Information about Ohio State University WORKING PAPERS IN LINGUISTICS
(OSU WPL)

Working Papers in Linguistics is an occasional publication of the Department of Linguistics of Ohio State University and usually contains articles written by students and faculty in the department. There are one to three issues per year. Information about available issues appears below. Nos. 1, 5, 10, & 23 are out of print and no longer available.

There are two ways to subscribe to WPL. The first is on a regular basis: the subscriber is automatically sent and billed for each issue as it appears. The second is on an issue-by-issue basis: the subscriber is notified in advance of the contents of each issue, and returns an order blank if that particular issue is desired.

OSU WPL #18, $2.00. 183 pp. (June 1975): Papers by Michael Geis, Sheila Geoghegan, Jeanette Gundel, G. K. Pullum, and Arnold Zwicky.

OSU WPL #19, $2.00. 214 pp. (September 1975), Edited by Robert K. Herbert: Patterns of Language, Culture, and Society: Sub-Saharan Africa contains eighteen papers presented at the Symposium on African Language, Culture, and Society, held at Ohio State University on April 11, 1975.

OSU WPL #20, $2.00. 298 pp. (September 1975), Edited by Robert K. Herbert: Proceedings of the Sixth Conference on African Linguistics contains twenty-seven papers presented at the Sixth Conference on African Linguistics, held at the Ohio State University on April 12-13, 1975.

OSU WPL #21, $2.50. 252 pp. (May 1976), Edited by Arnold Zwicky: Papers on Nonphonology. Papers by Steven Boer and William Lycan, Marian Johnson, Robert Kantor, Patricia Lee, and Jay Pollack.

OSU WPL #22, $3.50. 151 pp. (February 1977), Edited by Olga Garnica: Papers in Psycholinguistics and Sociolinguistics. Papers by Sara Garnes, Olga Garnica, Mary Louise Edwards, Roy Major, and John Perkins.

OSU WPL #24, $3.50. 173 pp. (March 1980), Edited by Arnold M. Zwicky: Clitics and Ellipsis. Papers by Robert Jeffers, Nancy Levin (OSU Ph.D. Dissertation), and Arnold Zwicky.

OSU WPL #25, $3.50. 173 pp. (January 1981), Edited by Arnold M. Zwicky: Papers in Phonology. Papers by Donald Churma, Roderick Goman (OSU Ph.D. Dissertation), and Lawrence Schourup.

OSU WPL #26, $4.50. 133 pp. (May 1982), Edited by Brian D. Joseph: Grammatical Relations and Relational Grammar. Papers by David Dowty, Catherine Jolley, Brian Joseph, John Nerbonne, and Amy Zaharlick.

OSU WPL #27, $4.50. 164 pp. (May 1983), Edited by Gregory T. Stump: Papers in Historical Linguistics. Papers by Donald Churma, G. M. Green, Leena Hazelkorn, Gregory Stump, and Rex Wallace.

OSU WPL #28, $5.00. 119 pp. (May 1983). Lawrence Clifford Schourup, Common Discourse Particles in English Conversation. (OSU Ph.D. Dissertation)

OSU WPL #29, $5.00. 207 pp. (May 1984), Edited by Arnold Zwicky & Rex Wallace: Papers on Morphology. Papers by Belinda Brodie, Donald Churma, Erhard Hinrichs, Brian Joseph, Joel Nevis, Anne Stewart, Rex Wallace, and Arnold Zwicky.

OSU WPL #30, $5.00. 203 pages. (July 1984). John A. Nerbonne, German Temporal Semantics: Three-Dimensional Tense Logic and a GPSG Fragment. (OSU Ph.D. Dissertation).

OSU WPL #31, $6.00. 194 pages. (July 1985), Edited by Michael Geis. Studies in Generalized Phrase Structure Grammar. Papers by Belinda Brodie Annette Bissantz, Erhard Hinrichs, Michael Geis and Arnold Zwicky.

OSU WPL #32, $6.00. 162 pages. (July 1986), Interfaces. 14 articles by Arnold Zwicky concerning the interfaces between various components of grammar.

OSU WPL #33, $6.00. 159 pages. (August 1986), Joel A. Nevis, Finnish Particle Clitics and General Clitic Theory. (OSU Ph.D. Dissertation).

OSU WPL #34, $6.00. 164 pages. (December 1986), Edited by Brian Joseph. Studies on Language Change. Papers by Riitta Blum, Mary Clark, Richard Janda, Keith Johnson, Christopher Kupec, Brian Joseph, Gina Lee, Ann Miller, Joel Nevis, and Debra Stollenwerk.

OSU WPL #35, $10.00. 214 pages. (May 1987), Edited by Brian Joseph and Arnold M. Zwicky. A Festschrift for Ilse Lehiste. Papers by colleagues of Ilse Lehiste at Ohio State University.

OSU WPL #36, $10.00. 140 pages. (September 1987), Edited by Mary Beckman and Gina Lee. Papers from the Linguistics Laboratory 1985-1987. Papers by Keith Johnson, Shiro Kori, Christiane Laeufer, Gina Lee, Ann Miller, and Riitta Valimaa-Blum.

OSU WPL #37, $10.00. 114 pages. (August 1989), Edited by Joyce Powers, Uma Subramanian, and Arnold M. Zwicky. Papers in Morphology and Syntax. Papers by David Dowty, Bradley Getz, In-hee Jo, Brian Joseph, Yongkyoon No, Joyce Powers, and Arnold Zwicky.

The following issues are available through either: The National Technical Information Center, The U.S. Department of Commerce, 5285 Port Royal Rd, Springfield, VA 22151 (PB) or ERIC Document Reproduction Serv. (ED) Ctr. for Applied Linguistics, 1611 N. Kent St., Arlington, VA 22209.

#2. November 1968, 128 pp. (OSU-CISRC-TR-68-3). PB-182 596.
#3. June 1969, 181 pp. (OSU-CISRC-TR-69-4). PB-185 855.
#4. May 1970, 164 pp. (OSU-CISRC-TR-70-26). PB-192 163.
#6. September 1970, 132 pp. (OSU-CISRC-TR-70-12). PB-194 829.
#7. February 1971, 243 pp. (OSU-CISRC-TR-71-7). PB-198 278.
#8. June 1971, 197 pp. (OSU-CISRC-TR-71-7). PB-202 724.
#9. July 1971, 232 pp. (OSU-CISRC-TR-71-8). PB-204 002.
#11. August 1971, 167 pp. ED 062 850.
#12. June 1972, 88 pp. (OSU-CISRC-TR-72-6). PB-210 781.
#13. December 1972, 225 pp. ED 077 268.
Parts of #14. April 1973, 126 pp. ED
#15. April 1973, 221 pp. ED 082 566.
Parts of #16. December 1973, 119 pp. ED

Please include payment with your order--make checks payable to "The Ohio State University". Send to: OSU WPL, Department of Linguistics, Ohio State University, 204 Cunz Hall, Columbus, OH 43210-1229.

