Identification of deleted plosives: the effect of adding noise or applying a time window: a reply to Ohde and Sharf

Pols, L.C.W.; Schouten, M.E.H.

Published in:
The Journal of the Acoustical Society of America

DOI:
10.1121/1.385312

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Identification of deleted plosives: The effect of adding noise or applying a time window (A reply to Ohde and Sharf)

L. C. W. Pols and M. E. H. Schouten

Institute for Perception TNO, Soesterberg, The Netherlands

(Received 5 June 1980; accepted for publication 8 September 1980)

A Letter in this Journal [Ohde, R. N. and Sharf, D. J., J. Acoust. Soc. Am. 69, 297-300 (1981)] deals with the identification of voiced and voiceless stops from vocalic transition + vowel stimuli edited from CV and VC syllables. The segmentation points were smoothed by applying a time window. With respect to an earlier study, in which straightforward segmentation without time window was applied, final plosives were still identified better than initial plosives. The present Letter shows that there is, however, one major difference between the with-window and the no-window condition, namely a substantial improvement in identification for initial voiced plosives. This improvement is similar to that found by the present authors in identification experiments with a noise burst adjacent to the segmentation point versus no-noise listening conditions.

Instead of a controversy reported by Ohde and Sharf, this Letter suggests a unifying interpretation.

PACS numbers: 43.70.Dn, 43.70.Ve, 43.66.Dc

INTRODUCTION

Several recent papers (Sharf and Hemeyer, 1972; Sharf and Beiter, 1974; LaRiviere et al., 1975; Ohde and Sharf, 1977; Pols and Schouten, 1978; Pols, 1979; Ohde and Sharf, 1980) deal with the identification of initial and/or final plosives in monosyllables from which the initial or final plosive burst + aspiration has been deleted. The stimuli for plosive identification, particularly identification of place of consonant articulation, then consist only of the consonant-vowel (CV) vocalic transition + vowel-segment for initial plosives, or of the vowel segment + vowel-consonant (VC) vocalic transition for final plosives. Ohde and Sharf (1977) and LaRiviere et al. (1975) also studied the vocalic transition alone. Pols and Schouten (1978) and Pols (1979) used various segmentation points.

The two discussion points we want to focus on in this Letter are:

1) Is there a substantial difference between initial and final plosive identification?

2) Does the deletion procedure itself introduce artifacts or lower the correct score, and if so are there ways of overcoming this?

In Pols and Schouten (1978) we, somewhat provocatively, claimed that there is no inherent difference in information content between CV and VC transitions. We contended that any such differences found (Sharf and Hemeyer, 1972; Ohde and Sharf, 1977) were due to the abrupt onset of the CV transitions after deletion of the initial burst + aspiration. As a result, deleted initial stop consonants were identified much worse than deleted final consonants. We demonstrated that for Dutch plosives the advantage of final over initial consonant identification could be reduced considerably by replacing the deleted consonants with noise, thus smoothing the abrupt onsets. As a result, identification of initial consonants improved, whereas that of final consonants remained unchanged.

In their recent Letter to the Editor, Ohde and Sharf (1980) set out to counter our claim that there is no inherent difference in information content between consonant-vowel (CV) and vowel-consonant (VC) transitions. They describe a new experiment, in which they eliminate a possible "click sensation" by smoothing the onsets of CV transitions and the offsets of VC transitions by means of an amplitude-smoothing time window. They claim that, although there are now no abrupt onsets anymore, the original advantage of final over initial transitions remains, and that the difference between the two is consequently an inherent one. The purpose of our present reply is, firstly, to indicate that we think that Ohde and Sharf's (1980) data do not differ essentially from ours, and secondly, to discuss some considerations they raise in their discussion.

I. INTERPRETATION OF THE DATA

In Table I, we present a summary of the figures on which Ohde and Sharf (1980) base their claim, and on which we base ours. We include the results of an experiment on American English plosives, recently conducted by one of us and described in Pols (1979).

A few things should first be said about differences between stimuli, experimental designs, and languages. Ohde and Sharf (1980) used CV-type syllables and separate subsets for voiced /b, d, g/ and voiceless /p, t, k/ stimuli, whereas we used CVC syllables and mixed voiced–voiceless presentations and response sets. Moreover, Dutch differs from English in that Dutch plosives are unaspirated and do not have initial /g/ or final voiced plosives. Dutch voiced plosives are usually prevoded, and it is not uncommon for a Dutch /t/ to be mistaken for a /d/ by a native speaker of English. So identification results will only be comparable in a relative sense.

Ohde and Sharf (1980) conclude from their new identification experiment with a time window that "the reliability and direction of the results ...," with respect
TABLE I. This table contains the identification scores for deleted plosive consonants from all the studies mentioned in this Letter. The deletion procedure resulted in stimuli consisting of vocalic transition + vowel for initial plosive identification or vowel + vocalic transition for final plosive identification. The first five columns specify the experimental conditions, namely: Segmentation procedure; whether voiced or voiceless subsets were used, or a full set; the syllable type (CV and VC or CVC); the language used (AE = American English, D = Dutch); and whether or not a time window or noise were applied. The next four columns give the detailed scores; the last two columns give combined scores.

