/-type. The explanation might be that the second type is remembered and is not so important for listeners may misjudge any voiceless plosive as /p/ when the release is absent.
ion of the place of English voiceless stops preceding /i/, /a/, /u/ vowels has been studied. It has been found that the voiced segment of the initial transition is neither a sufficient nor necessary cue for the place identification in the initial position, and it is the release that accounts for a correct consonant identification.

Similar conclusions were made for the Hungarian voiceless stops in VCV-syllables /I6/. The release was judged to be the most informative among the segments of closure, release and transitions due to /I/, where the distribution of stop cues has been studied.

A few studies of Russian voiceless stops have been described in /I0–I2/. It was found that in the final position of a stop the release is more important than the transition (the method consisted in transplanting bursts from one context into another) /I2/. The significance of bursts and final transitions in final positions of stops has been studied in detail in /I0/ on the CVc-syllables. It was shown that in most cases an isolated burst of final stops was sufficient for the identification of place. Final transitions can also carry information about the place of articulation.

In the article /I1/, however, VC-transitions were not found significant in the perception of plosives. (A more detailed comparative analysis is required to explain the disparity between /I/ and the present study).

The question remains – which of the two transitions, CV or VC, is more informative in the place identification of the plosives? The results of the present study suggest that VC-transitions are more informative than CV-transitions. Similar conclusions can be found in /I7/, where the role of CV- and VC-transitions in the place of articulation of English stops and fricatives in syllables with neutral /a/ in natural speech was determined. VC-transitions proved to be more informative than CV-transitions especially for voiceless stops: VC-transitions accounted for 92% of their intelligibility whereas CV-transitions accounted for only 52%. The corresponding values for voiced plosives were 92% and 71%. According to the authors the reason for VC-transitions being more informative than CV-transitions is their better physical manifestation. The conclusion has been substantiated by inverse listening data in particular.

SUMMARY

The results of the present study provide further evidence of the relative significance of transition and release as perceptual cues for Russian voiceless plosives. The regularities which have been found may prove useful in developing more efficient automatic recognition systems and high-quality speech synthesis.

REFERENCES
