Conversational Prosodic Interactivity When One Partner has Aphasia

Eugene H. Buder and Jamie L. Edrington

School of Audiology and Speech-Language Pathology
The University of Memphis, USA
{ehbuder; jledrngt}@memphis.edu

Abstract

The purpose of this study is to document prosodic interactivity, and its preservation in a woman with non-fluent aphasia. Fundamental frequency and sound pressure level records are examined for sinusoidal models. In-phase and anti-phase partner relationships in these models may represent synchrony. These measurements demonstrate that prosodic management of conversational rhythms is robust even when expressive language is disturbed.

1. Introduction

Speech behaviors are naturally cyclic at many levels [1], especially, it seems, when conducted by partners interacting synchronously in conversations. Some conversation analysts observe a markedness state for "rhythmic continuity" across the partners' contributions [2], but operational representations of rhythm, especially in casual speech, remain elusive.

Cyclicity in slowly-varying acoustic data such as fundamental frequency ("f_0", heard primarily as pitch) and sound pressure level ("SPL", heard primarily as loudness) can be determined using methods of statistical time-series analysis [3]. Frequency-domain analysis modeling of such records can be powerful guidance towards simple sine wave models of these prosodic variables [4].

Prior work on three speakers with fluent aphasia used cyclicity tools to analyze acoustic records of conversational materials, finding that alterations in rhythmic continuity from the unmarked continuous state to a break in rhythm were as observed by Couper-Kuhlen for unaffected individuals (based on differences in the initiation of the repair itself [2]). Speakers with fluent aphasia, however, prosodically managed such episodes according to partner familiarity and not the nature of the repair [5].

Accordingly, this study is designed to examine conversational speech behaviors of an individual with a language disorder–expressive language difficulties associated with non-fluent aphasia–with a familiar partner, with an unfamiliar partner, and between the two unaffected individuals in this triad. On face value, it would seem that these language difficulties might disrupt the natural flow of conversational interaction due to prosodic difficulties [5, 11]. However, it may be that an effective partner can help to induce fluency.

2. Method

The methodological details of the recording and analysis procedures utilized in this work are too numerous to be fully documented in this report, but an overview is provided here and the methods are also illustrated in sample results. The procedures are similar to those used in previous reports [6, 7] but some enhancements have been developed, which are emphasized here.

2.1. Participants

The central participant of interest in this study was a female individual ("PL"), recruited from a speech and hearing clinic located in Memphis, Tennessee, where she was receiving language therapy following a diagnosis of Broca's aphasia which resulted from a cardiovascular accident (CVA). Because it involves word formulation, Broca's aphasia is a type of "non-fluent" left hemisphere aphasia that usually results in prosodic difficulties [8, 9]. At the time of the study, she was 55 years of age and was 11 months post onset. Her scores on *The Western Aphasia Battery* [10] indicated expressive language deficits, characterized by frequent paraphasias; perseverations; word retrieval deficits; halting and hesitant speech; and a lack of grammatical completeness with reduced phrase length. PL is an African-American whose speech and language in the conversations were in the style of African-American Vernacular English [11].

PL was asked to recruit and bring with her to our recording session a person with whom she felt very comfortable speaking. She selected her niece, a young African-American woman ("N"). Another young African-American woman who worked at the speech and hearing clinic, but who was not a speech-language pathologist, was recruited to participate as a person who had never engaged in any extended conversation with either PL or N, and so was dubbed a stranger ("S").

2.2. Recordings

Participants wore several acoustic and physiological transducers: (1) high quality lavalier microphones (Sony ECM150) mounted approximately 10 cm below the mouth, (2) contact microphones (K&K Sound Hotspot) adhered to the neck wall somewhat superior and lateral to the vocal folds, in the vicinity of the thyrohyoid membrane, and (3) variable inductance bands (Inductotrace) worn beneath outer garments for monitoring rib cage and abdominal wall movements. Video was also captured in both split-screen face and upper body views and in a side view. All data were captured digitally via 8-channel Data Translation equipment (DT322) or video capture devices; the audio signals were low-pass filtered for anti-aliasing prior to digitization at 20 kHz for radiated signals and at 10 kHz for contact signals. Only the acoustic signals are considered in this report.

