Shimizu, K. (1961): "Experimental studies on movements of the vocal cords during phonation by high voltage radiograph motion pictures", Studia Phonologica 1, 111-116.

Žinkin, N.I. (1968): Mechanisms of speech, The Hague: Mouton.

J. Laver had four points to make about the issues raised in the three reports.

The first of them dealt with methods for estimating the different muscular forces acting on and in the tongue, as in the work of Fujimura, Kakita, and Perkell. He referred to a finding from speech error work based on an experiment to provoke subjects into making the kind of vowel-blend errors that Rulon Wells claimed almost never happen. The structure of the experiment was to push subjects just beyond the comfortable limit of accurate performance of target vowels. Facing them on a screen were two words - for example PEEP and PARP - and above the two words were two stimulus lights, and the task was to pronounce each word as accurately as possible immediately the associated light came on. The lights were programmed to come in random sequence, with 200 msec duration, with intervals of 200 msec. In this condition, all subjects made vowel errors, two types of diphthongs and one type of monophthong. When PEEP and PARP were in competition the two diphthong errors were either PAIP or PIAP. Laver proposed the following hypothesis to explain this result. One might imagine that the commands to the relevant muscle systems had a slight difference in the time course such that if the commands for AR preceded those for EE then one got PAIP and if the commands for EE preceded those for AR one got PIAP. But if the commands to the different muscle systems were issued perfectly simultaneously, then the monophthong [ɜ] as in PURP was the result as the mechanically joint product of the action of simultaneously activated different muscle systems. The relevance of this finding to the problem of estimating relative muscle system forces is, that if we look at the interactions of all pairs of vowels, then the "mechanically joint product" position of the intermediate vowel does not necessarily coincide with the geometric mean position between the two target positions. In the competition between PEP and POOP, for example, the intermediate monophthong was [œ] as in [pœ:p], in other words rather closer to the [ɛ] target than to the [u] target, as the lip position also, one might think, was slightly closer to the [ɛ] target

than to the close rounded [u] target. And this is, as far as the tongue is concerned, presumably because the genioglossus muscle has greater muscular force than the muscle system that raises and backs the tongue. Muscle system interactions of this sort in the balanced protagonist-antagonist situation in ordinary speech may well lie at the basis of the notion of "favoured articulatory zones" in the languages of the world. Laver concluded that we have here a very simple experimental paradigm of competition between two targets programmed in random sequence at high speed which can be applied in many areas of speech production and which can tell us perhaps a number of interesting things about the way speech is represented and controlled neuromuscularly.

Secondly, Laver had a comment about Ladefoged's suggested laryngeal parameters of glottal aperture, glottal tension, and glottal length. He pointed out that one aspect of the usefulness of this approach is that the six main modes of phonation - modal voice (Hollien's term), falsetto, creak, whisper, breathiness, and harshness - all have different specifications on these three parameters. And therefore, an explanatory basis is provided for the mutual compatibility or incompatibility between these six phonatory modes. It means that breathiness and harshness, for instance, are ruled out by that model as mutually incompatible, as they are in real life, because they need very different values on the glottal aperture and the glottal tension parameters.

Laver's third point concerned the habitual mode of phonation adopted by an individual which he found was an excellent example of a muscular setting (Honikman's term). The notion of a setting is extendable beyond the larynx to habitual adjustments of the supralaryngeal tract as well. We are all familiar with people using a particular long-term muscular adjustment of the supralaryngeal tract as part of their habitual voice quality. For example people who raise their larynxes and keep them raised throughout speech, people who have a tendency to maintain the lips protruded, qualities which characterize particular speech communities like velarization that one hears in the speech of Liverpool, and lastly habitual nasalization common among RP-speakers. The nice thing about muscular settings, in the context of MacNeilage's report, is that they furnish an excellent example of the Action Theory concept of co-ordinative structures, tuned to a long-term bias on segmental articulation - just like habitual gait.

The last point dealt with the problem of neuromuscular programming, when it is not just a matter of programming a sequence of segments as such, but rather of programming at least a triple layer of commands. Laver stated that if voice quality has a phonetic component which demands a particular controllable setting of the vocal tract and the larynx, then one has to take care of the neuromuscular programming for that component. Secondly, superimposed on that phonetic component of voice quality there will be the current tone of voice that the person is using, in other words the paralinguistic layer as well. And thirdly, the segmental and other components of the linguistic strand of speech. Laver concluded that neurolinguistic programming in real speech is at least three times more complex than would be needed for any single-layer control of segmental sequence.

