

On Perceiving Certain Voiceless Unaspirated Stops

Leigh Lisker

Haskins Laboratories
270 Crown Street
New Haven CN
USA

ABSTRACT

A common type of stop is voiceless and unaspirated, sometimes contrasting with a voiced one. In English this is true in certain contexts, but utterance-initially the two types vary freely, both heard as voiced by phonetically naïve native speakers, though linguists sometimes describe the first as "devoiced" and/or the second as "prevoiced." Moreover, wherever a voiceless inaspirate context is edited so the stop release + opening transition come to be utterance-initial, the stop is perceived as voiced. In Russian stops are either voiced or voiceless unaspirated, the two being in contrast initially and medially. Unlike the case in English, whose voiceless inaspirates have context-determined phonological and perceptual status, the voiceless inaspirates of Russian are everywhere reported as voiceless, a perception shared by speakers of Russian, English, and French.

INTRODUCTION

A commonly observed difference between languages involves their exploitation of differences in the articulation of stop consonants. Thus the stops of Russian fall into two sets, one voiced and the other voiceless unaspirated, which contrast with one another initially and medially. The stops of English, on the other hand, while they also include voiced and voiceless inaspirates, do not use these contrastively in utterance-initial position; instead, in that context the two types vary freely, both of them being opposed to voiceless aspirated stops. Further, the two languages resemble each other in that each includes clusters made up of [s]+voiceless unaspirated stop.¹ To be sure, when the stop of an English [s]-cluster is presented in the absence of the [s]-noise, it is heard as voiced by most English-speaking listeners, quite like the "devoiced" variant of the initial voiced stop, though phonologists of English more commonly draw attention to the absence of aspiration than the absence of closure voicing. In the case of the Russian voiceless inaspirates, while it may be assumed that those found initially before vowels will be heard as voiceless by native speakers of Russian, it is not known how they would be perceived by

¹ Russian also permits initial clusters made up of [z]+voiced stop.

phonetically naïve English speakers, nor is it known how either English or Russian speakers would judge the stops extracted from the Russian [s]-voiceless stop clusters and presented in initial position. One aim of the present study is to look into these questions.

EXPERIMENT ONE

Let us consider three English sentences which are alike in that all can be said to include a voiceless unaspirated labial stop. These are (1) Did he win this bout? (2) Did you fix the spout? (3) Didn't you drop out? Of the included labial stops, the first is often represented as [b], while the second and third are [p]. Phonologically they differ, the first being /b/, the third being /p/, while the second is phonologically not unequivocally one or the other. When the three sentences are deprived, by waveform editing, of the intervals preceding their labial closures, all are identified as instances of the word bout. In the cases of Sentences 1 and 2 this is no surprise, given that in the first this is the intended word, while in the second there is plenty of evidence that the word spout should yield bout upon deletion of the [s]-noise. Sentence 3 provides the surprise, for there is nothing in the literature to suggest that the post-closure release + transition of [p] will give rise to a "b" percept. To be sure, we might explain the "b" percept by the absence of any post-release aspiration, though we might have expected the [p] to be released with a burst and transition having the intensity and F0 and F1 believed to mark voiceless unaspirated as well as voiceless aspirated stops (Fischer-Jørgensen, 1968; Erickson, 1975; Ohde, 1984).

The finding that the three labial stops in our sample sentences were all heard as "b" by no means implies anything as to their phonetic discriminability. To test for this property three tests were carried in which the post-labial closure intervals were compared in pairwise fashion. In each test ten English speakers had the task of deciding from which of two source sentences each syllable was derived. The response data, when subjected to a repeated measures analysis of variance (ANOVA) showed that listeners were unable to distinguish among the stimuli derived from different source sentences. Thus the main effect of Source Sentence

was not significant, $F(1,9) < 2.1$; $p > .15$). Consistent with this result was the absence of any statistically significant difference in the timing of voice onset following the stop release.

EXPERIMENT TWO

A native speaker of Russian recorded three tokens of each of the three Russian trochaic words [dɔl'kə] [tɔl'kə] [stɔl'kə]. Two additional stimuli were generated by removing the closure voicing of the [d] and the [s]-noise of the [st] cluster. Four repetitions of each of the three tokens of the five stimulus types were presented in randomized order to three groups of listeners: 8 Russian speakers, 8 English speakers, and 8 French speakers. All these listeners had previous experience as subjects in experiments calling for auditory judgments of speech and speech-like stimuli, but none had a speaking command of a non-native language. In the present study they were asked to label the initial stops in so far as possible in terms of the categories of their native language.

As was expected, the Russian listeners agreed 100% in reporting Russian [d] as "d" and Russian [t] as "t." The artificially devoiced [d] were labeled "d" 89% and "t" 11%, while the [st] minus [s]-noise was judged 100% to be "t."