_______ Regular subscription _______ Issue by issue basis

_______ No subscription please, but send the following back issues.

_____	#18 ($2.00)	_____	#22 ($3.50)	_____	#26 ($4.50)	_____	#30 ($5.00)	_____	#34 ($6.00)
_____	#19 ($2.00)	_____	#23 ($3.50)	_____	#27 ($4.50)	_____	#31 ($6.00)	_____	#35 ($10.00)
_____	#20 ($2.00)	_____	#24 ($3.50)	_____	#28 ($4.50)	_____	#32 ($6.00)	_____	#36 ($10.00)
_____	#21 ($2.50)	_____	#25 ($3.50)	_____	#29 ($5.00)	_____	#33 ($6.00)	_____	#37 ($10.00)

Name: ______________________________

Address: ______________________________

____

____

Foreword

This volume includes papers that take an experimental approach to issues in linguistics and the psychology of language, or that are closely related to that enterprise. It is the second of a series of progress reports from the Linguistics Laboratory (see OSU Working Papers #36). The languages examined are Arabic, Korean and English. Most of the work reported here was done in the Linguistics Laboratory between 1987-1989. The contributors were faculty and students at Ohio State during this period. Some authors have since left Ohio State; current affiliations are given for each.

McAdams' paper, a revised version of his master's thesis, reports on work done on the lateralization of the mental resources that support the analysis of morphologically complex words. He uses a technique that presents words to subjects in such a way as to force the subject to do initial visual processing in only one hemisphere. The results seem to reflect equal facility for dealing with morphologically simple words of several kinds in each hemisphere, but markedly superior facility for dealing with morphologically complex material in the left hemisphere.

Beale and Cowart use a related technique with words presented in auditory sentence contexts. This study yields evidence that a complex cognitive function involved in anaphoric processing emerges only when information about a target word is simultaneously available to each hemisphere.

Scholz discusses an issue in the philosophy of language that seems to bear in interesting ways on the received view of the relation between grammars and parsers. 'Kriptenstein' (Wittgenstein/Kripke) puts forward a 'skeptical paradox' that calls into question the coherence of Chomsky's psychologistic conception of grammar. Scholz reviews Chomsky's discussion of the paradox and concludes that his attempts to escape the force of the paradox do not succeed. Scholz's analysis suggests that one way the psychologistic conception might escape the paradox would depend, in part, on positing a much more transparent relation between the grammar and parser than is typically assumed. Such a move would indicate, among other things, that experimental results of various kinds ought to bear much more directly on issues in syntactic theory than is commonly assumed.

Cowart reaches related conclusions. The paper examines some sentences involving extraction from **picture** NPs. He finds evidence of patterns of acceptability that are unlike those commonly assumed in syntactic research, but which are, nevertheless, quite stable. These results seem to raise questions about the relation between the specifically linguistic resources of the mind and others that may participate in language comprehension in a more peripheral way. These findings also suggest that experimental approaches to issues in syntax may be able to reveal phenomena that conventional informal approaches have failed to detect.

deJong uses X-ray microbeam data to test two accounts of articulator timing which attempt to eliminate the time specification from the mental input: those of Kelso et al. (1985) and Harris et al. (1986). The intent is to help draw the boundary between those aspects of the speech production system that are properly included in the grammatical system that governs production and those regularities of the system that are determined by the anatomy and physiology of the vocal tract itself. In general, his study -- which uses actual English words, as opposed to reiterant speech -- confirms the findings of the earlier experiments.

Hussein's study, based on an acoustical analysis of VCV utterances in Arabic, considers the source and nature of vowel-to-vowel coarticulation effects. He finds that models of coarticulation proposed by Fowler (1983) and Keating (1985) do not adequately account for vowel-to-vowel coarticulation in Arabic. The paper discusses additional factors that play a role in coarticulation.

Lee investigates the acoustic and articulatory differences between lenis and fortis stops in Korean. Using information on production differences found in past studies, she synthesized both types of stop. These synthesized tokens were later used in perception tests. However, the tokens were not perceived categorically; Lee suspects that this is partly due to dialect. Some participants were native speakers of Chonnam, a dialect which has pitch accent. They may have used the quality of the following vowel to determine whether the preceding consonant was lenis or fortis.

Jun examines the accentual patterns and prosodic structure of Chonnam, a dialect of Korean which has phrasal pitch accent. She demonstrates that the phrasal domain of the pitch accent is the Phonological phrase (P-phrase) in the Prosodic Hierarchy of Selkirk and Nespor & Vogel. Moreover, Chonnam contains two levels above the P-phrase: the Intermediate phrase and the Intonational phrase.

We are grateful to Eliza Segura-Holland for her help in preparing the final layout.

Gina Lee
Wayne Cowart
July 1990

Ohio State University Working Papers in Linguistics #38

Papers from the Linguistics Laboratory

Table of Contents

Ohio State University Working Papers in Linguistics 38 (1-23)

Asymmetries in Naming Accuracy and in Event-related Potentials for Laterally Presented Words of Variable Morphological Complexity

Brian McAdams

University of Minnesota

Abstract: Recent studies indicate that hemispheric asymmetries in lexical access exist, with the left hemisphere being superior in processing morphologically complex words. The present study looks for asymmetries in naming accuracy and in topographically displayed event-related potentials to laterally presented words. Data were collected from six subjects while words were presented to both visual fields. Right visual field superiority in naming accuracy was demonstrated for suffixed forms as expected, indicating a potential left hemisphere advantage for morphological processing. Event-related potential data revealed stimulus-relevant positive parietal peaks around 320 msec.

1. Introduction

Evidence has been accumulating for the past 150 years that the cognitive processing of language is asymmetrically localized within the cerebral cortex. The earliest evidence came exclusively from studies of individuals with language deficits resulting from stroke or trauma. Such aphasic individuals display a wide variety of disturbances of language comprehension, language production, or both; and, in nearly every case, this deficit results from a perisylvian lesion in the left cerebral cortex. Individuals with similar lesions in the right hemisphere usually did not display language deficits, or, if they did, their deficits in linguistic functioning differed from those resulting from left-hemisphere damage. For example, lesions to the ventrolateral frontal cortex and the dorsolateral temporal cortex (both near the sylvian fissure) generally only result in language impairment with a left hemisphere locus (Hecaen and Albert 1978).