Participants were seated (two at a time) in a room approximately 160 square feet in size residentially furnished. For each conversation, lasting approximately 20 minutes, the conversationalists were given no specified topic. Three conversations were recorded: first, between PL and N, second, between PL and S, and third between S and N. All conversations have been analyzed using the methods illustrated here for the conversation between PL (the woman with non-fluent aphasia) and N.

2.3. Extraction of data

Digital audio recordings were loaded into a custom version of TF32 [12] provided in the form of an ActiveX library. This version provided several features not yet available in the commercial release: (1) analysis parameters for pitch extraction for adapting to contact microphone signals, (2) features for editing Root Mean Square (RMS) energy contours, and (3) a facility for extracting time-series format data for both RMS and fundamental frequency (f_0) data.

For the present purposes, a one minute sample, representing the most extensive speech from both participants, was selected from each conversation. Radiated signals were used for RMS energy analysis. Editing of the RMS contours was applied to delete cross-talk, using the level of the noise background as a floor value for each participant's signal. In the subsequent conversion of the RMS voltage measures to Sound Pressure Level (SPL) on a decibel scale, each floor voltage was set to 30 dB, therefore equilibrating each participant's SPL data and providing a reasonable approximation of dB SPL.

Contact microphone signals were used for f_0 extraction in a three-stage procedure designed to maximize validity and replicability: (1) "global" analyzing parameter values were determined for a best overall fit to a given participants' session; (2) parameters were adjusted to accommodate local segments that deviated from overall norms (e.g., high or rapidly changing f_0 patterns, as in laughter); and (3) hand-editing, such as interpolations, was used for passages where the algorithm failed despite parameter adjustments (e.g., during segments with relatively aperiodic voicing).

Data from the TF32 interface was downsampled at 240 Hz and then processed in MATLAB [13] for conversion of RMS voltages to dB and then further downsampled using median-smoothing algorithms for two alternative modes of analysis: 16 sample-per-second (sps) data and 48 sps data.

Table 1: *Transcript of sample in Figure 1 ("[]" indicates overlap).*

PL:	I said, "Boy, you done had you [boy"]
N:	[That's] probably why she going to exercise and stuff. She's been uh going to work out 'cause she says [she picked up some weight. She. . .eating,]
PL:	[how she get like that? How she did that?]
N:	eating
PL:	It'll do you like that? (accompanied by laughter)
N:	yeah
PL:	as little as she is
N:	mmhmm
PL:	She was so little.
N:	I know she was.
PL:	What happened?
N:	eating, putting food in your mouth
BOTH:	Laughter ensues

2.4. Time-series analysis for identification of prosodic synchrony in conversational interaction

SPL and f_0 data were spectrally analyzed in MATLAB following general procedures for Fast Fourier Transform (FFT) analysis [4, 14]: 16 sps data were analyzed in 128 point transforms in 8 second (s) frames and 48 sps data were analyzed in 128 point transforms for 2 ⅔ s long frames. Frames are shifted through the data with 48 point (⅓ frame) overlap, and framed data were cosine tapered prior to FFT analysis. The illustrations in Figure 1 utilize 16 sps data. To guide inspection for dyadic-level phenomena, two bivariate representations of the dyadic data are also produced for each frame: (1) a cross-spectrum representing frequency-specific the selection of models that indicate significant relationships shared power between partners' data, and (2) a frequency-specific phase plot that maps dyadic phase relationships onto a scale from 0 to 1, 0 representing 0° or 360° relationship and 1 representing a 180° relationship. Either relationship (in phase or anti-phase) may be indicative of a "synchronous" relationship, depending on the scale and/or mode of conversational interaction (cf. the emerging interdisciplinary science of coordination dynamics [15]). We acknowledge Dr. Lesya B. Chorna for developing the MATLAB applications.

Prosodic synchrony analysis of the conversational interactions is guided by inspection of the univariate and bivariate spectral results in conjunction with the fitting of sinusoidal models to the original prosodic records, as indicated by prominent spectral components. An example selected from the conversation between PL and her familiar partner N is presented in the next section, with Table 1 containing a line-by-line transcript of the passage.

