Concurrency, Synchrony, and Temporal Organization

Eric Vatikiotis-Bateson

Department of Linguistics, University of British Columbia, Vancouver, Canada

evb@interchange.ubc.ca

Abstract

The organization of time-varying linguistic behavior, while controlled, is not precisely timed. This claim is supported empirically, but it is also motivated theoretically by an idea that will be fleshed out in the talk; namely, language behavior necessitates a neuro-cognitive displacement from the fine-grained spatial and temporal instantiation of language. This obviates the need for the precise synchronization that can be observed in that other peculiarly human activity – music.

Index Terms: concurrency, synchrony, multimodal coordination, instantaneous correspondence algorithm, language, music

1. Overview

In this talk, I attempt a critical look at the role of temporal organization in planning, producing, and perceiving spoken language behavior. Many of us, I am sure, believe that spoken language is highly constrained in its temporal patterning. After all, language exhibits rhythm and other characteristics that can be seen only through their repetitive patterning in time, which implies a certain stability. Furthermore, even if some of us have not given it much thought, a moment's attention to how speech is produced suggests the highly coordinated (including carefully-timed) activity of many physical structures (e.g., individual speech articulators) and sub-systems (e.g., respiration, phonation). However, stability and coordination do not necessarily require precision in either space or time. Indeed, I argue that, whether examined within or across modalities, coordinated behavior crucially must *not* be precisely timed or synchronized in order for it to be linguistically-relevant. Temporal coordination is argued to combine both emergent and controlled properties making the description of spatial and temporal coordination more a matter of proximity and overlap, which I term *concurrency*, than of strict synchronization.

The theoretical position that linguistic behavior necessarily entails concurrent rather than synchronized behavioral components will be motivated by a reinterpretation of some of my older research on articulatory dynamics (e.g., [1]) and speech motor control (see [2]), and additional empirical support from four types of study: i) multimodal studies examining the production and perception of declination and speech planning [3], and of Lombard Speech [4]; ii) combined linguistic and performance (kinematic and acoustic) analyses of story-telling in Shona [5]and Plains Cree [6]; iii) the effects of vocal effort during speech and operatic singing on postural and articulatory coordination; and, for contrast, iv) performance analyses of singer-audience interaction.

Key to this reinterpretation is an algorithm, recently completed and presented at this conference, that computes the instantaneous correspondence between signals across a range of temporal offsets that can be visualized as a time-varying function [7, 8, 9]. Combined with simple optical flow techniques for computing pixel motion, we clearly see the ubiquity of concurrent behavior within and across modalities, unimpeded by the noise of too much spatial and temporal resolution.

2. References

[1] E. Vatikiotis-Bateson and J. Kelso, "Rhythm type and articulatory dynamics in english, french, and japanese," *Journal of Phonetics*, vol. 21, no. 231-265, 1993.

[2] K. G. Munhall, M. Kawato, and E. Vatikiotis-Bateson, "Coarticulation and physical models of the speech production," in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, M. B.Broe and J. B. Pierrehumbert, Eds. Cambridge: Cambridge Univ. Press, 2000, pp. 9–28, invited Presentation, 7/96, Evanston Ill., Labphon 5.

[3] J. Tan, J. Dunham, C. Lee, and E. Vatikiotis-Bateson, "Initial pitch informs sentence duration," in *Journal of the Acoustical Society of America*, vol. 119, 2006, p. 3393.

[4] E. Vatikiotis-Bateson, A. V. Barbosa, C. Y. Chow, J. Tan, and H. C. Yehia, "Audiovisual lombard speech: Reconciling production and perception," in *International Conference on Auditory-Visual Speech Processing AVSP 2007*, J. Vroomen, M. Swerts, and E. Krahmer, Eds. The Netherlands: ISCA, 2007, pp. 45–50.

[5] R.-M. Dechaine, A. V. Barbosa, C. Mudzingwa, and E. Vatikiotis-Bateson, "Performance constraints in shona narratives," 14-17 April, 2008 2008.

[6] R.-M. Dechaine, J. Deschamps, A. Barbosa, C. Cook, J. Muehlbauer, T. Cardinal, J. Small, and E. Vatkiotis-Bateson, "Language as performance: multimodal communication in plains cree," 26-29 May 2007.

[7] A. V. Barbosa, H. C. Yehia, and E. Vatikiotis-Bateson, "Temporal characterization of auditory-visual coupling of speech," pp. EL1–13, 2008.

[8] ——, "Algorithm for computing spatiotemporal coordination," in *Auditory and Visual Speech Processing – AVSP08*, R. Goecke and P. Lucey, Eds. Moreton Island, Queensland, Australia: ESCA, 2008.

[9] ——, "Linguistically valid movement behavior measured non-invasively," in *Auditory and Visual Speech Processing – AVSP08*, R. Goecke and P. Lucey, Eds. Moreton Island, Queensland, Australia: ESCA, 2008.