More recently, other sources of evidence for the lateralization of linguistic processing have been utilized. For example, studies of hemispherectomy patients have shown that the removal of the left hemisphere can result in a variable amount of language loss; in some cases (e.g. if the left hemisphere is removed relatively early in life), the right hemisphere can perform phonological and semantic processing, but it tends to have considerable difficulty with syntactic processing (Millar and Whitaker 1983; Todorovic 1988).

Investigations of split-brain patients who have undergone partial or complete callosectomies have revealed that the right hemisphere does have a limited linguistic competence, but it is apparently unable to process many (if not most) function morphemes and it is unable to make many important syntactic distinctions, such as active/passive, past/future, singular/plural (Zaidel 1978; Todorovic 1988). Millar and Whitaker (1983) warn that we should be cautious in looking at evidence from split-brain subjects because their brains are not normal- due to the damage that caused the epilepsy (which the surgery was designed to treat), as well as because of the damage caused by the operation (i.e. one side of the brain must be pulled back to reach the corpus callosum, some arteries must be clamped (thus starving some tissues), and retrograde axonal degeneration may cause further damage). Also, neural plasticity may have allowed some cognitive functions to shift to the unimpaired hemisphere from the hemisphere with the epileptic locus in pre-operative split-brain patients, thus making the functional capacities of the post-operative commissurotomy patient studied by the researcher atypical. A direct comparison between the normal subject and the split-brain patient may therefore be inappropriate or at least regarded with caution (Millar and Whitaker 1983).

Lateralization of language processing has also been observed using the Wada test which reveals the hemisphere that is dominant for language functions in the pre-operative brain surgery patient. In this test, the patient receives an injection of sodium amytal in the inner carotid artery at the point where the blood flow is unilateral. Of over 36 patients tested by Millar and Whitaker (1983), one had bilateral language representation, one was unilaterally right hemisphere dominant for language, and the rest were unilaterally left hemisphere dominant for language processing (using a naming task).

Changes in metabolic activity have been measured using PET-scan technology to localize centers of language processing throughout the brain. Millar and Whitaker (1983) state that verbal analogy testing has been shown to result in elevated metabolic activation of Wernicke's area in the temporal lobe. Petersen et al. (1988) investigated the relationship between auditory and visual lexical processing using a PET-scan subtraction technique with promising results. They found different cortical areas (in both hemispheres, but primarily in the left hemisphere) support visual and auditory word processing (primary and "non-primary").

Brain stimulation studies utilizing electrode stimulation of cortical tissue of conscious brain surgery patients such as those of Ojemann and Whitaker (Millar and Whitaker 1983) have shown "that language lateralization can be quite complete," the right hemisphere contributing little to language processing at least as it is measured by their technique.

Non-invasive techniques have also been used to study the lateralization of language processing, the most popular being: dichotic listening, divided visual field stimulation, and event-related potential studies. The first of these utilizes a binaural presentation of linguistic stimuli. This dichotic listening task assumes that an individual with left hemisphere language dominance will have a right ear advantage in accuracy of reporting different words presented simultaneously to both ears. This occurs as a result of the majority of ascending auditory axons synapsing in the contralateral auditory cortex of the temporal lobe (Borden and Harris 1984). However, there are a number of commissures in the ascending auditory pathways, and in addition to these subcortical decussations, there is the large cortico-cortical commissure, the corpus callosum mentioned above, which can allow the temporal lobe auditory centers to transfer information (Durrant and Lovrinic 1984). Indeed, for language studies, these commissures may have a detrimental effect on the usefulness of dichotic listening tasks for the study of cerebral lateralization of language. For example, Todorovic (1988) points out that, in general, morphosyntactic studies using a dichotic listening task have failed to give consistent results. These same difficulties have been encountered by other experimenters with different stimuli using this task.

Perhaps the most widely used technique for studying language lateralization has been the divided visual field technique in which linguistic stimuli are briefly presented laterally to the left or right visual hemifield, or both, using a tachistoscope or video display terminal. Unlike the organization of the auditory afferents, the visual afferents remain isolated so that both visual fields do not receive extensive bilateral representation in the cortex. Instead, the temporal visual field, that part of the visual field that falls on the half retina furthest from the nose (the temporal hemiretina) in each eye, projects ipsilaterally through the optic nerve to the occipital cortex. By contrast, that part of the visual field that falls on the half retina closest to the nose in each eye (the nasal hemiretina) projects contralaterally to the occipital cortex. Thus, all optical stimuli appearing to the left of the point of focus are initially processed by the right hemisphere while the left hemisphere initially processes stimuli in the right visual field (Anderson, 1982).

Although the divided visual field technique does have the advantage of a relatively more isolated ascending neural pathway than does the auditory system, it does pose some technical difficulties for the lateralization experimenter. Young (1982) discusses these potential limitations at some length and their resolutions as proposed by various researchers. For example, it is not uncommon for investigators to present words outside of the central or foveal region of the visual field which has been shown to have some bilateral representation in the striate cortex of the occipital lobe (Young 1982; Beaumont 1982; Leventhal et al. 1988; Beaton 1985; McKeever 1986), although the precise size of the bilaterally projecting area and the perceptual significance of this bilateral representation are not well understood. McKeever (1986) points out that the overlap may be ignored at higher levels of the visual system, and Young (1982) and Leventhal et al. (1988) suggest that the overlap may be involved in stereopsis. Beaton (1985) supports the latter point with evidence from split-brain patients that have difficulty with depth perception of the stimuli placed in front of or behind the point of fixation. In hopes of avoiding any possible bilateral initial processing of the experimental stimuli, experimenters present stimuli to the parafoveal region of the visual field.

Presentation of stimuli to the areas surrounding the fovea, the part of the visual field with the highest visual acuity, results in another point of contention: can the loss of acuity in the parafovea and the resulting stimulus degradation lead to misleading results in laterality studies? Chiarello (1988) compellingly argues against the controversial claim of Schwartz and Kirsner (1986) that visual acuity gradients account for the commonly observed right visual field (RVF) advantage for laterally presented words. They claim that the initial syllable is crucial for lexical access, and that it is more difficult to obtain in a left visual field (LVF) presentation. Chiarello has found that this theory of lexical access has little support and that other parts of monomorphemic words can be just as important as the syllable in accessing the word; indeed, that the whole word is what must be used for access. In an earlier study, Chiarello et al. (1986) found that varying duration, retinal eccentricity, and size of the stimuli to determine the potential contributions of these factors to asymmetries in lexical decision tasks with lateralized stimulus presentation did not affect the RVF advantage; suggesting that both hemispheres can use degraded sensory information relatively equally.