3. Sample Result

Figure 1 presents a series of wave models that illustrate a high degree of conversational prosodic interactivity, even though one of the partners has been diagnosed with non-fluent aphasia. The most significant episode in this sample involves an anti-phase relationship between SPL cycles; the wave models of this episode were guided by the spectral results displayed in Figure 2. All the models depicted in Figure 1 were derived from 8 s frames of 16 sps data.

The panels of Figure 2 contain components that guided the SPL wave models of Figure 1. It is evident from the remarkably similar univariate spectra that both partners share cyclicity of fluctuations of SPL, particularly in the vicinity of 0.5 Hz. The cross-spectral panel verifies that the highest degree of shared energy is at exactly 0.625 Hz, which corresponds to a wave period of 1.6 s. Indicative of prosodic synchrony, the phase plot shows that the cycles in this frequency region are consistently in a very clear anti-phase relationship. Note that the waves also correspond to a phrase rhythm (see Table 1) in which the partners trade queries and affirmations culminating in laughter.

In the f_0 domain, there is also a 0.625 Hz peak in PL's spectrum (no f_0 spectra shown); the wave model at this frequency depicted in Figure 1 clearly corresponds to the phrase rhythms of partners' quick turns, but is shifted in phase relative to SPL cycles. This is because the most "power," in the phrase-final f_0 drops, coincides with the amplitude final margin. Finally, observe that prior to the quick phrase exchange pattern, there is also a slow fall-rise, modeled for B but which partners clearly share, culminating in the simultaneous termination of extended overlap.

4. Discussion

We have demonstrated that signal processing tools can be applied to detect rhythmic interactivity in the prosodic domain, and interactive prosodic rhythms can be robust, even in the

