Available online at www.sciencedirect.com

SCIENCE DIRECT®

Speech Communication 45 (2005) 425–434

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

Asymmetries in vowel perception, in the context of the Dispersion–Focalisation Theory

Jean-Luc Schwartz [a,*], Christian Abry [a], Louis-Jean Boë [a], Lucie Ménard [b], Nathalie Vallée [a]

[a] *Institut de la Communication Parlée, CNRS-INPG-Université Stendhal, 46 Av. Felix Viallet, 38031 Grenoble Cedex, France*

[b] *Département de linguistique et de didactique des langues, Université du Québec à Montréal, Montréal, Canada*

Received 21 January 2004; received in revised form 1 December 2004; accepted 7 December 2004

Abstract

In a recent paper in this journal, Polka and Bohn [Polka, L., Bohn, O.-S., 2003. Asymmetries in vowel perception. Speech Communication 41, 221–231] display a robust asymmetry effect in vowel discrimination, present in infants as well as adults. They interpret this effect as a preference for peripheral vowels, providing an anchor for comparison. We discuss their data in the framework of the Dispersion–Focalisation Theory of vowel systems. We show that focalisation, that is the convergence between two consecutive formants in a vowel spectrum, is likely to provide the ground for anchor vowels, by increasing their perceptual salience. This enables to explain why [y] is an anchor vowel, as well as [i], [a] or [u]. Furthermore, we relate the asymmetry data to an old experiment we had done on the discrimination of focal vs. non-focal vowels. Altogether, it appears that focal vowels, more salient in perception, provide both a stable percept and a reference for comparison and categorisation.

Keywords: Vowel discrimination; Peripheral vowels; Focalisation; Salience; Stability; Anchor

1. Introduction: Global vs. local constraints in substance-based theories of phonological systems

In the search for substance-based principles shaping phonological systems, both global and local constraints have been considered in the literature. *Global* constraints are based on the relations between elements in a system, so that a given gesture/sound/percept is included in the system not because of its own properties, but because of its contribution to a global function in relation with the other elements of the system. A prototypical example is provided in Lindblom's Dispersion Theory DT (Liljencrants and Lindblom, 1972;

* Corresponding author. Tel.: +33 4 76 57 47 12; fax: +33 4 76 57 47 10.

E-mail address: schwartz@icp.inpg.fr (J.-L. Schwartz).

0167-6393/$ - see front matter © 2005 Published by Elsevier B.V.

doi:10.1016/j.specom.2004.12.001

Lindblom, 1986), later becoming the Theory of Adaptive Variability TAV (Lindblom, 1990; and a revised version in Diehl et al., 2003). In this framework, distinctiveness of the units within the system is the driving force, and hence units must be as different as possible one from another, in perceptual terms. Therefore, "a selected unit is highly valued, not because of its individual qualities, but depending on its contribution as a team player" (Lindblom, 2003). On the contrary, *local* constraints should result in focusing the selection towards specific regions in the articulatory-acoustic-perceptual space, preferred for intrinsic properties, independent of the properties and configurations of the other elements in the set. This provides the basic rationale for Stevens' Quantal Theory QT (Stevens, 1972, 1989) in which non-linearities in the articulatory-to-acoustic or acoustic-to-auditory transforms define natural boundaries that would be exploited by phonological contrasts.

If such local attraction regions do exist, and in some sense "pre-exist" to linguistic systems, then it should be possible to display their existence through adequate non-linguistic or pre-linguistic experimental paradigms. This line of reasoning has produced some striking successes, particularly concerning two consonantal contrasts, that is place of articulation and voicing for plosives. Just to mention the second one, the existence of a "natural" boundary between unvoiced and voiced plosives has received support from VOT (Voice Onset Time) categorical experiments. These experiments involved either animals (Kuhl and Miller, 1978) or prelingual infants (Eimas et al., 1971). Both experiments displayed categorical perception with increased discrimination around the boundary between voiced and unvoiced plosives, though language was not (for animals) or not yet (for infants) present. The same kind of results was obtained with a non-linguistic continuum "mirroring" the linguistic one (TOT, Tone Onset Time: Miller et al., 1976).

The situation is not so clear for vowels, for which categorical perception does not seem to apply (Repp, 1984). Discrimination experiments were used by Kuhl to introduce the "perceptual magnet effect", according to which some regions of the acoustico-perceptual space could provide anchor points for categorisation (called "magnets"), both for adults and 6-month-old infants... though not for animals (Kuhl, 1991). But it appeared later that these regions were in fact the product of a *learning phase* from 0 to 6 months old. Indeed, different magnet regions were observed for 6-month-old infants of different languages, and these regions were related to adult prototypes in the corresponding language (Kuhl et al., 1992). Therefore, the magnet effect characterizes the tuning to a specific language under exposition, rather than universal local constraints on vowel systems.

A few years ago, we introduced a new theory for the prediction of vowel systems, integrating global and local peripheral constraints on the shaping of vowel inventories. This theory, called the Dispersion–Focalisation Theory (DFT), includes both the global dispersion ingredient exploited by Lindblom and colleagues, and an additional local property called focalisation, related to preferred regions in the perceptual space (Schwartz et al., 1997a; see Section 2.2). The DFT was shown to predict quite well the major characteristics of existing vowel inventories in the world languages (Schwartz et al., 1997b; Vallée et al., 1999).

In this context, we read with enthusiasm a recent paper in this journal by Polka and Bohn (2003) in which they summarise in a clear and striking way a series of experimental data from themselves and others, consistently showing the existence of asymmetries in vowel perception (see Section 2.1). The authors interpret these asymmetries as a predisposition for more peripheral vowels in the F1 – F2 space that could provide a perceptual anchor for vowel systems. These regions could in fact, in our view, be better described in the framework of the DFT, and seem to provide an interesting argument in favour of the theory, and particularly its "focalisation" component. The purpose of the present paper is hence to propose a reinterpretation of the paper by Polka and Bohn within the DFT. Their data, together with the DFT sketch, will be briefly recalled in Section 2. Reinterpretation will be done in Section 3, before a conclusion section.