Influence of Language Structure on Brain-Behavior Development

JENNIFER S. BUCHWALD, DONALD GUTHRIE, AND JUDITH SCHWAFFEL

Department of Physiology, Brain Research Institute and Mental Retardation Research Center, UCLA School of Medicine, Los Angeles, California

ROLAND J. ERWIN

Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania

AND

DIANA VAN LANCKER

Department of Neurology, University of Southern California School of Medicine and VA Outpatient Clinic, Los Angeles, California

Lack of exposure to specific sensory patterns during critical periods of development can result in a lack of responsiveness to those stimuli in adulthood. The present study extends these observations to native speakers of Japanese, a language which does not contain the contrastive /r/ and /l/ sounds present in English. Both electrophysiological (P3 event-related evoked potential) and behavioral results indicate deficient or absent discrimination of /r/ versus /l/ sounds in Japanese adults compared to native speakers of English. Thus, language structure appears to provide a subtle yet measurable effect on specific aspects of brain development and function.

Linguistic differences between English and Japanese are well documented and indicate that a phonological differentiation of /r/ and /l/ sounds occurs in the English language, whereas in the Japanese language, this differentiation is absent. The acoustics of the sonorant sounds /r/ and /l/ rely upon subtle contrasts in the position of the third formant and

This work was supported by USPHS Grants HD05958 and HD04612 and the Veterans Administration. We acknowledge N. Akatsuka and the UCLA Phonetics Laboratory for cooperation in stimulus preparation and we thank J. E. Bogen, D. Kempler, P. Ladefoged, and R. Strandburg for helpful discussions. Address correspondence and reprint requests to Diana Van Lancker, Ph.D., VA Outpatient Clinic (126), 351 E. Temple Street, Los Angeles, CA 90012. e-mail: DIANAVL@USCVM.BITNET.
the rate of its transition to be classified as distinctive phonemes in English (Dalston, 1975; Haggard, 1967; Lehiste, 1961; Lisker, 1957; Shoup & Pfeifer, 1976). These acoustic differences are not utilized phonologically (Jakobson & Halle, 1968; Ladefoged, 1975) in Japanese (Jorden, 1962; Mori, 1929); the Japanese sound transcribed as /r/ is pronounced most often as an alveolar flap but variants resembling English /r/, /l/, and /d/ also occur (the range of sounds classifiable as /r/ is unusually diverse—see Lindau, 1985). These sounds are all heard as /r/ consistently in the language culture to which native Japanese speakers are exposed throughout their early development. Such culturally driven determinants in the linguistic milieu provide specific differences in the auditory environments of the individual exposed to the Japanese vs. English language throughout childhood.

In developmental biology, numerous examples indicate that lack of exposure to specific sensory, e.g., auditory, patterns during critical periods of development can result in a lack of neurobehavioral responsiveness to those stimuli in adulthood (Nettebohm, 1979). Similarly, it is widely believed that a "critical" (Lenneberg, 1967) or "sensitive" period (Oyama, 1979; Scott, 1978; Flege, 1988) exists for human language learning. In this view, the capacity of the brain for reception, storage, and imitation of new linguistic sounds may commence as early as 6 months of age (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992) and peaks well before adolescence (Lenneberg, 1967; Flege, 1988; Scovel, 1988). This linguistic capacity diminishes with age as linguistic encoding becomes restricted largely to the left hemisphere. Dramatic demonstrations of this principle are provided by individuals who, as children, underwent surgical hemispherectomy to relieve intractable epilepsy. Longitudinal study of these individuals, after removal of the left cerebral hemisphere, indicates near-normal reacquisition of language (Dennis & Kohn, 1975; Dennis & Whitaker, 1976). In adults, in contrast, left hemispherectomy permanently abolishes propositional speech and most other language function (Smith, 1974; Van Lancker, 1988). Indeed, even restricted, focal left hemisphere lesions in adults can produce aphasia over a prolonged period of time (Benson, 1979; Bogen & Bogen, 1976), often permanently, despite extensive therapy (Naeser, Palumbo, Helm-Estabrooks, Stiassny-Eder, & Albert, 1989). These observations extend numerous earlier studies and further suggest that the acoustic-speech environment during childhood engages brain processing and storage systems that are more "open" or "plastic" than is the case in the postadolescent adult brain.