Stimulus duration is also a potentially crucial factor in divided visual field studies because of the potential that saccadic eye movements could bring the stimulus into the center of the visual field resulting in bilateral processing. This is especially important because it is hard to resist fixating on new stimuli. For this reason the stimulus must be presented for an interval shorter than that required to move the eyes such that the word is brought into the fovea. Stimulus location, intensity, size, and other factors can influence the latency and duration of a saccadic eye movement, but on average the saccade will begin about 180-200 msec post-stimulus onset in laterality studies and the saccade itself lasts about 20-30 msec (Young 1982). Young (1982) points out that for about 40-50 msec before and after, as well as during, a saccade, there is a significant loss of visual acuity. The range of stimulus durations used is highly variable, but the 50-200 msec range is most commonly used (Young 1982; Beaumont 1982). It might be noted that 100 msec are needed to "disengage . . . attention from any location in the visual field" and that saccade latency has been shown by Braun and Breitmeyer (1988) to depend on this variable rather than on the end of the fixation, and that attention can shift within the visual field without voluntary eye movement.

The divided visual field technique is frequently used with a lexical decision task (i.e. the subject is to say whether the stimulus is a word or not) or a naming task (i.e. the subject is to name the word they may have seen), and the accuracy and/or reaction time of the response is recorded. Both of these tasks and their measures are not looking at language processing as it occurs, but some time afterwards, potentially after other kinds of processing, even possibly interhemispheric transfer, could have occurred. The indirectness of the divided visual field technique can be corrected to some degree by coupling it with recordings of the brain wave activity during and after stimulus presentations. These event-related potentials (ERPs) "are changes in the electrical activity of the nervous system that are temporally associated with physical stimuli or psychological processes" (Picton and Stuss 1984). This method has the advantage of allowing the observer to indirectly "observe" the linguistic phenomenon of interest by measuring changes in the latency, amplitude, and/or spatial distribution of changes in the brain's electrical field that occur during and after stimulus processing rather than simply measuring behaviors long after processing has begun (and possibly been completed). The value of this approach for studies of the lateralization of different types of linguistic processing then is apparent, if differences in these variables can be correlated with differences in linguistic stimuli. For example, evidence for the lateralization of morphosyntactic processing to the left hemisphere might be proposed if a significantly larger potential could be localized to that hemisphere in conjunction with the processing of plural nouns but not with uninflected nouns.

ERPs recorded from scalp electrodes are not free of problems, however, in that many possible "sources with temporally and spatially overlapping fields" can generate them; thus, "each peak recorded in the scalp ERP does not necessarily reflect a separate cerebral process" (Picton and Stuss 1984). Considerable caution must therefore be taken in the interpretation of ERP data. Regan (1989) points out that "the more successful efforts at localizing intracranial (ERP) sources have involved large numbers of recording sites." A reliable method has also been developed to aid in ERP source derivation by Hjorth (1980). Thus, the problems arising from wave superposition, increase in amplitude of peaks of identical polarity or cancellation of peaks of opposite polarity (Allison et al. 1981; Regan 1989), can to a limited degree be overcome provided the generators of the summating potentials are at least as far apart as the electrodes. Studies such as those of Ducati et al. (1988) which demonstrated, using intracerebral recording as well as scalp electrodes in alert humans, that the VEP neural generators are completely within the striate cortex.

Previous studies of language processing using ERP techniques have primarily looked at phonetic segments, syllables, content/function words, and some sentence contexts (Picton and Stuss 1984; Molfese 1983; Garnsey 1985; Samar and Berent 1986; Kutas et al. 1988) with widely varied results. In part, this variability results from the lack of uniformity in divided visual field techniques, task selection, electrode montages, linguistic stimuli, and other variables as well (Molfese 1983). An example is seen in the work of Brown et al. (1976 (discussed in Picton and Stuss 1984)) who recorded ERPs to words in a common frame, "It was . . .", where the target word could be either a noun or a verb. Analyses found significant differences between nouns and verbs on the left anterior scalp. Principle components analysis of the same data (Brown et al. 1979, discussed in Picton and Stuss 1984) revealed peak latencies of 150, 230, and 370 msec on three components not in the left anterior scalp. In another study reported in Picton and Stuss (1984), Neville (1980) recorded evoked potentials to different words bilaterally presented at the same time. An N1 peak was found to be "significantly larger over the left parietal than over the right parietal scalp." No significant asymmetry was found at other electrode sites or if the stimuli were defocussed to the point of illegibility. Samar and Berent (1986), in an evoked response study of the prelexical locus of "the syntactic priming effect", found a presumably left hemisphere temporoparietal peak at 140 msec post-stimulus for centrally presented words, which seems to reflect whether a word occurred in an appropriate or inappropriate context.

2. Hemispheric Differences in Lexical Processing

The evidence for lateralization of linguistic processing and the techniques used to obtain this evidence as outlined in the previous chapter have led most investigators to conclude that, in the majority of the population, most language functions are handled by the left cerebral hemisphere (Caplan 1987). This left hemisphere dominance has been found to be strongest in most right-handed individuals and weakest in left-handed individuals. Millar and Whitaker (1983) report on a 1967 study by Zangwill who reviewed over 2000 brain damage cases reported in the literature and found that of those with left hemisphere damage 59.7% of right-handers were aphasic, 54.9% of left-handers were aphasic, and of those with right hemisphere damage, 1.8% of right-handers were aphasic, and 29.2% of left-handers were aphasic. As Caplan (1987) reports, it appears that familial handedness may also be relevant in determining the probability of language dominance being located in one or the other or both hemispheres. He states that "detailed studies of large populations have shown that strongly right-handed individuals from right-handed families have over a 98 per cent chance of being left-hemisphere dominant for language." Caplan is quick to add, however, that even in these individuals the right hemisphere does carry out some language-related processing.

Millar and Whitaker (1983) report that right hemisphere parietal lesions have been shown to significantly disrupt the comprehension of prosody (and thus affective speech). They also report that right-hemisphere lesions result in difficulties with metaphorical language, such as "a heavy heart"; these patients frequently interpret expressions such as this as meaning physically heavy rather than sad (n.b. other explanations relating to motivational states or affect may exist for the latter effect). Caplan (1987) points out that although split-brain studies, divided visual field studies, and dichotic studies show that the right hemisphere can do some other types of language processing to a limited extent, this does not mean that it performs these kinds of processing in normal language processing.

If there is such an extensive lateralization of linguistic functions to the left hemisphere, this raises several questions. Why is there a difference between the functioning of the hemispheres (e.g. the right hemisphere is often characterized as processing holistically and the left hemisphere as processing analytically), and how is this asymmetry anatomically realized? Also, is the lateralization as complete as Caplan (1987) and others claim (i.e. all linguistic processing is normally done by the left hemisphere except for the processing of prosodic elements), or are some linguistic functions normally bilaterally represented? In relation to this last question we might ask whether a process such as lexical access might be performed by the right hemisphere as well as the left if there is some capacity for performing linguistic processes such as prosodic analysis in the right hemisphere.

