

Compensation to articulatory perturbation: Perceptual data

Shari R. Baum^{a)}

School of Communication Sciences & Disorders, McGill University, 1266 Pine Avenue W., Montréal, Québec H3G 1A8, Canada

David H. McFarland and Mai Diab

Ecole d'orthophonie et d'audiologie, Université de Montréal, C.P. 6128, Succ. A, Montréal, Québec H3C 3J7, Canada

(Received 31 July 1995; revised 17 January 1996; accepted 2 February 1996)

The perceptual adequacy of vowels, stop consonants, and fricatives produced under conditions of articulatory perturbation was explored. In a previous study [McFarland and Baum, *J. Acoust. Soc. Am.* **97**, 1865–1873 (1995)], acoustic analyses of segments produced in two subtests (immediate compensation and postconversation) revealed small but significant changes in spectral characteristics of vowels and consonants under bite-block as compared to normal conditions. For the vowels only, adaptation increased subsequent to a period of conversation with the bite block in place, suggesting that compensation may develop over time and that consonants may require a longer period of adaptation. The present follow-up investigation examined whether the acoustic differences across conditions were perceptually salient. Ten listeners performed an identification and a quality rating task for stimuli from the earlier acoustic study. Results revealed reductions in identification scores and quality ratings for a subset of the vowels and consonants in the bite-block conditions relative to the normal condition in the immediate compensation subtest. In the postconversation subtest, quality ratings for the fricatives in the bite-block condition remained low as compared to those in the normal condition. Perceptual results are compared to the previous acoustic data gathered on these stimuli. © 1996 Acoustical Society of America.

PACS numbers: 43.71.Gv, 43.70.Bk

INTRODUCTION

Speech adaptation to perturbations of the oral-articulatory environment has been the subject of a great deal of research. Investigators have demonstrated, for example, that speakers are easily able to produce relatively accurate vowels despite fixation of the mandible by a bite block (Flege *et al.*, 1988; Fowler and Turvey, 1980; Gay *et al.*, 1981; Kelso and Tuller, 1983; Lindblom *et al.*, 1979; Lindblom and Sundberg, 1971). However, more recent work in our lab has revealed that there are small but significant differences in the acoustic characteristics of vowels and consonants produced under fixed-mandible conditions (McFarland and Baum, 1995; see also Flege *et al.*, 1988).

Specifically, we conducted acoustic analyses of vowels ([i a u]) and consonants ([p t k s ʃ]) produced by 15 speakers of (Québec) French under bite-block and jaw-free conditions. Two subtests were included to examine immediate compensation and compensation subsequent to a 15-min accommodation period of spontaneous conversation with a bite block in place. In the immediate compensation subtest, small but significant differences in vowel formant frequencies were found between bite-block and jaw-free conditions. Further, centroid frequencies (Forrest *et al.*, 1988) calculated for consonantal segments were lower in the bite-block relative to the normal condition. Fewer significant differences emerged across conditions in the postconversation subtest for the vowels, suggesting that some improvement in compensation

may have occurred during the accommodation period. In contrast, centroid values of both stop and fricative consonants continued to differ under the perturbed and normal conditions. Based on these results, McFarland and Baum (1995) hypothesized that compensatory strategies may develop over time via sensory feedback and that consonants are likely to require a longer period of practice before complete adaptation may be achieved (see also, e.g., Flege *et al.*, 1988; Hamlet and Stone, 1976). Although acoustic differences across normal and bite-block conditions were found, it is equally important to determine whether listeners are sensitive to these spectral shifts. Therefore, the present investigation was designed to assess whether the acoustic differences reported by McFarland and Baum (1995) are perceptually salient.

I. METHOD

A. Subjects

Ten adult female native speakers of (Québec) French (aged 18–28 years) participated in the perceptual experiment. Listeners were unaware of the purposes of the investigation, had received no training in phonetic transcription, were free from speech and/or language disorders and passed an audiometric screening.

B. Stimuli and procedures

A subset of the stimuli analyzed acoustically in the earlier investigation (McFarland and Baum, 1995) were selected for perceptual judgements in the present study. The reader is

^{a)} Author to whom correspondence should be addressed. E-mail: INSR@MUSICB.MCGILL.CA

referred to this earlier publication for details of experimental methods. The stimuli included the three vowels [i a u], the voiceless stop consonants [p t k] in the environment preceding the same three vowels, and the voiceless fricatives [s ʃ] in the same vowel environments. Speakers produced stimuli in two subtests: immediate compensation and postconversation. The immediate compensation subtest included three conditions: jaw-free (or normal, N), small bite block (SBB: 2.5 mm for vowels and 5 mm for CV stimuli) and large bite block (LBB: 22.5 mm for vowels and 10 mm for CV stimuli). In the postconversation subtest, a single bite-block condition (BB: 10 mm) was compared to a jaw-free condition, and stimuli were elicited after a 15-min period of conversation with the bite block in place.