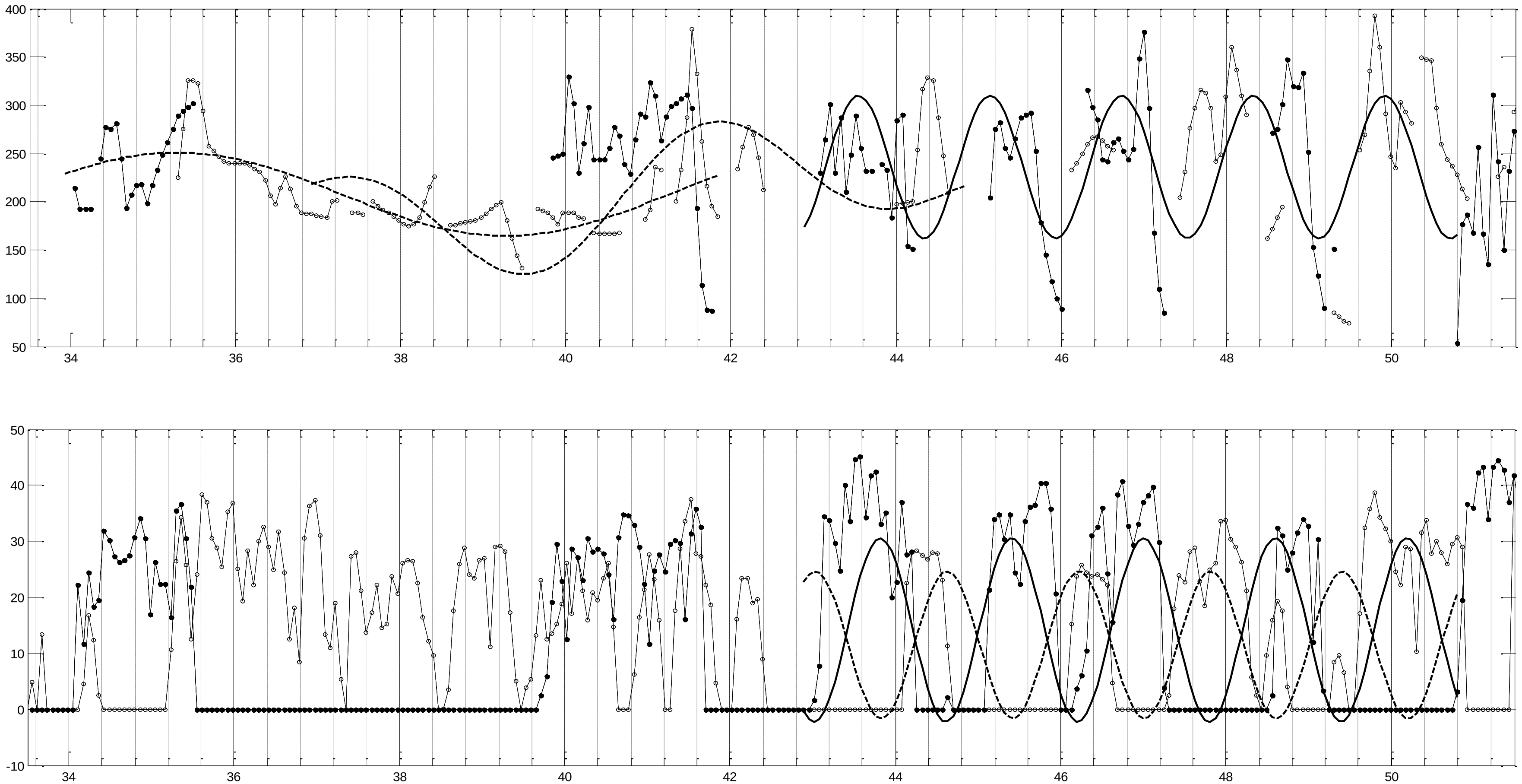

Figure 1: *A sample of prosodic data from the conversation between PL (a woman with aphasia) and N (her niece). The top panel represents fundamental frequency (f_0) data, in Hz, and the lower panel represents SPL data, in dB, (relative to the ambient noise floor of the recording environment). Pl's data points are represented with filled points and solid lines, and N's points with unfilled points and dashed lines. All models are based on 8 s frames which had been shifted by 3 s increments through the data.*

Comments: The first single long sine wave to the left models N's explanations that a mutual acquaintance was not pregnant. The long arc modeled by this frame, along with the same frequency in the next frame (summed with another incrementally higher frequency for a shared fit), bridges two occurrences of simultaneous speech. The overall segment is heard to be driven by N. At time 43, however, PL clearly take a leading role with a series of sharp questions; this is seen in the phase advancements of PL's SPL models relative to N (against which she also has greater overall SPL). The series of exclamations by PL also seen in an f_0 model, associated with f_0 declinations (though with a different phase alignment than in SPL). These lead to a request "what happened?". To take the floor for a humorous answer, N assumes the SPL cycle initiated by PL.

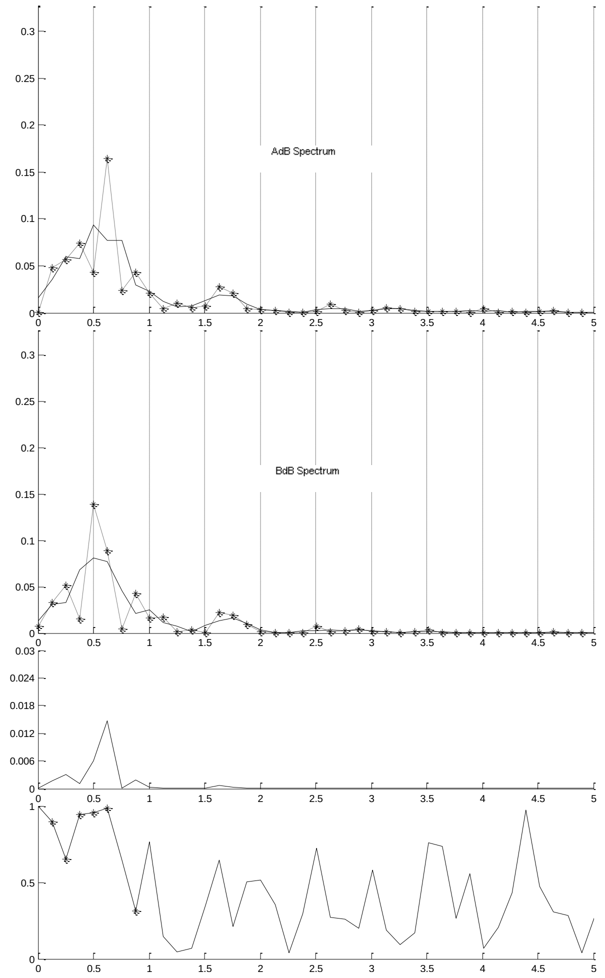

Figure 2: Frequency-domain results: All panels are on the same abscissa (Hz, or cycles per second). The top two panels are univariate FFT results (raw periodogram output in dashed lines and 3-point smoothed power estimates in solid lines).Variance in PL's SPL data is depicted as a function of frequency for PL (top) and N (next down). The third panel is a cross-spectrum, depicting covariance between partners' SPL data. The lowest panel depicts partners' phase relationships with substantial levels of cross-spectral energy coded by asterisks.

face of expressive language disturbance.

