A MULTI-PERSPECTIVE APPROACH TO THE DYNAMICS OF REAL-TIME PROSODY

Ruth Huntley Bahr, Elaine R. Silliman, Scott L. Silliman*, and Steven Everling
University of South Florida, Tampa, FL, USA, *University Health Science Center, University of Florida, Jacksonville, FL, USA

ABSTRACT

Much of what is known about prosody derives from clinical studies of adults with hemispheric lesions. Moreover, prosodic abnormalities tend to be interpreted with little attention to speech planning difficulties. This investigation describes a model of discourse and speech planning that utilizes an integrated methodology, incorporating neuroanatomical discourse, and acoustic-physiological domains. Videotaped samples of naturalistic spoken discourse patterns and more constrained speaking tasks were acoustically analyzed to identify correlates of prosody. Participants were two men who had similar discrete infarcts to the left inferior precentral gyrus and two otherwise healthy, age-matched peers. Discourse samples were analyzed within three types of encoding units: clausal, planning a speaker is engaged in, social factors, including who are the conversational participants, the nature of the task, and the communication goals to be accomplished, appear to influence the encoding units accessed (4, 12, 13). Fourth, turning to prosody, in the Levelt (5, 14-16) model of speech planning, distinctions between affective and non-affective prosody are not inherently categorical. Prosodic information resides at the level of phonological encoding, where an “articulatory or phonetic shape for all words and for the utterance as a whole” is generated (16, p. 91). Metrical information, as well as phrasal syntactic and affective prosody, while both higher level functions, are fed into a proposed prosody generator. This mechanism establishes temporal parameters for each successive syllable frame in terms of duration, loudness, and pitch movements prior to the formulation of the phonetic plan. The output at this level is described as an “intonational phrase,” a unit that descends from clausal structure, but is smaller than the clause. This unit functions to mark an intonational break, which is partially a pragmatic device “under the speaker’s intentional control” (5, p. 385) since the decision to break is an option that creates the intonational phrase. Thus, in the Levelt description, boundaries between affective and non-affective prosody appear to overlap from a functional perspective.

1.2. Problems in studying prosody. The present study derives from five problems connected with the clinical diagnosis and study of disruptions in the production of non-affective prosody in the left hemisphere: a) the absence of “theoretically grounded account(s) of...how the brain represents and processes language (or speech)” (17, p. 342); b) diagnostic procedures dependent on perceptual assessment of abnormality which often results in vague or imprecise descriptions; c) the use of speaking tasks that involve sentence repetition and oral reading of sentences which do not access on-line discourse or speech planning; d) the lack of information on the range of normal variations in the natural uses of non-affective prosody (2, 18); and e) lack of consensus about the functional units of prosody.

2. REVIEW OF NON-AFFECTIVE PROSODY

2.1. Definition. Most data about the boundaries between normal and atypical prosody derive from clinical studies on the production of non-affective prosody in patients with either left or right hemisphere infarcts or in neurologically healthy adults (e.g., 1, 17-21). In regard to left hemisphere damage due to ischemic etiology and based primarily on autopsy data, only two studies have clearly identified discrete lesions associated with aphemia (22, 23 [patient #2]). Key features that appear to define aphemia, in contrast to Broca’s aphasia, dysarthria, apraxia, or foreign accent syndrome, are: (a) involvement of the inferior portion of the left precentral gyrus, (b) intact language comprehension in the oral and print domains, (c) a slowed rate of speaking, and (d) prosodic abnormalities. Right hemisphere damage following ischemic infarction that results in difficulty modulating affective prosody appears associated with the right inferior frontal and anterior-inferior parietal regions (24). Although aphasia is not common, problems with inferencing and planning may be present because of the frontal lobe involvement (25, 26). Pertinent findings from the right hemisphere research...
on production suggest that variations in fundamental frequency are essential to the signaling of affective meanings (1, 21, 27) and that both cerebral hemispheres may contribute to the non-affective prosodic aspects of language expression (e.g., 21). Moreover, Positron Emission Tomography (PET) data on patterns of pitch perception in a tonal language (Thai) versus English show that the acoustic properties of prosody may not be as relevant for their encoding, relative to hemispheric specialization, as is functional experience with a particular language (28).

