Phonetic Explanations for the Devoicing of High Vowels
Author(s): Hector R. Javkin
Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society (1979), pp. 413-418

Phonetic Explanations for the Devoicing of High Vowels[1]

Hector R. Javkin
U.C.L.A.

Introduction

Mathematical models of phonetic processes can be an important tool in phonetic research. They can serve as a test of the current state of knowledge. For example, the success of speech synthesis by rule serves as a stringent test of what we know about segmental spectra, durational effects, co-articulation, etc. Models can also be useful in filling in gaps in those areas in which empirical data is difficult to obtain. The intricate mechanical forces occurring in vocal cord vibration, for example, are usually not accessible during speech. Even when experimental data from real speakers is available, models can suggest the areas in which that data would be useful. For questions in which all the relevant variables can be taken into account, models can actually serve as explanations.

It has been observed by Greenberg (1966) and others and convincingly demonstrated by Jaeger (1978) that high vowels such as /i/ and /u/ are much more likely to occur voiceless than are low vowels such as /ae/ and /ɔ/. In this paper, I will suggest that some of the relevant variables cannot yet be fully determined, but that some of the data, and some of the work with models of speech production, suggest what the explanation probably is.

An explanation was proposed by Ohala (1975) on the basis of a model of speech aerodynamics. High vowels such as /i/ and /u/ have a somewhat closer constriction than low vowels. The output of the model suggested that high vowels will therefore be associated with higher oral pressures and higher air velocities through the constriction than low vowels. Both of these effects should lead to vowel devoicing.

Assuming that sub-glottal pressure is equal for different vowels, higher oral pressure will result in a smaller pressure difference across the glottis. Since pressure differences are what force air through the glottis and provide the power for vocal fold vibration, this reduction in the pressure difference could reduce the chances of a vowel being voiced.

The greater air velocity through the constriction should result in greater turbulence on those occasions when vowels are devoiced. The turbulence will therefore be more noticeable. This greater perceptibility should make it more likely that listeners will "pick up" the devoicing and start to make it part of the linguistic code.

A further possibility which I investigated is that the different vocal tract shapes for high and low vowels cause a greater amplitude of fricative noise, not as a result of greater air velocity, but as a result of the effects of vocal tract resonances and losses. In other words, high vowels, such as /i/ and /u/ which have a place of maximum constriction further front in the mouth than low vowels, would transmit more of whatever fricative noise was produced. This is suggested by a model constructed by Stevens (1971) in which he found that fricatives closer to the front of the mouth would have greater transmission of fricative noise. A further consideration is the fact that the damping of noise that occurs as sound travels through the vocal tract is more or less proportional to the distance the sound wave travels in the tract itself. To the extent that high vowels develop fricative noise at their constriction maxima, the transmission of the frication produced in front vowels might be greater than the transmission of the frication of vowels further back in the vocal tract.

Discussion

There are reasons to believe that the oral air pressure for high vowels cannot be the reason these vowels devoice. First, the difference in pressure that Ohala's model provides for close versus open vowels is only 1 cm. of water. It is not clear that such a difference would impede voicing. Second, some exploratory measurements by Ohala (1973) yield a pressure difference of approximately one-half a centimeter of water. Third, it might be expected that an oral pressure that could impede voicing would lead to a lowering of pitch. Once again, high oral pressure, by reducing the difference in pressure across the glottis, will tend to reduce airflow through the glottis. Other things being equal, a reduction in airflow leads to a reduction in pitch. We would therefore expect that high vowels would have an intrinsically lower pitch than low vowels. They do not. In fact, high vowels have an intrinsic pitch which may be as much as 10 Hz higher than low vowels. It was therefore clear that the predictions regarding pressure should be tested, by taking direct measurements from speakers, using an oral pressure tube.