<table>
<thead>
<tr>
<th>Study</th>
<th>Segment procedure</th>
<th>Stimulus response</th>
<th>Syllable</th>
<th>Language</th>
<th>Window noise</th>
<th>Initial (CV)</th>
<th>Final (VC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharf and Better (1974)</td>
<td>splicing</td>
<td>subset</td>
<td>C</td>
<td>AE</td>
<td></td>
<td>64</td>
<td>39</td>
</tr>
<tr>
<td>Sharf and Hemeyer (1972)</td>
<td>splicing</td>
<td>subset</td>
<td>C</td>
<td>AE</td>
<td></td>
<td>66</td>
<td>26</td>
</tr>
<tr>
<td>Ohde and Sharf (1977)</td>
<td>erasure</td>
<td>subset</td>
<td>CV</td>
<td>AE</td>
<td></td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>Ohde and Sharf (1980)</td>
<td>digital</td>
<td>subset</td>
<td>CV</td>
<td>AE</td>
<td>window</td>
<td>72</td>
<td>40</td>
</tr>
<tr>
<td>Pols (1979)</td>
<td>digital</td>
<td>full set</td>
<td>CVC</td>
<td>AE</td>
<td>noise</td>
<td>54</td>
<td>17</td>
</tr>
<tr>
<td>Pols (1979)</td>
<td>digital</td>
<td>full set</td>
<td>CVC</td>
<td>AE</td>
<td></td>
<td>82</td>
<td>26</td>
</tr>
<tr>
<td>Pols and Schouten (1978)</td>
<td>digital</td>
<td>full set</td>
<td>CVC</td>
<td>D</td>
<td></td>
<td>74</td>
<td>34</td>
</tr>
</tbody>
</table>

We find it difficult to include in this discussion the data from similar experiments by Sharf and Hemeyer (1972) and Sharf and Better (1974). As can be seen in the first three rows of Table I, there is considerable inconsistency among these scores and those found by Ohde and Sharf (1977) whereas the experimental conditions were quite similar. There were two differences, however, which may have influenced the results: (1) Sharf and Hemeyer (1972) and Sharf and Better (1974) used only the vowel /o/, whereas Ohde and Sharf (1977) used a larger vowel set; (2) the segmentation procedure used in the two earlier papers was tape splicing, replaced in 1977 by tape erasure. We consider the more recent data of Ohde and Sharf (1977) to be the more reliable ones; moreover, in that study, the set of vowels is more similar to the one used by Ohde and Sharf (1980).

II. IMPLICATIONS

In our 1978 paper, we used the term "click sensation" for want of a better term. What we meant by it was the effect produced by the inevitable spectral consequences of an abrupt onset. We felt that this effect differed only in degree, but not in principle, from the clearly audible click produced by cutting and splicing at a high amplitude point on a waveform. Combining Ohde and Sharf's (1980) and our own results, we think it may be said that the addition of noise or of a time window reduces, perhaps even cancels, the unwanted onset spectrum. Such an effect is to be expected if a time window is used, provided not too much information is smoothed away. With noise, however, the effects are perhaps less predictable, although the net results appear to be at least as effective. In addition to supplying the required spectral continuity, added noise could also (1) serve as an alerting signal to subjects (this could be investigated by having noise bursts of various durations); (2) mask the no-signal situation, with unpredictable effects. If no noise is applied there is silence instead, with the con-
sequence that no information is present at all; (3) allow subjects to suppose that the original information could be present but has been made inaudible ("hear through the noise") (Plomp, 1980).

The third possibility is the one we briefly hinted at in our paper, where we said that this "possibility must not be excluded." The truth is, however, that we do not really know exactly what happens when noise is added to a signal; how, for example, would it provide "spectral continuity?" The fact that Ohde and Sharf's (1980) and our results are of the same magnitude could indicate that the effects caused by a time window or noise are the same, but it is too early to say that.


Maurice J. M. Martens
Section Experimental Ecology, Department of Botany, Catholic University, 6525 ED Nijmegen, The Netherlands

Axel Michelsen
Institute of Biology, Odense University, DK 5230 Odense M, Denmark
(Received 6 May 1980; accepted for publication 22 July 1980)

We investigated the vibration of leaves of four plant species in a sound field using a laser-Doppler-vibrometer system. All leaves behave as linear mechanical systems when driven by sound and noise at sound pressure levels (SPL) of up to 100 dB re 20 μ Pa. The modes of vibration are complex in the investigated frequencies (0.5-5.5 kHz), and change with the orientation of the leaf in the sound field. The vibration velocities of the leaves varied between 10^-5 and 3 X 10^-4 m/s, while the vibration velocity of the air particles is 5 X 10^-6 m/s at 100 dB SPL. Although the amount of sound energy absorbed in this way by a single leaf is very small, this mechanism may anyhow contribute to sound attenuation by plants and plant communities, since the number of leaves of one fullgrown tree equals 2 X 10^5.

PACS numbers: 43.80.Gx, 43.28.Hr

INTRODUCTION

The acoustic properties of the environment are important with regard to noise pollution problems and noise abatement, and for the understanding of sound communication by animals. Vegetation is a complex medium consisting of air, soil, and the stems, trunks, branches, and foliage of herbs, shrubs, and trees. Despite both theoretical and experimental work, the mechanisms of sound attenuation in various types of plant communities remain obscure. Among the mechanisms which have been suggested are the ground effect, the thermoviscous absorption of sound at the soil and in the boundary layer of air at the surface of the leaves, scattering from trunks, branches, and leaves, and the vibration of thin branches and leaves. The aim of this paper is to investigate the vibration of plant leaves in a sound field using a laser-Doppler-vibrometer system. The acoustic force acting upon a leaf is determined by the vectorial difference between the sound pressures at its two surfaces, cf. a ribbon microphone and other pressure gradient receivers. A part of the sound energy received by a leaf may be lost by friction and thus absorbed, and some of the energy may be re-emitted. At low frequencies the force acting on the leaf is small, and the re-emission of sound is likely to be inefficient. At higher frequencies the driving force is larger, and a considerable part of the energy received