THE PERCEPTION OF SILENT-CENTER SYLLABLES IN NOISE

Susan L. Hura

University of Texas, Austin, Texas

ABSTRACT

The perception of vowel-less /b/-vowel-/t/ syllables was tested at various signal to noise ratios. Contrary to what has been shown in previous studies [7], vowels in these "silent-center" syllables were not identified at the same accuracy as vowels in full syllables. This calls into question the degree to which the perception of silent-center syllables can be seen as evidence for the theory of dynamic specification.

1.INTRODUCTION

Dynamic specification is a recent theory of vowel perception proposed by Strange [7] in which vowels are conceived of as gestures having intrinsic timing parameters. Dynamic specification is in opposition to a traditional target theory which states that vowel recognition is based upon characteristic frequency values for the first 2 (or 3) formants taken from a single time slice in the syllable nucleus. Strange cites certain perceptual results in support of dynamic specification. Of specific interest here is the correct identification of vowels in vowel-less syllables.

Using a wavefrom editing technique, "silent-center" syllables were generated, in which the vowel nucleus was attenuated to silence, leaving 3 or 4 pitch periods of consonant transition at either edge of the syllable. Strange found that subjects were able to identify the vowels in silent-center [SC] syllables with nearly the same accuracy as they could the vowels in full syllables, thus refuting a simple target theory. If recognition can proceed in the absence of vowel nucleus information, then this information is not the determining property of vowel identity.

In this paper, Strange's SC result is reconsidered. Let us begin with the assumption that target formant values, formant transitions and other dynamic attributes all play a role in the identification of vowels. These factors normally provide redundant and overlapping information about vowel identity, thus it is not surprising that identification can be relatively accurate in the absence of some of this information. SC syllables are an example of a stimulus where some vowel information is absent, but a great deal remains. In favorable listening situations, it is possible to make up for the lack of one sort of information by focusing on remaining information. Because vowel identification is a familiar task and an experimental setting is relatively free of distractions and ambient noise, the listening conditions in Strange's experiment were close to ideal. Under degraded listening conditions, however, listeners may rely more on each source of information than they otherwise would. My claim, then, is that Strange's SC result is due to the favorable listening conditions under which she tested identification performance. To investigate this claim, Strange's Experiment 3 [7] was partially replicated.

2. PERCEPTUAL STUDY

2.1. Stimuli

Stimulus materials consisted of /b/-vowel-/t/ syllables in the carrier phrase "I say the word /bVt/ somemore," for each of 10 vowels. The speaker was an adult male with a midwestern dialect. Stimuli were digitized and waveform edited to produce SC syllables according to criteria defined by Strange [7].

Full and SC syllables were then embedded in wide-band noise. Two

sorts of SC syllables in noise were created: in the first (SC1), the amplitude of the initial and final components has been boosted such that their peak amplitude is equal to the full syllable peak; in the second set (SC2), amplitude of initial and final components is the same for full and SC versions of a syllable at the same S/N. Six S/N were created for each syllable type for each /bVt/ by varying the amplitude of the signal in relation to constant amplitude noise. All stimuli were embedded into the same carrier phrase.

2.2. Subjects & Procedure

Stimuli were randomized, and a listening test was created. Stimuli were presented in blocks of 21; interstimulus interval was 4 seconds, interval between blocks was 8 seconds. The task of the subject was to circle the /bVt/ word they heard on a preprinted answer sheet. There were a total of 252 items in the test--it lasted approximately 30 minutes.

The 26 subjects tested were U.T. undergraduate volunteers who were paid for their participation. 20 of the subjects spoke a Texas dialect; dialects of the remaining 6 varied. Subjects were tested in a laboratory setting in groups of 3 to 6.

Table 1: Overall percent correct by syllable type.