Air pressure measurements

Although measurements were attempted for three speakers, only one of these, a female speaker of American English, provided reliable data. The measurements were made at the Phonology Laboratory of the University of California, Berkeley, using a Statham PM15-ECT pressure transducer, whose output was recorded on an FM channel of a Vetter tape recorder. Calibration signals at 0 and 4 cm. of water, measured with a water manometer, were recorded on the same channel just prior to the measurement tokens. The audio signal was recorded on the direct channel of the Vetter tape recorder. The output of the FM tape recorder was continuously monitored on an oscilloscope during the reading of the tokens. The measurements were stopped and the word list re-started five tokens back on three occasions when the oscilloscope signal suggested that the tube was clogged. The recordings were measured, using the calibration signal as a reference, on a Linc-8 computer, which sampled the pressure data at 500 Hz and integrated the signal to approximately 125 Hz. The words used were: *lead* (v.), *lewd*, *laud*, *lad*; *seat*, *suit*, *sought*, *sat*; *beat*, *boot*, *bought*, and *bat*, read within the frame sentence: 'I will say ____ again' following the token number. Ten tokens of each word were used, with all words in a pseudo-random order which was modified so that each token occurred five times in the first 60 utterances. This was done so that some of the data could be salvaged in case the experiment had to be terminated early.

Results

Table 1
Average oral pressures for vowels /i/, /u/, /ɔ/, /a͡e/ (N=10)

	i	u	ɔ	a͡e	
l_d	-.01	.61	.57	.19	
s_t	.69	.88	.91	.64	(in cm. H_2O)
b_t	.51	.68	.55	.47	

The results are shown in Table 1. The negative values for the word *lead* are somewhat difficult to understand. The holes in the pressure tube were transverse to the airstream, and it is possible that a small Bernoulli effect caused the negative pressure. The measurements as a whole are within the range of other studies (e.g. Ohala 1973), but they are not consistent with the results of the aerodynamic model. High and low vowels show very similar pressure values, as can be seen in the averages for all environments, shown in Table 2.

Table 2
Oral pressures for vowels /i/, /u/, /ɔ/, /a͡e/
averaged over all environments (N=30)

i	.40	
u	.72	(in cm. H_2O)
ɔ	.67	
ae	.43	

The measurements, taken together with some observations made earlier, suggest that differences in oral air pressure cannot be the cause of the preferential devoicing of high vowels.

Vocal tract transmission

In order to test the hypothesis that the vocal tract configuration for high vowels makes these vowels noisier and more noticeable when they are voiceless, a computer model for synthesizing speech from vocal tract shapes was used. The program (VOCALTS) was developed by Jim Wright on the basis of specifications by Klatt (1971). The description here is a highly abbreviated summary of the description in Wright (1976). The program models the vocal tract by simulating a series of cylindrical sections of uniform length. Values representing a traveling wave can be entered. The program traces the progress of the wave through each cylindrical section. In the program, as in the real world, partial reflections occur at boundaries between sections of different cross-sectional area. Damping of the wave due to friction and cavity wall movement, which were estimated in Fant (1960), are roughly simulated by assuming that the damping is proportional to the distance that the wave travels. The boundary conditions at the lips are modelled by adding a very large section and those at the glottis by adding a very small section. In addition, the model assumes that the wave from the glottis section is completely damped by the pulmonic system, so that there is no reverse wave reflected from the glottis section. The program differs from the description given in Wright 1976 in that the source function representing voicing at the glottis is removed. A random-number generating algorithm is added in order to simulate broad-band fricative noise. The fricative noise was added at the place of maximum constriction for each of the vowels examined. For vowels in which the maximum constriction was several sections long, noise was added at the forwardmost section. The vocal tract shapes used were those for the Russian vowels /i/, /a/ and /u/, taken from Fant (1960). Restrictions in the program required that the sections be .85 cm. long. The values given by Fant were quantized into these sections. This resulted in the vowels /i/ and /a/ each having 20 sections totalling 17 cm., which represents a lengthening of half a centimeter for the /i/. The frication source for /u/ was placed at the place of maximum constriction at the lips. The output of the program was analyzed with a Fast Fourier Transform.

Of the three vowels, /i/ had the highest peak amplitude, which occurred at 5100 Hz. /u/ had a peak amplitude about 2 dB lower, which occurred at 7250 Hz, and /a/ had a slightly lower peak energy, occurring at 3700 Hz. The summed amplitude values for the entire spectrum showed even smaller differences, with all of the vowels falling within 1 dB of each other.[2]

Conclusion

It seems, from the similarity in output amplitudes for the three vowels, that differences in vocal tract transmission do not result in differences in devoicing between high and low vowels.