2.2. Difficulties in description. In general, there are two problems hampering understanding of the functional-anatomical contributions of the left and right hemispheres to the production of non-affective and affective prosody. The first pertains to sample selection. There is wide variability in defining lesion size and sites in individuals with left anterior infarcts (29). The second involves differences in procedures (30). The tasks selected commonly involve highly constrained speaking activities, such as isolated sentence or word reproduction, which do not require planning comparable to natural discourse and speech production. Second, variations in speaking tasks and their level of cognitive demand, as well as features of the social context, can also lead to tradeoffs in the allocation of attentional resources to speech planning versus discourse production (6, 7, 31, 32). As a consequence, greater variability in speaking rates, pausal patterns, and articular precision may be noted. Moreover, for clinical studies in the past 13 years, when acoustic parameters of prosodic production have been instrumentally analyzed (e.g., 21, 27, 30, 33-38), these measurements have been interpreted as single dimensions independently of the discourse or speech planning situation.

2.3. Hesitation phenomenon. A different picture of prosodic production may emerge when interactions among the discourse and speech planning systems are considered (37, 39, 40). Discourse planning has primarily been studied through the prism of hesitation phenomenon research. The basic premise is that cognitively and socially difficult tasks influence speakers to pause more as information processing demands increase (13, 41), while highly familiar, repetitive, or memorized content contains less hesitations and results in less pausing (4, 42, 43). Moreover, the discourse production process evolves in a rhythmic, or cyclical, manner (6, 7, 31, 41). In encoding a chunk of speaking, a planning phase characterized by more intervals of pausing alternates with an execution phase, where phonological encoding takes place for the transformation of discourse and linguistic elements into speech (41). A point of unresolved contention concerns whether the unit of discourse planning is semantic, syntactic (clausal), supraclausal, or, even, prosodic.

It is the purpose of this project to: (1) Contrast the information obtained through clausal, prosodic, and supraclausal methods of investigating prosody; (2) Investigate the role of acoustic measurement in the determination of prosodic patterns in discourse that might correspond with a disruption in non-affective prosody; and (3) Compare variations in prosodic correlates as obtained from three discourse samples with those obtained from three traditional sentence level tasks.

3. METHOD

3.1. Participants

3.1.1. Clinical Cases. Two patients, HC, age 57 years, and AI, age 50 years, were studied. Each patient presented with prosodic and other speech disturbances due to brain infarctions involving the precentral gyrus. Both patients underwent brain MRI examinations. Localization of their infarctions to the inferior portion of the left precentral gyrus was verified by the mapping technique of Sobel et al. (44).

An MRI performed on HC the second day after his stroke showed that the upper two thirds of the inferior portion of the left precentral gyrus was damaged. He exhibited severe disruptions in prosody and articulation that were typical of traditional definitions of aphemia. HC was then seen again at 57 days poststroke for administration of discourse and speech production tasks by a neurology resident.

With AI, MRI findings on the third day poststroke showed an infarction isolated to the middle third of the inferior section of the left precentral gyrus. AI presented with a disturbance of nonaffective prosody. AI was first videotaped two days poststroke and then videotaped at 55 days, 62 days, and 117 days poststroke by the attending neurologist.

3.1.2. Controls. Two men, matched by age and educational level, served as control subjects. Their performances on the discourse and speech tasks would serve as an index of normal variation.

3.2. Stimuli and Procedures.

3.2.1. Expository Tasks. Three expository tasks were randomly presented, videotaped, and digitally audiotape recorded in a quiet room. These samples were analyzed with respect to planning cycles as an overall measure of discourse planning (32, 45, 46) and with acoustic measures traditionally associated with speech planning (21).

Using the pitch contour subroutine of the Computerized Speech Laboratory (CSL), Henderson’s (46) procedures for the determination of planning cycles were updated. Planning cycles were defined by periods of relative fluency in which no pause exceeded 250 msecs. and periods of relative pausing in which silent pauses were ≥ 250 msecs. (12, 47-49). The durations of fluent and pause periods were computed and plotted as a step-graph with Claris Works. The planning cycles then were visually compared within and between subjects to verify the existence of planning cycles independent of individual differences and to note differences in the shape of the cycles that may be attributed to style of discourse planning (proximal vs. distal planning), discourse complexity and/or size and location of cerebral lesion.

In order to permit more detailed acoustic analysis, each sample was transcribed orthographically and then segmented into Communication Units (C-Units) (50), a clausal unit, and Intonation Units (IPs) (5), a prosodic unit. To assess the consistency of C-Unit and IP segmentations, two independent raters evaluated all of the transcripts by relistening to the video/audio tapes and judging the accuracy of the transcriptions and unit boundaries. Interobserver reliability was determined with the kappa formula (51).

Acoustic analysis of the discourse-based data for individual C-Units and IPs was carried out. The following measures were obtained: (1) normalized pause durations between units or the sum of all pre-unit pauses divided by overall duration (both in msecs.); (2) pause-to-speaking ratio (PSRAT) (32); (3) coefficient of variation for frequency and amplitude values (21); and (4) speaking rate in syllables per second for each discourse sample (52). The first measure considered individual planning style while the second focused on local planning. The last two measures covered the more traditional view of prosody as including variations in f0, amplitude and timing.