One answer that has been proposed to explain, in part, why hemispheric differences appear to exist is that of Geschwind and Levitsky (1968). They observed that the previously reported anatomical asymmetry of the temporal lobe was quite consistent within a large sample (i.e. 100 brains), and that this was a portion of the lobe that had previously been associated with linguistic processing, Wernicke's area, or, more precisely, the superior portion of the lobe in the insula known as the plana temporale. In 65% of the brains studied, the left planum temporale was larger, 11% of the brains had a larger right hemisphere planum temporale, and the rest (23%) had roughly equivalent plana temporale. Witelson (1983) reviewed a number of studies of planum temporale asymmetry and concluded that "all studies found the left planum to be larger . . . [to the extent that it is] a difference easily observable by gross visual inspection." She also points out that this asymmetry is reliable (in 70% of cases), a statement which cannot be made of other cerebral asymmetries such as those in the ventricular system, and vascular pattern asymmetries, and asymmetries "in the breadth and alignment of the frontal and posterior region of the hemispheres," asymmetries which are more difficult to relate to observed functional asymmetries.

On the cellular level, differences have also been observed between the right and left hemispheres in Broca's area (Scheibel et al. 1985). Broca's area in the left hemisphere was found to have a greater number of more extensively branched dendrites on layer III pyramidal neurons than other areas studied. They suggest that this may allow for more "degrees of freedom for the soma-dendrite complex with respect to a given input . . . [in that] each branch point represents a potential locus of enhancement or suppression of local electrical activity in the dendritic tree."

Such observations of macroscopic and microscopic asymmetries between hemispheres in language-relevant areas may be intriguing, but at this time they have not yet been definitively shown to correlate with hemispheric asymmetries in linguistic processing, although this may be an area that can be addressed on a gross scale with ERP and related technologies. For example, Garnsey (1985) used ERP to analyze the timing and localization of processes correlating with lexical access of function and content words. She found that the waveforms for these word types diverged after 200 msec, and that the difference was strongest near Broca's area, a finding consistent with function word localization evidence from aphasics. If function/content word distinctions can be identified using ERP data, can other evidence of linguistic asymmetries be demonstrated with this technology (e.g. high vs. low word imagability, inflectional vs. no inflectional morphology)? If linguistic functions that are known to be distinct are localized in one hemisphere, is it possible to demonstrate distinct anatomical localization of these processes using ERP mapping technology?

The present study attempts to answer, in part, these questions in relation to theoretically interesting neurolinguistic questions: 1) Can both hemispheres process morphologically complex words? 2) Is there an observably distinct localization of morphological processing within the cerebral hemispheres? That is, can differences between lexical access of monomorphemic and bimorphemic words be identified? 3) Can differences such as the distinction between inflectional and derivational morphological processes (Miceli and Caramazza 1988) be observed? A divided visual field study using a vocal reaction time measure by Goodall (1984) apparently demonstrates that the left hemisphere has a superior capacity for processing morphologically complex (agentive and plural) nouns, a finding which supported the earlier finding of Gazzaniga with split-brain subjects. Also, Todorovic (1988) conducted two divided visual field experiments with morphologically complex words which demonstrated that the right hemisphere is less sensitive to case-marking. He suggests that this may result from the right hemisphere having a different type of lexicon, or it may lack the necessary processing capabilities to handle morphology and syntax. These findings along with the ERP data of Garnsey and the researchers mentioned in section 1 give the impression that the objectives of this study may be within reach with this relatively modest technology.

To aid in isolating the linguistic capacities of the cerebral hemispheres, the present experiment utilizes a divided visual field presentation of the stimuli, and in order to identify the location of morphological processing in the brain, Brain Electrical Activity Mapping (BEAM) of ERP data is utilized. To determine the potential validity of the conclusion that the ERP data is actually measuring linguistic processing and not simply noise, accuracy of the responses are also recorded. It is expected that the right hemisphere will show a lower accuracy and less variation from the baseline electrical potential than the left hemisphere for those conditions with suffixes (i.e. where morphological processing is required). Increased left hemisphere activity is expected in or near the P3 and F3 electrode sites, with the P3 activity corresponding perhaps to Wernicke's area processing of content morphemes and F3 activity corresponding to function morpheme processing (i.e. the suffixes in this case).

3. Methods

3.1. Subjects

Nine right-handed subjects (8 female, 1 male) with 20/20 (or near) vision and right eye dominance were used. Subjects were screened for history of neurological, psychiatric, and visual disorders (none reported), as well as for familial sinistrality (FS+/FS-). Data from one FS+ subject (DB) were included in the results. Subjects were native speakers of English and were between 18 and 29 years old. Subjects were paid $10 and, in some cases, also received course credit for their participation.

Three subjects (2 female, 1 male) were unable to perform the task due to an inability to read the words (as determined by the continued naming accuracy below 10%) and/or the high level of artifact leading to nearly all trials being rejected; thus their data were not included in the results.

3.2. Equipment

Stimuli were presented on an Amdek 310A amber monitor by an IBM AT. ERPs were collected and analyzed using a Bio-Logic Brain Atlas system in the Neuroimaging Laboratory of the Department of Psychiatry, Ohio State University. Brain Atlas amplifiers were calibrated for a DC offset of $< 0.2\ \mu V$.

Electrode placement was performed according to the 10-20 system as described by Harner and Sannit (1974). Twenty-two gold-plated scalp electrodes were used with impedances $< 10\ k\Omega$ (generally $< 5\ k\Omega$) and within $3\ k\Omega$ of each other. A "rostral" nose electrode was used as a reference electrode, and the Fpz electrode served as the ground electrode. The nose electrode was used to avoid lateralized cerebral and heart artifacts that can be a problem for ear lobe, Cz, and linked mastoid reference sites. Gain was set at 30 k and filters at 30 and 0.1 Hz. The ERP analysis window was set for 1024 msec (the first 100 msec for pre-stimulus baseline). The automatic artifact reject was used to eliminate the noisiest samples. The computer sent a trigger pulse 100 msec before stimulus presentation to the Bio-Logic Brain Atlas to initiate ERP recording.

3.3. Stimuli and Procedure

Stimuli were single words presented horizontally in upper case letters for 150 msec. Since lateral word presentations make saccades to the stimuli difficult to resist, presentation times were kept short so that the end of the stimulus presentation would likely precede the completion of the saccade. Stimuli appeared randomly five character spaces to the left or right of a central distractor (a '+') which remained on the screen throughout most of the task. The fixation symbol disappeared 1.5 sec after stimulus onset as a signal for the subject to say the word aloud. The interstimulus interval was 3 seconds. The section of each visual field where words were presented was, in terms of visual angle, between 2° (inner boundary) and 6° (outer boundary of the longest word) from the center of the screen. Words were presented with the innermost edge 2° from the center. Thus words were presented beyond the regions of highest acuity and on the edges or outside of the retinal areas with bilateral projections.