Not at all unexpectedly the English-speaking listeners also reported Russian [d] entirely as "d," while Russian [t] yielded 88% "t" and 12% "d" responses. The devoiced [d] evoked 84% "d" responses and 16% "t"s, while the stimuli derived from the cluster were reported 12% as "d" and 88% as "t."

The French listeners also judged the Russian voiced stop entirely as "d," while Russian [t] elicited the same proportion of "t"s and "d"s as were derived from the American listeners. The devoiced [d]s were judged 96% "d" and the stimuli from the cluster were reported 96% as "t."

EXPERIMENT THREE

As we have seen, English speakers, unlike Russian and French speakers, quite regularly produce variants of initial voiced stops with voiceless closures. A new set of tokens of the Russian words tested above was recorded by the author, this one including [dɔl'kə] tokens, i.e. tokens whose initial stops were phonetically voiceless unaspirated apical stops like those commonly found in a word such as *during*. These stimuli, which we might call "pseudo-Russian," were submitted to our three sets of listeners for judgment. The distributions of "d" and "t" responses elicited by those stimuli intended to match the Russian speakers productions were not very different from those evoked by the latter, a major difference being that the [dɔl'kə] stimuli (i.e. the naturally voiceless "d"s, were heard largely (88%) as "t" by the Russians, but mostly as "d" (88%) by the Americans, while the French divided their responses fairly evenly (52-48%) between "d" and "t." On the other hand, the artificially devoiced

[d]s were heard by the Russians as 46% "d" and 54% "t," and were generally reported more often as voiced than were the naturally "devoiced" [d] stimuli. In the case of the stimuli derived from the words with initial clusters, however, the pseudo-Russian items elicited about as many "t" responses as did those tokens recorded from the Russian speaker. Comparison of the responses evoked by the naturally and artificially devoiced [d] tokens revealed that for each group of subjects the first set yielded fewer "d" judgments than did the second, though the difference was least in the case of the Americans, for whom both types of [d]s were heard mostly as voiced, quite in line with what we should expect of English-speaking listeners, in whose language the "voicing" of initial inaspirates does not depend on the presence of closure voicing.

DISCUSSION

A comparison of the data reported in the two above described experiments shows that the voiceless unaspirated stops of English and Russian elicit rather different percepts. The question might be posed whether speakers of the two languages simply view this class of stop differently, e.g. that for Americans it is enough that an initial stop not be aspirated for it to be "voiced," while for Russians the absence of voicing during closure, provided it is naturally achieved, is enough for it to be "voiceless." A closer look at our results shows that this cannot be the whole story, since from Experiment Two it is clear that English speakers do not require aspiration in order to report hearing a voiceless stop, unless we are willing to allow that Russian voiceless stops, while not as aspirated as English initial voiceless stops, are nevertheless released with a modicum of aspiration. Moreover, Russian speakers, although they regularly produce their "voiced stops" with closure voicing, do not absolutely require that feature in order to report a stop as voiced.

In order to check whether the Russian and the English voiceless inaspirates manifested any clear acoustic difference, measurements of voice onset timing (VOT) were carried out on twenty tokens each of the Russian [p], the English [b] of the first sentence of Experiment One, the English [p] of the second sentence, and the English [p] of the third sentence. These yielded mean VOT values of +17 ms for the Russian [p] as against values of +13ms for English [b] and +10ms for each of the [p]s of English sentences two and three. The difference between the means of Russian [p] and English [b], when subjected to unpaired t-test, yielded a t-value of 2.02, with $p = .05$, while mean differences among the three English voiceless inaspirates, when similarly tested, failed to demonstrate statistical significance. In other words, it would appear that Russian [p], while not produced with a VOT value anywhere nearly as great as mean values reported for

English /p/ in the literature,² has a value great enough for it to pass muster as English /p/.

REFERENCES

- [1] D. Erickson. "Phonetic implications for an historical account of tonogenesis in Thai," in *Studies in Tai Linguistics in Honor of William J. Gedney*, J. H. Harris & J. R. Chamberlin, Eds. pp. 100-111. Bangkok: Central Institute of English Language Office of State Universities. 1975.
- [2] E. Fischer-Jørgensen, "Les occlusives françaises et danoises d'un sujet bilingue," *Word*, vol. 24, 112-153, 1968.
- [3] L. Lisker and A. S. Abramson. "A cross-language study of voicing in initial stops: Acoustical measurements," *Word*, vol. 20, pp. 384-422, 1964.
- [4] R. N. Ohde, "Fundamental frequency as an acoustic correlate of stop consonant voicing," *Journal of the Acoustical Society of America*, vol. 75, pp. 224-230, 1984.

ACKNOWLEDGMENTS

The work reported here was supported by the National Institute of Health NIDCD under grant DC-02717 to the Haskins Laboratories

² Thus in Lisker and Abramson 1964 a mean VOT value of +58ms was reported for a total of 102 isolated words with initial /p/ recorded from four American English speakers.