3.2.2. Constrained Speaking Tasks. To assess the speaker’s use of prosody in isolated sentences, three more constrained speaking tasks were administered for comparison with the expository tasks. These three tasks involved the production of sentences that manipulated or violated certain prosodic parameters, such as intonation, word stress and pause placement.
(adapted from 53). The tasks were designed to examine volitional production of certain aspects of prosody. These performances were analyzed using the pitch and amplitude contour subroutines of CSL, as well as for the durations of specific units.

4. RESULTS

4.1. Methods of investigating prosody. Proximal and distal planning styles were evidenced in both clinical and control cases. This would suggest that planning style was not interrupted by cerebral infarction. Hesitations were noted in both groups and were indicative of lexical search or thematic shifts in focus. Location of hesitations seemed to correspond more closely with the IP than the C-Unit. Hence, the unit of encoding seems to be something other than the clause.

4.2. Role of acoustic measurement. In terms of the measurement of non-affective prosody, the largest pre-unit pauses for all participants occurred in two contexts, when they initiated speaking on the topic and when they shifted to a new thematic focus. In addition, shifts in thematic focus were marked by predictable increases in PSRAT. In other words, at these points, more pauses were present when new thematic units were being planned for expression.

For AI, who had the most circumscribed lesion, timing differences were evident only in the clause level communication unit (C-unit) data, while HC had consistent timing difficulties within both the C-unit and IP data. However, there were overlaps in values for syllables per second and PSRAT between AI and the two controls. These overlaps suggest that no single measure was sufficient to diagnose disturbed prosody. Coefficient of variation for frequency also differed with AI having more difficulty controlling the frequency range of his C-units. These values differed between the two cases and the two controls, suggesting that the type of prosodic disturbance, clinically described as aphemia, affected pitch use. However, amplitude values did not differ for the two cases nor did amplitude measures differentiate the two cases from the controls. This result may indicate that amplitude does not function as a key acoustic parameter in aphemia.

4.3. Comparisons among tasks. Speech production task results indicated that both AI and HC were experiencing a prosodic disturbance that was not attributable to a demonstrable dysarthria or apraxia. Both men were able to produce the sentences varying prosodic elements successfully. However, individual variability was noted in how each produced the desired prosodic target. For example, AI had more difficulty modulating pitch as the primary element of prosody and relied more on manipulating segmental durations. On the other hand, HC experienced trouble regulating duration. Instead, he produced the target prosodic elements by using pitch variation strategies. The control subjects experienced no difficulties with these tasks. While more robust, these findings were similar to those evidenced in the discourse level data.

5. DISCUSSION

The advent of neuroimaging techniques in the 1970’s made it possible to identify pathologic anatomy in patients with aphasia. Observational studies of patients who have undergone neuroimaging studies have lead to revised notions about the patho-anatomic localization of Broca’s area. Mohr, et al. (54) observed that infarctions restricted to Broca’s area did not produce a persistent disorder of language, but rather a “mutism that is replaced by a rapidly improving dyspraxia and effortful articulation” (p. 311). The complete syndrome of Broca’s aphasia is now thought to be a sequela of damage to several contiguous areas located within the peri-Sylvian region of the language dominant hemisphere (55).

Our research supports observations that small lesions at different sites within the peri-Sylvian region produced distinctive clinical syndromes. Lesions that involve the inferior frontal gyrus, but spare the precentral gyrus, produce impairments in language and speech initiation, but do not interfere with repetition or articulation. Lesions limited to the inferior portion of the precentral gyrus and adjacent pars opercularis are associated with impaired articulation, but normal initiation of language and speech (55). In the two patients described, who had lesions confined to the inferior portion of the precentral gyrus, dysprosody was the primary language impairment. The few prior case studies that included detailed speech and language analyses, performed on patients with postmortem or neuroradiographic evidence of lesions confined to the precentral gyrus, also suggested damage confined to the inferior portion of this gyrus produces a slow, dysarthric, and dysprosodic speech pattern (22, 56-58). Thus, elements of Broca’s aphasias can appear in isolation from other clinical signs that are typically found in this type of aphasia.

Since small, discrete lesions of the peri-Sylvian region produce impairments of specific components of linguistic expression, it appears that individual systems governing different aspects of expressive language are located within anatomically distinct sites. The “classical” Broca’s area is comprised of several smaller anatomic foci which interact to generate spoken language. Precisely how these foci interact is not known.

REFERENCES