Individuals with Broca's aphasia suffer from word-finding difficulties and typically display prosodic difficulties when speaking [8, 9, 16]. However, very few prosodic difficulties were noted when listening to PL engage in conversations with N. The wave modeling results presented here substantiated our auditory impressions of the conversation between PL and N. Furthermore, comparisons of prosodic wave models for the conversation between PL and her familiar niece N to those for PL and the unfamiliar stranger S have suggested no difference between PL and the other women in terms of prosodic interactivity.

Interaction seemed to help induce fluency in this person with aphasia [16]. In fact, using a conversational setting to improve the social and expressive language skills of an individual with aphasia is one of the major tenets of a therapy program known as Promoting Aphasics' Communicative Effectiveness [16]. Despite the presence of expressive language deficits, PL successfully engaged in conversations where she was able to match her conversational partner's prosodic rhythms.

References

[1] E. H. Buder, "Dynamics of speech processes in dyadic interaction," in *Dynamic patterns in communication processes*, C. A. Vanlear, Ed. Thousand Oaks, CA.: Sage, 1996, pp. 301-325.

[2] E. Couper-Kuhlen, *English speech rhythm: Form and function in everyday verbal interaction*. Amsterdam: John Benjamins, 1993.

[3] J. M. Gottman, *Time-series analysis: A comprehensive introduction for social scientists*. Cambridge: Cambridge University Press, 1981.

[4] R. M. Warner, *Spectral analysis of time-series data*. New York: The Guilford Press, 1998.

[5] R. A. Dressler, "Rhythmic patterns during conversational repairs in speakers with aphasia (Doctoral dissertation, The University of Memphis, 2004)." *Dissertation Abstracts International*, vol. DAI-B 65/11, p. 5673, 2004.

[6] E. H. Buder and A. Eriksson, "Prosodic cycles and interpersonal synchrony in American English and Swedish," in *Eurospeech 1997 Proceedings*. vol. 1, E. Dermatas, Ed. Grenoble, France: European Speech Communication Association, 1997, pp. 235-238.

[7] E. H. Buder and A. Eriksson, "Time-series analysis of conversational prosody for the identification of rhythmic units," in *Proceedings of the 14th International Congress of Phonetic Sciences*. vol. 2, A. C. Baile, Ed. San Francisco, CA: University of California, Berkley, 1999, pp. 1071-1074.

[8] K. P. Kearns, "Broca's aphasia," in *Aphasia and related neurogenic language disorders*, 2nd ed, L. L. LaPointe, Ed. New York, NY: Thieme, 1997, pp. 1-41.

[9] J. C. Payne, *Adult neurogenic language disorders: Assessment and treatment*. San Diego: Singular Publishing Group, Inc., 1997.

[10] A. Kertesz, *The western aphasia battery*. New York, NY: Grune and Stratton, 1982.

[11] A. McCabe and L. Bliss, *Patterns of narrative discourse: A multicultural, life span approach*. Boston, MA: Allyn & Bacon, 2003.

[12] P. Milenkovic, "TF32," Madison, WI: University of Wisconsin- Madison, 2001.

[13] MathWorks, "MATLAB," R2007a ed Natick, MA: The MathWorks, Inc., 2007.

[14] P. Bloomfield, *Fourier analysis of time series: An introduction*, 2nd ed. New York: John Wiley & Sons, 2000.

[15] J. A. S. Kelso and D. A. Engstrøm, *The complementary nature*. Cambridge, MA: MIT Press, 2006.

[16] J. C. Rosenbek, L. L. LaPointe, and R. T. Wertz, *Aphasia: A clinical approach*. Boston, MA: College-Hill Publication, 1989.