The isolated vowel and consonant segments defined in McFarland and Baum (1995) were used to create six perception tests, one for each phoneme class (vowels, stops, and fricatives) and subtest (immediate compensation and postconversation). Vowels were defined from the onset through the end of periodicity. Stop consonants were demarcated from the burst through the end of aspiration (equivalent to VOT), and fricative consonants were delimited by the onset and offset of frication noise. Thus consonants were presented isolated from their vowel environments. Three productions for each of the 15 speakers in each condition were selected randomly for each for the six tests.

For perceptual judgements, stimuli were presented in random order to listeners via headphones at a comfortable loudness level. The order of the six perceptual tests was counterbalanced across listeners. Listeners were asked to both identify the sound presented from a limited set of alternatives¹ and rate its quality. For example, for the stop consonant tests, subjects were provided with the choices [p t k] and were instructed to select from among these choices the sound that most closely matched the sound they heard. They were then asked to rate the quality of the sound on a five point scale, with the anchor words being “unintelligible” and “perfect.” A 6-s intertrial interval separated the stimuli, and each test was preceded by six practice trials to familiarize the listeners with the experimental procedures.

II. RESULTS

Percent correct identification responses and quality ratings were calculated for each of the vowel and consonant stimuli in each condition. Overall means and standard deviations for the immediate compensation and postconversation subtests are presented in Figs. 1 and 2, respectively. Analyses of variance (ANOVAs) were computed both by speaker and by listener for identification responses and quality ratings for each of the three sound types. Adopting a conservative approach, we will report only those differences that were significant in both the analyses by speaker and listener.

A. Immediate compensation

In addition to the accuracy data for vowels shown in Fig. 1(a), confusion matrices revealed that [i] was most often misperceived as [u], and misidentification increased across bite-block conditions (N: 6%, SBB: 10%, LBB: 21% [u] responses). Statistical analyses revealed no significant

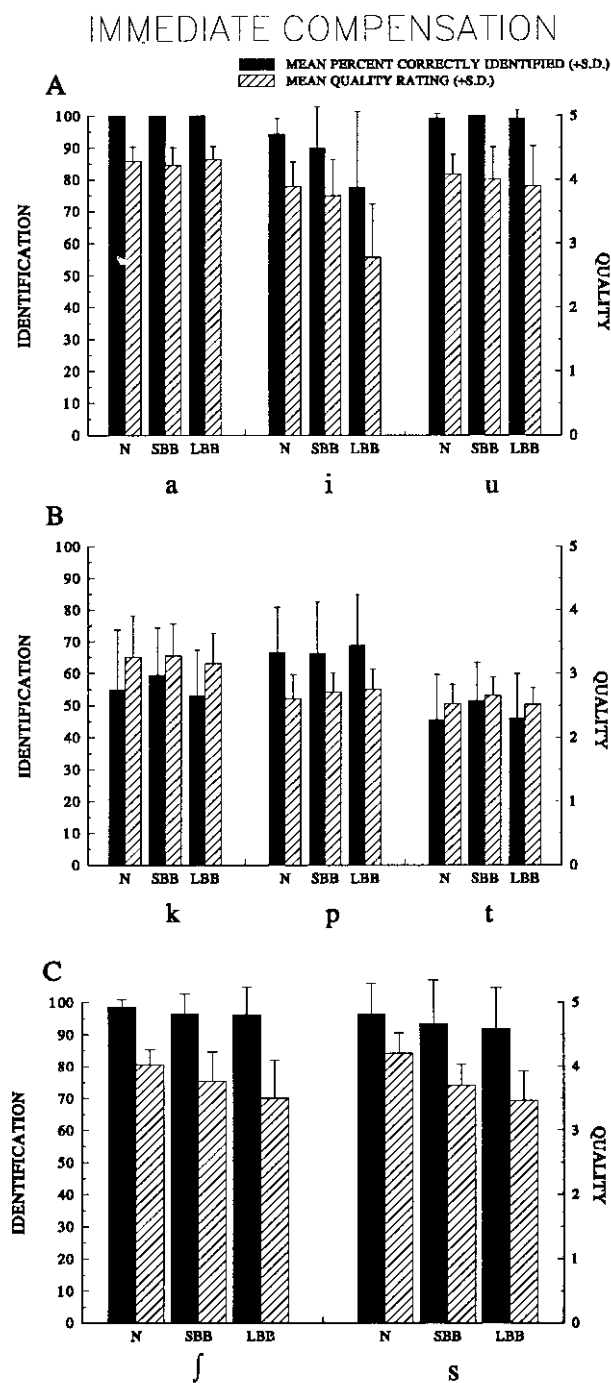


FIG. 1. Percent correct identification (and standard deviation) and quality ratings for vowels (a), stop consonants (b), and fricatives (c) in the immediate compensation subtest. LBB=large bite block, SBB=small bite block.

differences in vowel identification across conditions, but a significant vowel by condition interaction for vowel quality ratings [$F(4,52)=18.073$, $p<0.001$]. Newman-Keuls *post hoc* analyses ($p<0.05$) revealed significantly lower quality ratings for [i] in the LBB as contrasted to both the N and SBB conditions, which did not differ. No significant differences were found between conditions for the other two vowels.