Based on the fact that language learning by Japanese children does not include phonological differentiation of acoustic cues underlying /r/ and /l/, as well as on the finding that Japanese adults cannot reliably identify synthesized /r/ and /l/ speech-like stimuli (Mann, 1986; Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975; Mochizuki,
1981; Strange & Dittmann, 1984), we hypothesized that the adult Japanese brain would be unable to discriminate English word-pairs (i.e., "minimal pairs") differing only in /r/ versus /l/ phonemes (e.g., "liver/river, rip/lip"). In order to test this hypothesis, as well as to examine the more general question of linguistic influences upon brain–behavior development, the following study of Japanese and American adults was undertaken. To provide objective assessments, four different experiments were designed, one that measured electrophysiological responses during auditory exposure to sound contrasts and three that measured overt discrimination, identification, and recognition of these sounds.

To examine brain processing of /r/ and /l/ stimuli, electrophysiological recordings of evoked response potentials (ERPs) were utilized, with the cognitive P3 component receiving particular emphasis. The P3 appears as a large, long-latency positive potential in response to a rare, unanticipated stimulus which intermittently occurs during repeated presentations of a frequent, anticipated stimulus (Squires, Halgren, Wilson, & Crandall, 1983). While its generator system is still not clearly understood, this potential provides a powerful tool for examining the brain’s ability to discriminate between "rare" and "frequent" stimuli at processing levels beyond those of the primary sensory systems (Erwin, Van Lancker, Guthrie, Schwafel, Tanguay, & Buchwald, 1991).

METHODS

Subjects

The Japanese subjects were native to Japan and spoke Japanese as a first language but could converse in, read, and write English with proficiency. All were Japanese university faculty members or their spouses and had been in the United States less than 1 year at the time of testing. The group consisted of 7 females and 7 males in the 29–46 year age range (mean age = 35.4 years). The American group consisted of 14 adults, 7 females and 7 males, in the 20–30 year age range (mean age = 23.4 years), who were all native to the United States and spoke English as a first language. All subjects had normal hearing. Informed consent was obtained from each subject.

Measures

For electrophysiological recordings of the P3, two intermixed "rare" (20%) and "frequent" (80%) stimuli were presented pseudorandomly (i.e., no two rare stimuli occurred sequentially). The stimuli occurred in blocks of 300 trials each and were counterbalanced so that the rare stimulus became the frequent stimulus in a second 300-trial block. Four stimulus pairs were used: (1) two minimally contrasting consonant–vowel syllables "ba/pa"; (2) a word pair, "rip/lip"; (3) a linguistic–prosodic pair, “Bob! (exclamation, i.e., high falling intonation)/Bob?” (question, i.e., low rising inflection); and (4) an affective–prosodic pair, “Bob (angry)/Bob” (happy), differentiated only by acoustic cues signalling affective meanings (Williams & Stevens, 1972; Van Lancker & Sidtis, 1992). Stimulus pair one, the linguistic–syllabic contrast "ba/pa," sounds similar whether pronounced in Japanese or English, and English words containing these sounds are readily learned by Japanese
second-language speakers of English. Stimulus pairs three and four were included to examine Japanese abilities to discriminate, identify, and recognize nonphonological, prosodic contrasts carried in the English speech signal. The stimuli were digitized from tape recordings of the words spoken naturally by a female voice and were equalized with regard to duration (420–440 msec) and peak intensity (60 dB SPL + 5 dB). The stimuli were delivered binaurally through Sony Nude earphones at a rate of 1/1.5 sec.

All recordings and behavioral tests were carried out in an electrically shielded, sound attenuated recording room. Recordings of evoked potential activity were obtained from electrodes placed centrally and laterally using an Electrocap and were referred to linked mastoid electrodes. Electrodes were also placed at the canthal and supraorbital regions of the left eye to monitor eye movements. EEG activity was amplified 50K using Grass P511 amplifiers with the bandpass set at 1–100 Hz; eye movement activity was amplified 10K with the bandpass set at 1–100 Hz.