M. Sawashima, responding to Ohala's last point, claimed that he did agree that the up and down movement of the larynx is highly correlated to the F_0 change. But Sawashima found it difficult to explain that the up and down movement of the larynx directly can affect the vocal fold tenseness if we consider the mechanical and structural properties of the larynx. Maybe we can explain it by saying that the up and down movement of the larynx indirectly can provide a change in the longitudinal tension of the vocal folds, which was said many years ago by Sonninen and others. Sawashima concluded that what we want to find is a reliable physiological correlate to the change or control of the vocal fold tension, and in that sense we can't say that the change of the vocal fold tension is caused by the up and down movement of the larynx.

S. Smith drew the attention of the audience to some of his works dealing with the functional dichotomy of the vocal folds (membrane-cushion, cover-body) and which were done before the works made by van Berg and Hirano.

P. Ladefoged presented a series of slides showing the laryngeal behaviour for different voice qualities in a Bushman language. In his co-report Ladefoged pointed out that the laryngeal parameters normally used are completely inadequate for a description of the six voice qualities found in this language. A very interesting finding was that the speakers of the language all had

a thickened interarytenoid muscle, which helps them to produce the ventricular phonation. The bulge seen on the interarytenoid muscle is not genetically controlled, because one of Ladefoged's colleagues has developed a thickened interarytenoid muscle, working with the language.

O. Fujimura had two points to make. The first one dealt with spoonerisms as evidence for the phoneme size segment as the functional unit. He pointed out that no unit whether phoneme, distinctive feature, syllable, or word can freely exchange with another unit in any environment. The facts are more complicated, and there are constraints and contextual conditions that have to be considered. Fujimura found that there is a confusion between the elements for exchange and the environment set up for the exchange of the elements, and he proposed to consider not only one unit for everything, the phoneme for instance, but also larger units as well. Typically, the exchange occurs in syllable initial position, and why is it so if the phoneme is really the functional unit for exchange?

The other comment concerned the vertical movement of the larynx, which Fujimura found is a very interesting phenomenological fact in correlation with pitch control. This is quite useful in finding out what the control signals are for "pitch control" in devoiced portions of speech. He referred to Japanese which has vowel devoicing according to certain contextual conditions. Fiberoptic observations have shown some vertical movement, qualitatively, in relation to the lexical accentual patterns and also to the phrase boundary phenomenon. In the case where the second syllable of the phrase is devoiced and should be high in pitch according to the general rules, some native speakers feel that the second syllable in those devoiced cases is low in pitch. And fiberoptic observations seem to support this feeling in terms of the vertical movement of the larynx.

N. Waterson had some comments concerning the question of babbling as preparation for speech. Early babbling or cooing usually begins spontaneously as a type of unstructured vocalization and is generally mainly vocalic in nature with perhaps a few sounds in the velar and uvular regions. This stage seems to be

non language-specific. But Waterson pointed out that the interactions between the baby and his caretakers play an important role in preparing the baby for linguistic communication.

The vocalic type of vocalization is replaced by more complex vocalizations containing various consonantal sounds, and they become structured and repetitive. This suggests that the baby is developing processing skills which enable him to recognize samenesses and differences in vocal stretches - something that is essential for the development of language. When structured babbling begins, mothers tend to imitate those stretches which seem to them to be similar to their own language, so the baby is encouraged to work on the sounds of the language of the environment. The child is thus prepared for the sounds he will use in his first words.

The protolanguage stage, which usually overlaps with babbling, is generally articulatorily much less complex than what has been achieved in babbling but represents the development of the functional use of vocalizations. When vocalization is first used functionally, the production is very simple as if articulatory complex production and functional use cannot be coped with by the child's processing system at the same time at this early stage. When he has learnt how to use simple vocalizations functionally, he is ready for the use of the more complex production of the actual speech, and the first words soon follow.

B. Lindblom pointed out that the interest in the biological basis of speech, brought up by MacNeilage, is an interest in the most general phonological universal of all, namely in the difficult topic of speech sounds being a subclass of all sounds and gestures. In this context Lindblom had a question for Ladefoged, Fujimura, and Perkell, which had to do with our articulatory modelling: "Why leave out the jaw?" Lindblom had earlier argued that with the aid of the notion of neutral tongue shape and the jaw parameter we can perhaps explain the origin of the distinctive feature open and close. Furthermore, Lindblom referred to some jaw data presented in his symposium report showing how consonants resist coarticulation in the environment of maximally open vowels. He found that this illuminates some of the phonetic background on phonotactic syllable structure, on strength hierarchies, and such abstract notions from phonology.

