Control and representations in speech production

Pascal Perrier
Institut de la Communication Parlée – UMR CNRS 5009
INP Grenoble & Université Stendhal Grenoble

In this talk, two aspects of the concept of *representations in speech* will be addressed. In the first part, classical hypotheses that have inspired speech production and speech perception research for the last 20 years will be presented. These hypotheses have aimed at clarifying, how the objectives of the speech production task are represented in the speaker's mind. The second part will focus on the nature of the representations that speakers might have of their physical speech apparatus, in order to elaborate the motor strategies permitting the achievement of their speech objectives.

Speech production is by essence a semiotic task: its ultimate objective is the exchange of information with listeners sharing the same linguistic code. At the same time, speech production is a motor task aimed at producing movements and sounds. Trying to understand what the representations of the speech production task in the speaker's mind are requires understanding the link between the abstract semiotic aspects of the task and its motor, physical and perceptual characteristics and suggesting how speakers transpose the abstract objectives into the physical and more proximal domains. This is not a simple task since the link between the different domains in which speech production can be analyzed is far from bi-univocal. Indeed, to produce a phoneme, a speaker has the choice between a large number of muscle recruitments, articulatory positions and even acoustical productions: what are the domains in which the task is specified and how? To address this issue, the talk will first present the pioneering works by Stevens (Acoustical Invariance), Liberman & Mattingly (Motor Theory of Speech Perception), Fowler (Direct-realist Theory of Speech Perception), and Lindblom (Adaptive Variability), all of whom have elaborated strong and controversial hypotheses about the physical correlates of speech, considering it in a perception-action framework. These hypotheses will then be discussed in light of recent contributions based on perturbation experiments or on modeling studies (Guenther *et al.*, Max *et al.*, Munhall *et al.*, Ostry *et al.,* Perkell *et al*., Savariaux *et al.*, Sussman *et al*.).

To achieve these objectives, speakers have to activate a large number of muscles in order to control their vocal source and vocal tract articulators with enough accuracy. The physical characteristics of the speech articulators are quite complex and various: they can be bony or soft tissues, they can move under the influence of external or internal muscles, and their masses can be very different from each other. In addition, speech movements are fast (transitions between phonemes can be as short as a few tenths of a second, which does not permit an ongoing cortical monitoring of the linguistic task. Hence, understanding how speakers manage this complexity and achieve the required accuracy is a very important issue for speech motor control. A classical hypothesis assumes that the brain builds up *representations* of the physical speech production apparatus, the so-called *internal models.* Many of the above-mentioned experimental works have also contributed to addressing this issue. It will be shown that the large majority of these studies support the existence of internal representations. The question of the degree of complexity of these internal representations (do they include static, kinematic or dynamic descriptions?) will also be briefly discussed on the basis of modeling and experimental works in the speech production domain as well as in the motor control domain (Kawato *et al.,* Ostry *et al.*, Perrier *et al.*) **.**