Responses to the rare and frequent stimuli were digitized and averaged online over a 50 msec pre- and 750 msec post-stimulus time period using a 5-msec sampling rate (160 data points). In order to equalize the averaged number of trials, frequent stimulus averages included only trials which immediately preceded the rare stimulus. Thus, each average was composed of 60 trials. Because of the obviously different ERP latencies and morphologies elicited by the various stimuli, the data set for each stimulus was analyzed separately.

The averaged ERPs recorded at Cz, the locus of maximum P3 amplitudes, were compared among subjects within each group using paired r tests to provide a group statistic reflecting latency intervals across which ERP amplitudes differed significantly between rare and frequent stimulus presentations. For each stimulus, the 160 data points obtained when that stimulus was rare were compared with the 160 data points obtained when that stimulus was frequent. To avoid artifacts of multiple r test comparisons, our interpretation of "significance" required that ten or more successive points (>50 msec) show significant difference (p < .05). Monte Carlo results indicate that sequences of data points with the autocorrelation structure exhibited by these samples would produce ten consecutive "significant" r values less than once in 200 times (Guthrie & Buchwald, 1991). To compare between groups, unpaired r tests were similarly computed for each of the 160 data points.

For the first behavioral measure, during P3 recordings, the subjects were asked to press a button to each rare stimulus presentation. Each button-press response was registered by the computer as a correct response or as an incorrect response due to omission or commission. Data were collated for each subject and pooled for the group. Across-group comparisons of error rates were made using a conditional binomial test based on a Poisson distribution of errors. This task addresses auditory–perceptual discrimination of the pairs, measuring whether a given target stimulus ("rip" or "lip") is "heard" as "same as" or "different from" the previous stimulus or sequence of stimuli.

In a second behavioral task, the subjects were presented with pairs of pictures with word captions, which corresponded to the P3 stimulus pairs (Fig. 1). They were asked to place a mark under the picture that corresponded to the stimulus that they heard. This task includes auditory discrimination as well as identification, given a choice of two items. The stimuli, presented via tape recording delivered by a female voice at intervals of 8 sec, were randomly intermixed and each stimulus was presented four times. Data were analyzed as for the button-press task.

A third behavioral task tested the subject’s ability to discriminate between word pairs beginning with /r/ or /l/. A printed list of five word pairs ("rip/lip, rice/lice, rain/lane, red/lead, rock/lock") was given to each subject who was asked to circle the word that was heard on each trial. The taped auditory stimuli, spoken by a female voice at intervals of 8 sec, were randomly intermixed and each stimulus was presented four times during the session. This discrimination-identification task extended the "r/l" query (in behavioral task two) to additional word pairs. Data were analyzed as for the other behavioral tasks.
RESULTS

The results of this study indicated that in all three behavioral tasks, the Japanese and Americans performed equally well, except for those trials utilizing /r/ and /l/ stimuli.

In the first task, the Japanese and American button-press responses to the rare P3 stimuli during P3 recordings were highly accurate with one exception: the Japanese performance on the "rip/lip" stimulus pairs. In contrast to only 29 errors in the Japanese group made on the other six stimuli combined, the Japanese subjects made 345 errors of button-press omission or commission to the "rip/lip" stimuli. This result produced a significant ($p < .001$) group difference between Japanese and American error rates.
In the second behavioral measure, both Japanese and American subjects matched the stimuli used in the P3 task ("rip/lip, ba/pa, Bob!/Bob?, Bob-angry/Bob-happy") to the appropriate picture/word (Fig. 1) with relatively few errors. Again, the single exception appeared in the Japanese response to "rip/lip." For the Americans, an error rate of 0/112 trials accompanied these stimuli while the Japanese error rate was 23/112 trials, a significant ($p < .001$) group difference.

In the third behavioral test, American subjects discriminated individual words from the five /r/ and /l/ minimal pairs with essentially no errors (2/280 trials). In contrast, the Japanese error rate was 55/280 trials which again resulted in a significant difference between the American and Japanese groups ($p < .001$).