Words in 10 conditions (five linguistic variables x two visual fields) were used: 1) 16 monosyllabic nouns and verbs, 2) the same nouns and verbs with inflectional suffixes (i.e. the regular plural ending **-s**, and the third person singular present tense ending **-s**), 3) the same nouns and verbs with derivational suffixes **-less** and **-able** (these suffixes derive adjectives from the base nouns and verbs), 4) long, low frequency adjectives matched for log word frequency and word length (both letters and syllables) with the derived adjectives of condition three, and 5) short, high frequency adjectives matched for log frequency and word length (letters and syllables) with the words in conditions one and two. Word length varied from three letters (one syllable) to ten letters (two or three syllables). The words in each condition are shown in Table I along with the log frequency of each as derived from the word frequency list of Francis and Kucera (1982). Some low frequency words used either did not occur in this list or occurred only once in over a million words.

Subjects were seated in a comfortable chair 350mm from the screen. Eye distance was maintained by having the subjects rest their foreheads against a cushioned headrest board attached to the top of the monitor. Subjects were cautioned not to lean too heavily upon the headrest to avoid producing additional

Simple Nouns & Verbs		Inflected Nouns & Verbs		Derived Adjectives	
MEAT	1.65	MEATS	1.08	MEATLESS	0
MOOD	1.57	MOODS	0.90	MOODLESS	0
TREE	1.76	TREES	2.00	TREELESS	0
LAW	2.48	LAWS	1.94	LAWLESS	0
DOOR	2.49	DOORS	1.56	DOORLESS	0
FRIEND	2.11	FRIENDS	2.21	FRIENDLESS	0
SONG	1.83	SONGS	1.77	SONGLESS	0
LEG	1.76	LEGS	1.83	LEGLESS	0
WEAR	1.51	WEARS	0.78	WEARABLE	0
BRING	2.20	BRINGS	1.60	BRINGABLE	0
GROW	1.80	GROWS	1.34	GROWABLE	0
SPEND	1.72	SPENDS	0.90	SPENDABLE	0
LEARN	1.92	LEARNS	1.00	LEARNABLE	0
KEEP	2.42	KEEPS	1.28	KEEPABLE	0
ASK	2.11	ASKS	1.26	ASKABLE	0
SERVE	2.03	SERVES	1.57	SERVABLE	0
mean log frequency	1.96		1.44		0

Short, High Frequency Adjectives		Long, Low Frequency Adjectives	
FRESH	1.91	INTREPID	0
RICH	1.85	VALIANT	0
VAST	1.79	SHODDY	0
SOFT	1.78	FLAXEN	0
WILD	1.73	PENSIVE	0
GREEN	1.93	BENIGN	0
THIN	1.95	JUBILANT	0.30
MERE	1.67	PRISTINE	0.30
mean log frequency	1.83		0.06

Table I: The materials for the five linguistic conditions with log word frequency.

neck muscle artifact in the signal. Subjects were instructed to remain as still as possible to avoid producing muscle artifacts, and they were instructed to avoid looking at the words. Subjects were regularly asked if they needed a break to drink some water, stretch, etc. to avoid the possibility of fatigue introducing undesirable variation in the data. This was particularly important, as the experimental session often took over five hours (including time for electrode application and removal). The first block of words presented to the subject served as a practice block, unless the subject excelled from the outset, in which case, it was included in the actual data.

The task was to name the word during the interval after its disappearance and before the next word appeared. A naming task was chosen in the hopes of avoiding the potentially longer semantic processing that may be involved in a lexical decision task, and because it has been suggested that the naming task only involves lexical access without extensive processing or influences of such variables as imageability (McMullen and Bryden 1987).

3.4. Data Analysis

The Bio-Logic Brain Atlas automatically averages EEG samples as they are recorded. This results in some noise being introduced into the averaged data, as waveforms for misreported words are averaged in also. However, a large sample size for each subject may partially compensate for this noise (provided that the across subjects error rate is not too high). Averaging across subjects, too, should help correct this problem. Averaging is used to eliminate muscle artifact noise and other non-task-relevant noise from the signal. It is assumed that the processing remains time-locked to the stimulus presentation across the presentation period (Picton and Stuss 1983; Regan 1989).

For each subject, all stimulus blocks of each condition were averaged together using the Bio-Logic Bank Mathematics Package of the Brain Atlas. No facility was available for recording trials individually. There were 32 words presented per block and up to 16 samples were collected. Two linguistic conditions were presented per block to each visual field, giving four experimental conditions per block overall. Data were collected from only two conditions per block because the Bio-Logic unit used has sufficient memory to average only two conditions at once. The results of these averages were then transformed using the Brain Atlas version of the Hjorth source derivation technique for topographic electrode data described by Hjorth (1980). This technique "preserves all information available from the International 10-20 system, in spite of the simplified presentation of local activity in the scalp field" (Hjorth 1980). Source derivation involves the subtraction of "overlap" from surrounding electrodes in the signal recorded from one electrode. This was done for each subject, when the performance of individuals was of concern, as well as for the six subject total.

Waveforms were then compared between conditions for each subject and across subjects. Comparisons included latency and amplitude measures, as are traditionally made for small arrays. The waveforms of greatest interest in this study were those considered least likely to contain muscle artifact, namely: F3, F4, C3, C4, P3, P4, Fz, Cz, and Pz. In addition to the multichannel waveform data, evoked potential maps of interpolated inter-electrode data were also used to better display spatial distribution and dynamics. These analyses were both performed after the Hjorth source derivation transformation was applied to reveal a more accurate image of the localization of the potential sources.

4. Results and Discussion

4.1. Accuracy Data

The accuracy results reported here are for five subjects.[1] The data described are for the words included in the waveform results. The accuracy data from the filler conditions of each block are not included because the possibility that eye movements could have occurred is more difficult to exclude, as the automatic artifact rejection did not affect these items. When an item in the experimental set was rejected due to artifact, the item was recorded as an error regardless of whether the word was reported correctly or not. Also, words that were reported without the suffix (if they were in the morphologically complex conditions) were not counted as correct responses even if the root (noun or verb) morpheme was reported correctly. This did occur for some subjects several times during the experimental session, particularly when the subject began to tire; however, the data on the frequency of this error have not yet been analyzed.