Identification accuracy tended to be lower for stop consonants [Fig. 1(b)] when contrasted with vowels [and fricatives, Fig. 1(c)]. Confusion matrices revealed that [k] and [p]

POST CONVERSATION

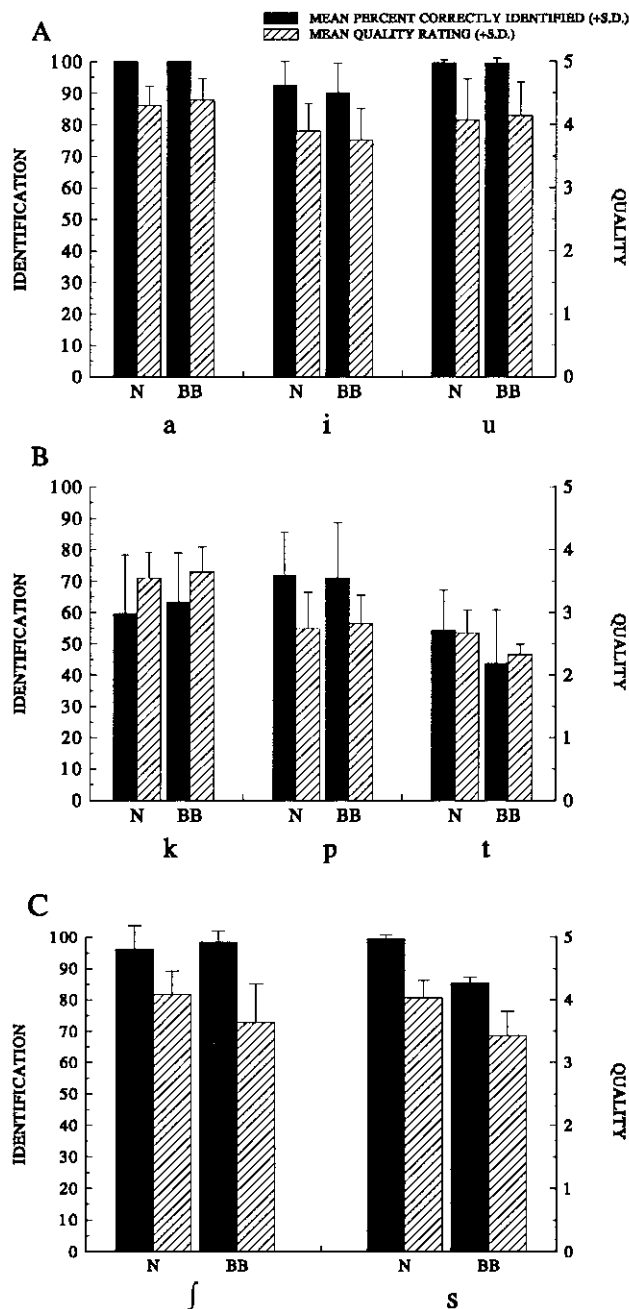


FIG. 2. Percent correct identification (and standard deviation) and quality ratings for vowels (a), stop consonants (b), and fricatives (c) in the postconversation subtest. BB=10-mm bite block.

were most often misperceived as [t] ([k]—N: 33%, SBB: 27%, LBB: 30%; [p]—N: 25%, SBB: 27%, LBB: 25%). [t] was most often misidentified as [p] (N: 43%, SBB: 36%, LBB: 41%). No significant effects of condition were observed in stop identification or quality ratings.

As illustrated in Fig. 1(c), a significant main effect of condition emerged for the fricative quality ratings [$F(2,26)=15.931$, $p<0.001$]. *Post-hoc* analyses revealed significant differences between N and both SBB and LBB conditions; the latter two conditions did not differ.

B. Postconversation

Percent correct identification and quality ratings of vowels in the postconversation condition [Fig. 2(a)] were overall very high and varied little across bite-block conditions. There was a tendency, as in the immediate compensation condition, for [i] to be misidentified as [u] (N: 7%, BB: 8%) more than any other confusion.

As illustrated in Fig. 2(b), identification scores and quality ratings for stops varied considerably, and no significant trends emerged across bite-block conditions. Again, [k] and [p] were most often misperceived as [t], which in turn was most often misidentified as [p].