Electrophysiological recordings from both the Japanese and American subjects showed typical auditory evoked potentials, with a larger endogenous P3 response to the rare than to the frequent stimulus. Rare-frequent difference potentials, a measure which demonstrates the magnitude of the P3 discrimination of rare versus frequent stimuli, are shown in Fig. 2 for the American and Japanese groups. In each of the grand averages of the American group, a large P3 difference potential in the 250–550 msec latency range is evident. Paired t test comparisons of the American data showed that significant rare/frequent differences were produced in the latency range of the P3 by each of the eight stimuli (Table 1). The Japanese grand averages also show large difference potentials in the latency range of the P3 to most of the stimuli; however, in response to "lip" and "rip," the difference potentials are notably reduced or absent (Fig. 2). Paired t test comparisons of the Japanese data showed that significant rare/frequent differences were not produced either by "lip" or by "rip" (Table 1; Fig. 3). Significant rare/frequent differences in the P3 latency range were produced by all other stimuli except "ba." This last statistic was surprising since the grand averages showed clear P3 difference potentials to both "ba" and "pa"; response variability may have precluded statistical significance in this case. This interpretation is supported by the fact that the Japanese subjects showed the same highly accurate level of behavioral discrimination of "ba" as did the American subjects. Group comparisons did not show any additional differences.

In contrast to the virtually nonexistent Japanese P3 evoked responses to the "rip/lip" stimuli, the 100 ms negative N1 and 200 ms positive P2 potentials, which precede the P3 in latency, were consistently produced (Fig. 3). These potentials do not require rare/frequent stimulus discrimination for their appearance—i.e., they are not endogenous. The normal appearance of N1 and P2 in response to all stimuli in all Japanese subjects indicated, indeed, that the /r/ and /l/ stimuli were being centrally processed. However, the acoustic differences between these stimuli were not differentiated by the P3 generating system of the Japanese group.
Fig. 2. Difference potential grand averages for the American and Japanese groups to each of the eight stimuli. The grand averages reflect “rare–frequent” response differences for each of the 14 subjects in each group. Positive voltage above the baseline; stimulus onset occurs at 0 ms.
TABLE 1
Paired t Test Comparisons of Rare vs Frequent Stimulus Responses

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Latency intervals (ms) exhibiting significant differences*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Americans</td>
</tr>
<tr>
<td>Angry</td>
<td>290–445</td>
</tr>
<tr>
<td>Happy</td>
<td>290–475</td>
</tr>
<tr>
<td>Ba</td>
<td>300–500</td>
</tr>
<tr>
<td>Pa</td>
<td>310–445</td>
</tr>
<tr>
<td>Statement</td>
<td>335–480</td>
</tr>
<tr>
<td>Question</td>
<td>395–560</td>
</tr>
<tr>
<td>Lip</td>
<td>385–525</td>
</tr>
<tr>
<td>Rip</td>
<td>415–480</td>
</tr>
</tbody>
</table>

* (p < 0.05).

DISCUSSION

Taken together, the present data indicate significant differences between the Japanese and American subjects in their ability to discriminate between /r/ and /l/ sounds that are not systematically contrasted in the Japanese language but which do occur contrastively in English. Other verbal stimuli, making up linguistic, linguistic–prosodic, and affective–prosodic contrasts, produced no differences between the groups either behaviorally or (except for “ba”) electrophysiologically. An intact P3 for the two prosodic pairs for both experimental groups merely means that the pair members were discriminable to both the native Japanese and the native English speakers. However, the result is also consistent with the observation that some Japanese linguistic–prosodic contrasts under some conditions bear a similarity to those in English (Abe, 1955) and it is harmonious with an oft-expressed belief in the cross-language universality of prosodic information in speech.