J.S. Perkell mentioned, responding to Lindblom, that the actual contribution of the structure of the jaw - i.e. the lower teeth - to directly determining the area function is minimal, but that the jaw serves more as a framework for carrying other articulators around and thereby has an indirect influence. Perkell pointed out that we can't answer the question concerning the importance of the jaw without including the jaw in our physiological models.

P. MacNeilage, replying to Fujimura, mentioned that what he really wanted to say was to stress the prominence of the segment assuming that the larger the number of areas that involve a unit the more important it is at a particular stage of the modelling process to which one thinks the areas are relevant. He agreed that one has to take into account many units in the modelling process and that contextual influences are extremely important.

Replying to Waterson, MacNeilage pointed out that by babbling he did not mean cooing but just what he liked to call the canonical form, the open-close alternation with time locking. He found that maybe he disagreed with Waterson about the onset of that stage. MacNeilage was of the opinion that it happens rather suddenly. It is an important point that has to be explored in the light of the role of imitation. If the adults imitate the child's forms but the child's initial forms occur suddenly, then imitation may have a rather minor role in the onset of the phenomenon, even if it may be important it its subsequent development.

MacNeilage concluded by saying that he was impressed with the lack of disagreement that there had been about the speech production aspect of the phonetic discipline. He liked to believe that it is a very healthy sign and that the heat of the argument is related to the state of the knowledge in this area.

P. Ladefoged returned to the problem dealing with the jaw. His evidence to say that one should leave out the jaw is that what is controlled is the vocal tract shape, referring to Lindblom and his colleagues, who have shown quite effectively that we can produce very similar shapes with the jaw in different positions. If we look at mathematical techniques for reducing the amount of variance between a group of speakers we come out with factors that

reflect the cavity shapes and do not reflect the jaw positions. This is another evidence that the jaw has no role to play. But Ladefoged pointed out that it is just so for vowels and that he might have to put the jaw back again for consonants, referring to Lindblom's new jaw data for consonants (cf. vol. II, p. 33-40).

H. Fujisaki mentioned that we have to treat the jaw as an independent motor unit when we are dealing with the dynamics of articulation. When Ladefoged speaks about tongue control it is a combination of independent or dependent control of the jaw and the tongue. The fact that one can produce many speech sounds without moving the jaw does not exclude the fact that the jaw plays an important role in articulation.

N. Waterson, responding to MacNeilage, replied that if he by "sudden" meant over two or three weeks then there was probably no disagreement, but if he meant from one day to another then they did disagree. But she pointed out that there is not quite enough data on babbling to be able to make a categorical statement about it.

MacNeilage admitted that he did not have very much data and that much of it was informal, but it was his impression that it happens virtually from one day to the next.

Fujimura advocated the independent function of the jaw. He referred to his tongue model, which actually includes an independent variable corresponding to the jaw angle. Fujimura found that the jaw has important functions particularly with respect to the inflection of stress patterns referring to some of his jaw data, which show that jaw height does not correlate clearly either positively or negatively with tongue height and it is not random either. He concluded that the jaw constitutes a very important articulatory dimension.

J. Ohala made a comment dealing with the interpretation of speech errors. He did agree with Fujimura's call for caution in the interpretation of speech error data for what they may reveal about units of speech production. He did this with an analogy.

Let us imagine the following domestic accident: a cook stores spices in a spice cabinet in alphabetic order, i.e., cumin is after corriander, and tumeric is after thyme, etc. In reaching for the thyme to add to a dish, the cook accidentally grabs tumeric instead, thus making a culinary analogue of a speech error. The analyst trying to interpret this error would look in vain for any chemical or physical similarity between tumeric and thyme. What is the point of this? Simply that for the purpose of retrieval or general "housekeeping" functions of manipulating the stored units of speech, it is possible that the addresses or labels used bear only an arbitrary relationship to the substance of the units themselves. Ohala concluded that until we have some general idea of how speech is "programmed" he did not think that the data from speech errors can unambiguously rule out features, phonemes, or syllables - or something else - as possible units of production.