It might also be noted that some subjects reported words in the same syntactic category as that of the experimental items, but these were also counted as errors due to the fact that they were not identical with the stimuli presented. This type of error may be a result of the subject recognizing that many of the words in a block of stimuli belong to a particular syntactic category and guessing a word that belongs to that category, or that the subject has extracted some syntactic information from stimuli but not a sufficient amount of orthographic/phonological information to make a correct response. Further study is needed to select between these alternatives.

Figure 1 reveals a considerable difference in response accuracy across the two visual fields. Overall, the left visual field percent response accuracy is 14%

---

[1] A data file for one further subject (TB) recording the actual words presented to this subject was lost, making it impossible to assess the accuracy of the subject's responses. It was the experimenter's impression (based on the written record of subject's responses) that the responses given were highly accurate throughout the experimental session (i.e., across conditions).

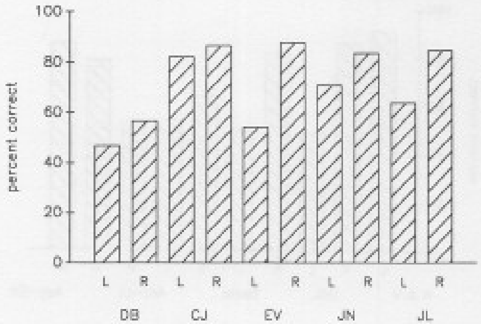


Fig. 1: Accuracy by Visual Field for each subject.

less than that of the right visual field. The direction of this difference is consistent with that reported by other investigators in previous divided visual field studies (Todorovic 1988; Beaumont 1982). Note also that the direction of the difference in visual field accuracy is the same for all subjects (LVF < RVF), although it varies greatly in magnitude -- from 4% (CJ) to 33.7% (EV). This suggests that this effect is reliable.

Subjects reported greater ease in reading words presented to the right visual field and some reported greater ease in reading the initial letter of the words. This may be due to a task-specific strategy adopted by some, if not all, of the subjects, namely: subjects appeared to attend to whatever cues they could to identify the word with the least effort and the highest accuracy. Once subjects noticed that the initial letters could be used to identify a unique word in a block of stimuli[2], they used this information (perhaps along with word shape/length information) to guess the word they had just seen. Unfortunately, the additional delay provided by the wait for the naming cue may have allowed time for a scan of memory for an appropriate match with the fragmentary data extracted from the flashed word. If words were presented several times to the same visual field (and this did happen due to the random presentation), additional information about a word could have been obtained by the occasional saccade to the word site in preparation for the appearance of the next word, if a word did then appear at the expected site. The artifact reject did sometimes reject waveforms that followed multiple presentations to the same visual field.[3]

Evidence of guessing may also be seen in the frequent occurrence (at least for some subjects) of mistakes such as reporting "mood" for "door" (both having "oo" medially), "learn" for "wear" (both having "ea" medially), "meat" for "mood" (both with an initial "m"), "leg" for "law" (both with an initial "l"), etc. Some

[2] Note that the same 16 monosyllabic nouns and verbs formed the basis for three experimental conditions. There was thus considerable repeated use of each stem.

[3] It is perhaps also worth mentioning that Tomlinson-Keasey et al. (1983) found that the initial letter of a word "does not play a critical role in word recognition" in tachistoscopic lateralized presentations.

subjects also reported that the words seemed to be displayed for a much shorter duration when a new condition was used; this was especially true for low frequency adjectives and long words.

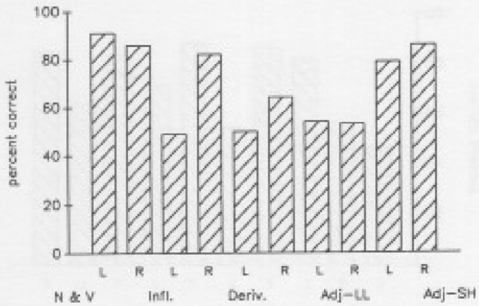


Fig. 2: Accuracy by Visual Field and Experimental Condition averaged across all subjects.

Accuracy results for the two visual fields for each experimental condition are displayed in Figure 2. In the simple nouns and verbs condition, there was little difference (5%) between visual fields, and what difference there was favored the LVF. This is perhaps small enough to be noise; indeed, as we will see, not all subjects show this LVF advantage. The inflected nouns and verbs condition shows the largest visual field difference (33%), with the LVF much lower, and the RVF at nearly the same level of accuracy as seen in the simple nouns and verbs condition. The derived adjectives (formed from the simple nouns and verbs) likewise show lower accuracy in the LVF, but with the performance for the RVF below the level for the inflected forms and monomorphemic forms.

This pattern for the suffixed words (i.e. the RVF superiority and LVF at lower accuracy is not surprising if the left hemisphere is handicapped by either a lack of function morphemes in its lexicon, or a lack of the rules for processing suffixes. An acuity gradient is not likely to be a reasonable explanation for these results, as the LVF suffixes are closer to the higher acuity regions of the visual field than those in the RVF, unless the subjects used a combination of length cues and distinctive letters (i.e. initial letters or medial "oo," etc.); the latter explanation seems unlikely, however. Long, low frequency adjectives occur at low accuracy

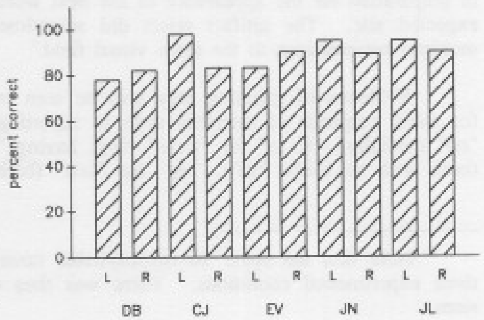


Fig. 3: Accuracy by Visual Field in the Nouns and Verbs condition for each subject.

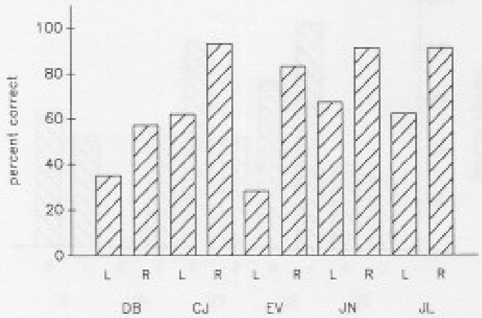


Fig. 4: Accuracy by Visual Field in the Inflected Nouns and Verbs condition for each subject.

levels regardless of visual field of presentation. This is again not surprising. However, the short, high frequency adjectives again show little difference across visual fields, with the RVF only 7% better than the LVF. These results are similar to those for the other monomorphemic, short, high frequency content words (i.e. the simple nouns and verbs condition); indeed, for the RVF the accuracy is identical for these conditions.