Other studies focussing upon production of new phonemes indicate that pronunciation success is significantly greater for individuals who begin second language learning before the age of 12 and that, for these children, length of residence in the nonnative country is a significant predictor of pronunciation success (Oyama, 1982; Purcell & Suter, 1980; Snow & Hoefnagel-Hoehle, 1982). Adults, in contrast, show relatively poorer second language pronunciation and little or no benefit from prolonged residence (Flege, 1988), so that adult pronunciation in the second language may be said to “fossilize” at a relatively early stage of second language learning (Selinker, 1972). Since speech production ability is related to but not fully dependent on perception ability (Ladefoged, 1967), the inability of the Japanese to produce words containing /r/ or /l/ cor-
Fig. 3. P3 grand averages for American and Japanese groups to the word-pairs "lip" and "rip." Each trace represents the average of 60 trials for each of the 14 subjects in each group. The first pair of traces compares responses to stimuli present in the "rare" condition vs. the "frequent" condition; the difference potential resulting from subtraction of the frequent from rare response is displayed beneath. Positive voltage above the baseline; stimulus duration is indicated by dots above the baseline.
rectly may only partly reflect their inability to discriminate these sounds: in fact, some native Japanese can more accurately produce than perceive the English /r/-/l/ contrast (Goto, 1971; Sheldon & Strange, 1982). It has been suggested that individuals who begin learning a second language by the age of 5 years, but not as adults, may form a new phonetic category for sounds that differ physically from counterparts in the native language, and that after the age of 5, it may be impossible for learners to produce the authentic second language sounds (Flege, 1988; Borden, Gerber & Milsark, 1983). The effects of formal training on adult perceptual learning of a new phonemic contrast in the second language have been variable, with reportedly little change (Strange & Dittmann, 1984; Flege, 1989; Henly & Sheldon, 1986) to moderate improvement (Goto, 1971; MacKain, Best & Strange, 1981) especially when certain conditions are met in the training environment (Logan, Lively & Pisoni, 1991; Pisoni, Lively, & Logan, 1993) with exceptions attributable to various influences, including natural aptitude (Underbakke, Polka, Gottfried, & Strange, 1988). The effect of production and/or perception practice on adult ability to discriminate and produce “new” phonological contrasts merits further study.

Finally, in considering physiological mechanisms underlying the diminished ability of our Japanese subjects to discriminate /r/ and /l/ words, brain systems responsible for P3 generation appear to be relevant. That the P3 system was generally functional in the Japanese subjects was clearly indicated by the normal P3 responses elicited by most stimuli. Moreover, the words forming minimal pairs beginning with /r/ or /l/ were “heard” by the Japanese subjects, as was shown by the consistent appearance of the non-endogenous N1 and P2 evoked potentials to these stimuli (see Fig. 3). Both animal (Buchwald, 1990; Harrison, Buchwald, Kaga, Woolf, & Butcher, 1988) and human (Squires et al., 1983; Halgren et al., 1980; Okada, Kaufman, & Williamson, 1983) experiments suggest that structures within the limbic system may be important or even essential for P3 generation. Other studies have suggested a role for temporoparietal (Goff, Allison, & Vaughan, 1978; Smith, Halgren, Sokolik, Baudena, Musolino, Liegeois-Chauvel, & Chauvel, 1990) or frontal cortex (Wood & McCarthy, 1986), subcortical centers, e.g., thalamus (Yingling & Hosobuchi, 1984), or possibly for some cortical–subcortical interaction (Turetsky, Raz, & Fein, 1990). The present study suggests that at least some components of this P3 generating system are “plastic,” with a different discriminative capacity in the American subjects, exposed during childhood to systematically contrastive /r/ and /l/ sounds, than in the Japanese who lacked this exposure. Whether intensive training in perception and/or production of the phonemic contrast would alter brain structure, as has been reported following altered afferent input to adult primate cortex (Pons, Garraghty, Ommaya, Kaas, Taub, & Mishkin,
1991; Pons, Garraghty, & Mishkin, 1988), remains to be studied. In either case, language structure appears to provide a subtle yet measurable biological effect on brain development and, thereby, upon subsequent brain function and behavior.

REFERENCES


Pons, T. P., Garraghty, P. E., Ommaya, A. K., Kaas, J. H., Taub, E., & Mishkin, M.