The only viable explanation seems to be the second suggestion of Ohala (1975), that the higher air velocities and narrower constrictions for high vowels yield more frication noise at the source. Furthermore, the sharp bends in the airstream near the place of maximum constriction for high vowels, particularly /i/, should yield higher turbulence. This is the part of the explanation which had not yet been adequately tested. Frication noise can be calculated by determining the amount of turbulence generated by the complex shapes of the constrictions for vowels, but this has not yet been done. At the moment, the answer that the devoicing of high vowels is due to the noisiness of frication source has been arrived at by eliminating the other possibilities.

The fact that it is not the increased pressure of high vowels, but their noisiness, contains some phonological predictions. We might expect to find a number of languages in which high vowels are associated with the development of frication in surrounding consonants. Jaeger (1978) on the basis of the sample in the Stanford Phonology Archive, found a large number of languages in which such frication develops in the environment of high vowels versus a much smaller number of languages in which frication develops in the environment of low vowels. Jaeger also claimed, however, that a number of languages showed the <u>devoicing</u> of consonants in the environment of high vowels. If the devoicing associated with high vowels is a by-product of frication, it would be expected that there would be many more cases of consonants fricating in the environment of high vowels than there would be consonants devoicing in the same environment. In fact, none of the cases of devoicing cited by Jaeger actually represent devoicing. There are two languages which appear to have devoicing, Kunimaipa and Totonac. The rules given by Jaeger for these languages are the following.

Kunimaipa: l → dƚ/high vowels

Totonac: l → ƚ/i___

Although /ƚ/ is voiceless in some transcriptions, the Archive sources for these two languages do not specify these sounds as voiceless. Aschmann (1946) describes the /ƚ/ in Totonaco as fricative. In Kunimaipa, the Archive notes that the /dƚ/ cluster sometimes becomes voiceless phrase finally, which is very different from saying it becomes devoiced in the environment of front vowels. Furthermore, the source given by the Archive for this (Pence 1966:61) does not mention this devoicing. This is not to say that high vowels could not be associated with the devoicing of consonants, only that such devoicing must be expected to occur much less frequently than frication.

This is yet another case in which the phonetic and phonological facts fit extremely well. The devoicing of high vowels is a result of the fact that these easily become fricatives, not because of a direct effect on vocal fold vibration. In conclusion, perhaps the point needs to be made once again that the kind of models discussed here, intricate though they sometimes are, will provide the kinds of answers to many of the questions that phonologists ask.

Footnotes

1. I would like to thank Johanna Nichols, John Ohala, Sandra Pinkerton, Carol Riordan, and James Wright for very useful advice on this study. Jim Wright deserves additional thanks for making his programs available. I would also like to thank the members of the Phonetics Laboratory at U.C.L.A. (Work supported by NIH and NSF.)

2. Lehiste and Peterson (1959) found that, among voiced vowels, low vowels show the greatest amplitude. This suggests that a frication noise at the glottis would fail to provide greater amplitudes for high vowels. Nevertheless, I tested the model with frication at the glottis. The results once again failed to support the hypothesis that vocal tract transmission is responsible for the preferential devoicing of high vowels.

References

Aschmann, H. P. (1946) "Totonaco phonemes," IJAL 12:34-43.

Pence, A. (1966) "Kunimaipa phonology: hierarchical levels," Pacific Linguistics, Series A 7.49-67.

Fant, G. (1960) Acoustic theory of speech production, The Hague: Mouton.

Greenberg, J. (1966) "Synchronic and diachronic universals in phonology," Lg. 42:508-17.

Jaeger, J. (1978) "Speech aerodynamics and phonological universals," Proc. Berkeley Ling. Soc. 4:311-29.

Klatt, D. (1971) Unpublished, untitled ms.

Lehiste, I., and G. Peterson (1959) "Vowel amplitude and phonemic stress in American English," JASA 31.4:428-435.

Millardet, G. (1911) "Insertions de consonnes en suedois moderne," Rev. Phon. 1:309-46.

Ohala, J. (1973) "Explanations for the intrinsic pitch of vowels," Monthly Internal Memorandum of the Phonology Laboratoratory, UC Berkeley.

_____ (1975) "A mathematical model of speech aerodynamics," in Speech Communication. Proceedings of the Speech Communication Seminar, Stockholm, Aug. 1-3, 1974. Vol. 2, pp. 65-72.

Stevens, K. N. (1971) "Airflow and turbulence noise for fricative and stop consonants: static considerations," JASA 50:1180-92.

Wright, J. T. (1976) "Vocalts: a computer model for synthesizing oral and nasalized vowels." Unpublished ms.