For fricatives [Fig. 2(c)], a significant main effect of condition [$F(1,14)=5.141$, $p<0.05$] and a condition by fricative interaction [$F(1,14)=15.232$, $p<0.002$] were found for identification scores. *Post-hoc* analyses revealed that identification scores were significantly lower in the BB as contrasted to the N condition for [s] only. Quality ratings were lower in the BB as contrasted to the jaw-free condition for both fricatives, as confirmed by *post-hoc* analyses [$F(1,14)=20.005$, $p<0.001$].

III. DISCUSSION

The results of the present investigation indicate that many of the acoustic changes caused by perturbation of the articulators via jaw fixation (McFarland and Baum, 1995) were, indeed, perceptually salient. The spectral deviations found in our previous acoustic study were reflected mainly in the quality ratings in the present investigation. In general, the findings are in accord with the acoustic analyses reported in McFarland and Baum (1995), particularly for the vowel and fricative data.

Perceptual identification scores for the stop consonants were low, averaging 57% and 61% in the immediate compensation and postconversation subtests, respectively. The low rates of identification even in the normal condition may reflect the difficulty of identifying voiceless stops extracted from varying vowel contexts due to coarticulatory effects (e.g., Winitz *et al.*, 1972). Crucially for the purposes of the present investigation, there were no differences in identification performance across bite-block conditions. Quality ratings for the stop consonants were much lower than those for the vowels, but again did not differ across perturbation conditions. The pattern of results for the stop consonants is not consistent with the acoustic data reported by McFarland and Baum (1995), which showed significant differences in centroid frequencies in bite-block as compared to normal conditions. It may be that perceptual differences were obscured in the very brief stop consonant stimuli. Perhaps the within category nuances in spectral cues across conditions were too short to be perceived in the current paradigm. Or, as noted above, presenting the stop consonants isolated from neighboring vowels may have made the task too difficult to be sensitive to the effects of bite block perturbation.

Taken together, the present findings reinforce the conclusions of our acoustic study (McFarland and Baum, 1995; see also Flege *et al.*, 1988), indicating that compensation for fixation of the jaw by a bite block is not as immediate or as

complete as previously hypothesized (cf. Kelso and Tuller, 1983; Lindblom *et al.*, 1979; Gay *et al.*, 1981). By examining these rather subtle shifts in acoustics and perception in a range of sound classes, we may explore the development of new motor programs and their potential changes over time in response to articulatory perturbation.

ACKNOWLEDGMENTS

We are grateful to P.-A. Ménard for technical assistance. This research was supported in part by the Fonds de la Recherche en Santé du Québec (FRSQ).

¹Although for the vocalic stimuli one might expect misperceptions to involve vowels other than the three point vowels tested (e.g., [ɪ] for [i]), as a first step we decided to limit the possible choices to the vowels actually produced. This was done in order to maintain as much consistency as possible across the sound classes and to keep the untrained listeners' identification task as simple as possible so that the quality rating task could be performed quickly and easily.

Flege, J. E., Fletcher, S. G., and Homiedan, A. (1988). "Compensating for a bite block in /s/ and /t/ production: palatographic, acoustic, and perceptual data," *J. Acoust. Soc. Am.* **83**, 212–228.

- Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. (1988). "Statistical analysis of word-initial voiceless obstruents: Preliminary data," *J. Acoust. Soc. Am.* **84**, 115–123.
- Fowler, C. A., and Turvey, M. T. (1980). "Immediate compensation in bite-block speech," *Phonetica* **37**, 306–326.
- Gay, T., Lindblom, B., and Lubker, J. (1981). "Production of bite-block vowels: acoustic equivalence by selective compensation," *J. Acoust. Soc. Am.* **69**, 802–810.
- Hamlet, S., and Stone, M. (1976). "Compensatory vowel characteristics resulting from the presence of different types of experimental dental prostheses," *J. Phon.* **3**, 199–218.
- Kelso, J. A. S., and Tuller, B. (1983). "'Compensatory articulation' under conditions of reduced afferent information: a dynamic formulation," *J. Speech Hear. Res.* **26**, 217–224.
- Lindblom, B., Lubker, J., and Gay, T. (1979). "Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation," *J. Phon.* **7**, 147–161.
- Lindblom, B. E. F., and Sundberg, J. E. F. (1971). "Acoustical consequences of lip, tongue, jaw, and larynx movement," *J. Acoust. Soc. Am.* **50**, 1166–1179.
- McFarland, D. H., and Baum, S. R. (1995). "Incomplete compensation to articulatory perturbation," *J. Acoust. Soc. Am.* **97**, 1865–1873.
- Winitz, H., Scheib, M., and Reeds, J. (1972). "Identification of stops and vowels for the burst portion of /p,t,k/ isolated from conversational speech," *J. Acoust. Soc. Am.* **51**, 1309–1317.