REPORT: SPEECH PRODUCTION
(see vol. I, p. 11-56)

Reporter: Peter F. MacNeilage
Co-reporter: Peter Ladefoged
Co-reporter: Masayuki Sawashima

Chairpersons: Antti Sovijärvi and Hiroya Fujisaki

REPORTER'S ADDITIONAL REMARKS

P. MacNeilage, in his presentation, commented on the question of the control of speech production and the biological basis of speech.

The first comments dealt with the role of feed-back. P. MacNeilage claimed that if one considers how we produce speech under the various postural circumstances, we are forced to conclude that peripheral somatic feed-back plays a virtually continuous role in the control of speech production. It must be a system that can sense at the periphery what the present posture is and that is required to monitor the attempts of the control system to produce speech in any particular posture. We can't be assumed to be infinitely versatile in terms of preprogramming at all postural circumstances. Furthermore, P. MacNeilage pointed out that the concept of normal speech production is perhaps misleading, since most of our work is done in the laboratory with the subjects looking straight ahead and in a fixed position. This is not the normal posture and very little of our work has dealt with postural variations. P. MacNeilage continued by saying that we know very little about how the feed-back works and that we need more information which may perhaps come from people doing research in dentistry. He warned against conclusions drawn from physiological studies of animal limbs, since the human somatic sensory system differs from the animal system in many significant ways. In addition, P. MacNeilage found that the results of experiments where the posture is artificially manipulated, such as in the bite-block studies and in studies where the jaw movement is impaired, support the argument about the necessity for feed-back.

Then P. MacNeilage raised the question: This feed-back is feed-back to what? Among other things he pointed out that it seems necessary in speech production to recognize a multiplicity of levels of organization, some of which are quite accessible to us

and others which are not. But it is nevertheless crucial for us to understand those higher levels if we want to come up with a plausible theory of speech production. In this connection, P. MacNeilage stated that there has to be a distinction between a context sensitive system at a lower level of organization and some kind of context independent entity or set of entities at a higher level, referring among other things to segmental spoonerism. We produce sequences with spoonerisms fluently which means that subsequent to the permutation, the context sensitive control system makes the appropriate adjustments. He noticed that very often spoonerisms involve single segments, and very few can be unequivocally labeled distinctive feature movement type errors, and relatively few involve whole syllables. This means that at least at one level of organization the segmental unit is an extremely important one for speech production.

Before leaving the topic of control, MacNeilage stated that our rather simple algorithms do not account very well for the dynamic aspects of speech production, referring to differences in stress and speaking rate, and to coarticulation. The same speaker can use different strategies in changing the speaking rate, for instance, which also proves that we are dealing with an extremely versatile control system.

Turning now to the question of the biological basis of speech production, P. MacNeilage emphasized - as he does in his paper - that we have very much neglected the study of prelinguistic vocalization in our studies of speech production. This neglect may be due to R. Jakobson's theory of language acquisition which assigned babbling to "external" phonetics. P. MacNeilage claimed that the phonetic forms of early speech with reference are extremely similar or identical in many cases to the babbling forms that immediately preceded them. This means that the same production system that has been working earlier in the proto-language stage is still an extremely important component in early referential speech. P. MacNeilage claimed that babbling begins at a particular time on a particular day. Finally, he stated that babbling is some kind of innate movement control organization that is "there" in relation to speech.

DISCUSSION

John Ohala and John Laver opened the discussion.

J. Ohala stated that from his point of view one of the very promising and most essential developments in current work on speech production is the large number of models, including various aspects of the articulatory apparatus, which have been developed in the past decade or so. He believes that the rise of model-making is a development of the computer revolution in the laboratory and that it has come of age where we have become familiar with and have used computers to develop models which in many cases are conceptually simple, but which require computationally rather complex activity. Some objections have occasionally been raised against model-making, usually along the lines of: "Well, you have made the model, you have put the properties into it that it has, why can't you figure out what it is supposed to do in advance, why bother with it? It is simply making explicit what you already know or what you assume to be true." In order to parry off this kind of objection, Ohala referred to the Nobel Prize winner H. Simon, who indicates that it may very well be true that in model-making-like abstract logic and didactic logic and so on - the consequences of a particular set of assumptions must naturally follow in an automatic, perfectly regular way. But when our models and the assumptions in them get sufficiently complex, really only God can figure out what the consequences of these assumptions may be. The rest of us have to work them out painstakingly, teasing them out for understanding, and this is why we make models. Furthermore, Ohala pointed out that our models serve a very interesting heuristic purpose in that they tell us what to look for in the data. This was made evident to him in working with an aerodynamic model revealing that if one is going to have production of a fricative or some kind of fricated segment one should not have nasal leakage, obviously because the air flowing out of the nasal cavity would prevent the build-up of the high pressure drop necessary to produce the turbulence. And Ohala asked whether this has phonological consequences. He pointed out that he had never seen any observation of this in the literature, but when he searched for it he was able to come up with a number of examples from sound change and allophonic variation. For example, English has a palatal fricative

as an allophone of /h/ before the palatal glide /j/ in words like Hugh and human. But that same allophone is no longer a fricative if we embed it in a heavily nasalized environment as in the word inhuman. With this example Ohala illustrates how models can tell us what to look for and in that sense even help us to enhance our naturalistically obtained data base.

Then Ohala addressed one comment to Sawashima concerning the vertical tension of the vocal folds. Sawashima said in his co-report that there is no evidence for the existence of any physiological mechanism whereby vertical compression or tensing of the cords could affect F_0. However, it is well known that the average F_0 of vowels is positively correlated with the "height" of vowels. But, to date, no one has found any significant difference in the degree of muscle activity of the intrinsic laryngeal muscles during the production of various vowels. On the other hand, van den Berg (1955), Shimizu (1960, 1961), and additional workers cited in Žinkin (1968:353) have found that the laryngeal ventricle is larger, both in width and vertical depth, during the production of high vowels such as [i] and [u] - thus showing greater separation between the ventricular folds and the vocal folds - but smaller during the production of low vowels. Also, Luchsinger and Arnold (1965:223) describe a patient with bilateral paralysis of the cricothyroid muscles but who could nevertheless vary F_0 over a few semitones. X-rays revealed no change in the angle of the cricothyroid visor but the whole larynx was higher in the neck during the production of high F_0. (More detailed arguments for F_0 variation due to vertical tension have been given in Ohala 1972, 1977, 1978.)

References

Berg, J. van den (1955): "On the role of the laryngeal ventricle in voice production", Fol. phon. 7, 57-69.

Luchsinger, R. and G.E. Arnold (1965): Voice-speech-language; clinical communicology: its physiology and pathology. Belmont: Wadsworth.

Ohala, J. (1972): "How is pitch lowered?", JASA 52, 124.

Ohala, J. (1977): "Speculations on pitch regulation", Phonetica 34, 310-312.

Ohala, J. (1978): "The production of tone", in Tone: a linguistic survey , V. Fromkin (ed.), 5-39. New York: Academic Press.

Shimizu, K. (1960): "On the motions of the vocal cords in phonation studied by means of the high voltage radiograph movies". [In Japanese; English summary]. Oto-Rhino-Lar. Clinic [Zibi-Inko-Ka Rinsyo] 53, 446-461.