	-6	-3	0	3	6	9
full	60.3	74.2	85.0	87.7	90.4	89.6
SC1	53.1	48.8	76.5	86.5	75.4	80.0
SC2	56.5	58.1	60.4	73.5	78.9	84.2

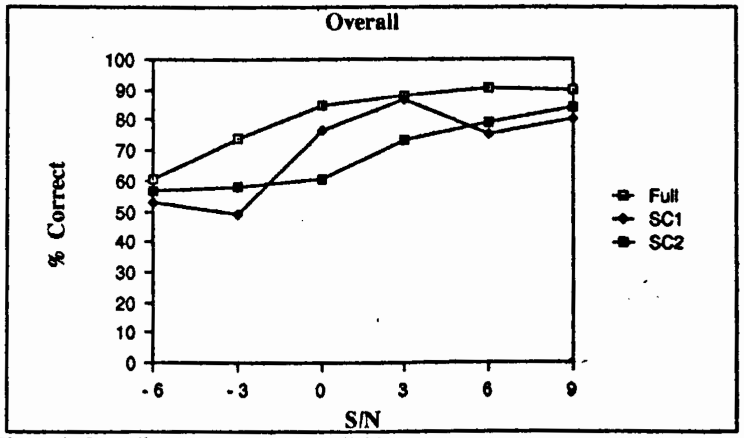

Figure 1: Overall percent correct by syllable type.

2.3 Results & Discussion

Table 1 below gives percent correct by syllable type and S/N, collapsed across vowels. Examining this data it is apparent that vowels were perceived more accurately in full syllables than in either SC1 or SC2 syllables. Figure 1 shows these same results graphically.

A two-way analysis of variance on syllable type and S/N was performed; both of these factors were shown to have an effect, but their interaction does not. Syllable type (full versus SC1 versus SC2) is significant for $F(2,162)=4.22$ at $p<.025$; S/N is also significant for $F(5,162)=8.01$ at $p<.001$. T-tests for differences among the means of the 3 conditions were performed, which showed that full syllables are significantly different from either type of SC syllable ($p< .025$), and that the two types of SC syllables are not significantly different. Thus it appears that given a more difficult identification task, vowels are significantly more difficult to perceive in SC syllables.

3.CONCLUSIONS

The results of this study show that while it is easy to identify the vowels in silent-center syllables, this identification is not as accurate as for full syllables. It was shown that under degraded listening conditions, when the listener is more dependent upon the redundancies of the speech signal, identification performance is significantly better for full syllables than for silent-centers. The ability to accurately perceive a vowel in a syllable where it one is not physically present is certainly remarkable. The current data show that even at low S/N, subjects identify vowels in SC syllables at well above chance level. However, given the poorer identification performance on SC syllables in difficult listening situations, we are not justified in claiming that transition information has greater importance than the nucleus in the specification of vowels. This is not to say that nucleus information is privileged, for the representation of a vowel as a static point in F1/F2 space is clearly insufficient to account for the present results. It is clear, however, that syllables containing nucleus information are better perceived than those without it.

Perhaps a dual-target model of vowel specification [1] can provide an explanation for the current results. If a vowel is specified by formant values in both the nucleus and offglide, identification should be more accurate when both of these are present in the stimulus, as is the case for full syllables. When some of this information is missing, as in SC syllables, a dual-target model predicts the poorer identification performance shown here.

4. REFERENCES

[1] ANDRUSKI, J. & NEAREY,T. (forthcoming), "Comparing the perception isolated vowel and /bVb/ syllables in hybrid silent center stimuli" *JASA*.
[2] DIEHL,R., McCUSKER,S. & CHAPMAN, L. (1981), "Perceiving vowels in isolation and in consonantal context" *JASA, 69*, 239-248.
[3] JOOS, M.(1948). "Acoustic Phonetics" *Language Supplement, 24*, 1-136.
[4] LINDBLOM, B. & STUDDERT-KENNEDY, M. (1967), "On the role of formant transitions in vowel recognition" *JASA, 42*, 830-843.
[5] NEAREY, T. and ASSMANN, P. (1986), "Modelling the role of inherent spectral change in vowel identification" *JASA, 80*, 1297-1308.
[6] PARKER, E. & DIEHL, R. (1984), "Identifying vowels in CVC syllables: effects of inserting silence and noise" *Perception & Psychophysics, 34*, 369-380.
[7] STRANGE, W. (1989), "Dynamic specification of coarticulated vowels spoken in sentence context" *JASA,85 (5)*, 2135-2153.
[8] STRANGE, W., VERBRUGGE, R., SHANKWEILER, D.& EDMAN, T. (1976), "Consonantal context specifies vowel identity" *JASA, 60*, 213-224.