Data for the Nouns and Verbs condition for individual subjects are displayed in Figure 3. Little variation occurs in the accuracy of response for this condition for any subject apart from, perhaps, subject CJ who was 15% better in the LVF than in the RVF.

The Inflected Nouns and Verbs condition for individuals is summarized in Figure 4. A large RVF superiority was seen for all subjects ranging from 22% (DB) to 55% (EV) more accurate in the RVF than in the LVF. A Wilcoxson-signed test showed a significant ($p<.05$) tendency of the LVF to demonstrate lower accuracy than the RVF.

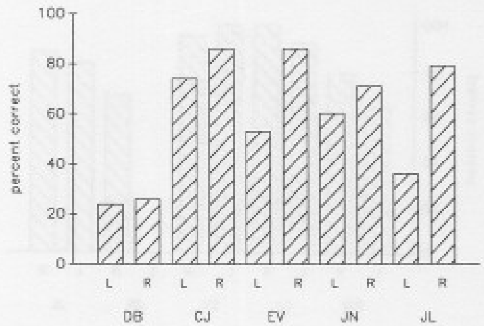


Fig. 5: Accuracy by Visual Field in the Derived Adjectives condition for each subject.

Individual results for the Derived Adjectives appear in Figure 5. Again there is a RVF superiority for all subjects, and LVF accuracy is near 50% except for CJ. The RVF accuracy is not far below that of the simple nouns and verbs condition except for DB.

The individual data for the Long, Low Frequency Adjectives condition appear in Figure 6. LVF accuracy in this condition was below 50% for all subjects except for one (CJ), and RVF accuracy was generally low as well except for EV and perhaps

CJ. This may be due to the very low frequency of these words; in effect, these were long non-words for several subjects. JN was unable to complete enough of the experiment to obtain any RVF data for this condition. It should be noted that the RVF accuracy is below that of the previous condition, derived adjectives, for some subjects. This condition was used to reveal whether adjectives derived from nouns and verbs would be processed in the same manner as the monomorphemic adjectives matched for length and word frequency. These results suggest that, at least for the left hemisphere, these adjectives are treated differently. It may be that the difference is due to no more than the lower frequency and greater difficulty of coping with the long, low frequency adjectives. Arguably, some of the derived adjectives were also non-words for some of the subjects; nevertheless, the derived adjectives were higher frequency forms due to the higher frequency of their noun and verb base forms.

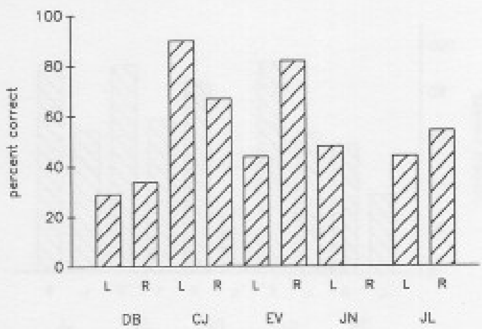


Fig. 6: Accuracy by Visual Field in the Long, Low Frequency Adjectives condition for each subject.

Results for the Short, High Frequency Adjectives condition for each subject are summarized by Visual Field in Figure 7. The pattern of this condition is not too different from that of the simple nouns and verbs condition as one would expect. Again, as in the last condition, JN was unable to complete the experiment, thus her data for the LVF is missing for this condition.

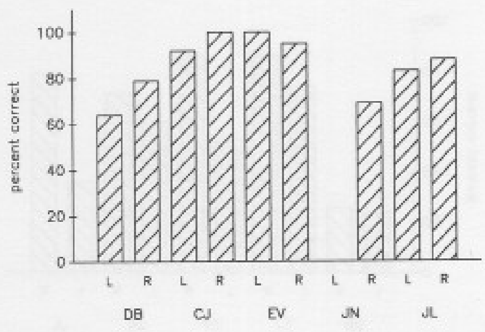


Fig. 7: Accuracy by Visual Field for the Short, High Frequency Adjectives condition for each subject.

4.2 Waveform and Spatial Distribution Variation

Waveform data from six subjects were analyzed. Subject TB's waveform data were included although the accuracy data for this subject were lost (see discussion above). Because of this small sample size and a lack of variance statistics for the waveform data reported below, caution must be exercised in interpreting these data. Because the monomorphemic adjective conditions were not crucial to the hypothesis and the data from these conditions were especially sparse, only six experimental conditions were analyzed: nouns and verbs (LVF and RVF), inflected nouns and verbs (LVF and RVF), and derived adjectives (LVF and RVF).

A sample of the waveform data is displayed in Figure 8. The cursor time was 100 msec ahead (± 976 μsec) of the actual initiation of stimulus presentation due to the ERP baseline collected before each trial. Data points were separated by 4 msec intervals, and voltages and latencies were taken from cursor locations. Waveform variation evident in this sample could be the result of variation in sample size or condition, variation in accuracy in different conditions, or any of a number of other factors.

A number of features of the waveform data seem worthy of consideration.[4] There was a P320 prominence that was clearly lateralized, with peak activity at P3 or P4 dependent upon visual field of presentation. This effect was evident across linguistic conditions. There was also a P333 peak centered at Pz which extended laterally toward P3 and P4. This peak sometimes remained more lateralized toward the hemisphere that initially processed the stimulus. Either of these effects may be a manifestation of the well-known P300 associated with visual processing.

At the Cz locus there was an N348 effect evident across all six of the analyzed conditions. This event seemed to spread over the interval from 300 to 600 msec post stimulus. Mean amplitude of the difference for this effect was -3.73 μV.

There was also an effect that appeared to be a manifestation of the phenomenon known as Contingent Negative Variation, or CNV, at F3, F4, and most especially at Fz (over the interval 400-900 msec). This may be a slow wave correlate of motor planning for the ensuing utterance of the target word. Orgogozo and Larsen (1979) report a significant increase in cerebral blood flow in the dorsomedial frontal lobe (i.e. the supplementary motor area) during vocalization. They claim that it is likely important in initiating and controlling some voluntary motor activities in man and they suggest that it may act as a higher order motor center. Regan (1989) also reports that neurons in the medial and lateral premotor cortex will become more active seconds prior to the initiation of a motor act, and in the premotor cortex changes in activity can anticipate motor activity by several minutes. He also states that the supplementary motor area appears to be crucial "in

---

[4] As is conventional in the ERP literature, these features will be identified by polarity and latency. Thus, P300 is a positive-going wave that diverges from some reference wave (e.g., that derived from some other experimental condition) at about 300 msec after stimulus onset. P300 at Fz is a peak that occurs at an electrode on the scalp over the frontal lobe at the zenith. Scalp site P3 is over the parietal lobe on the left side while P4 is over the corresponding site on the right side.

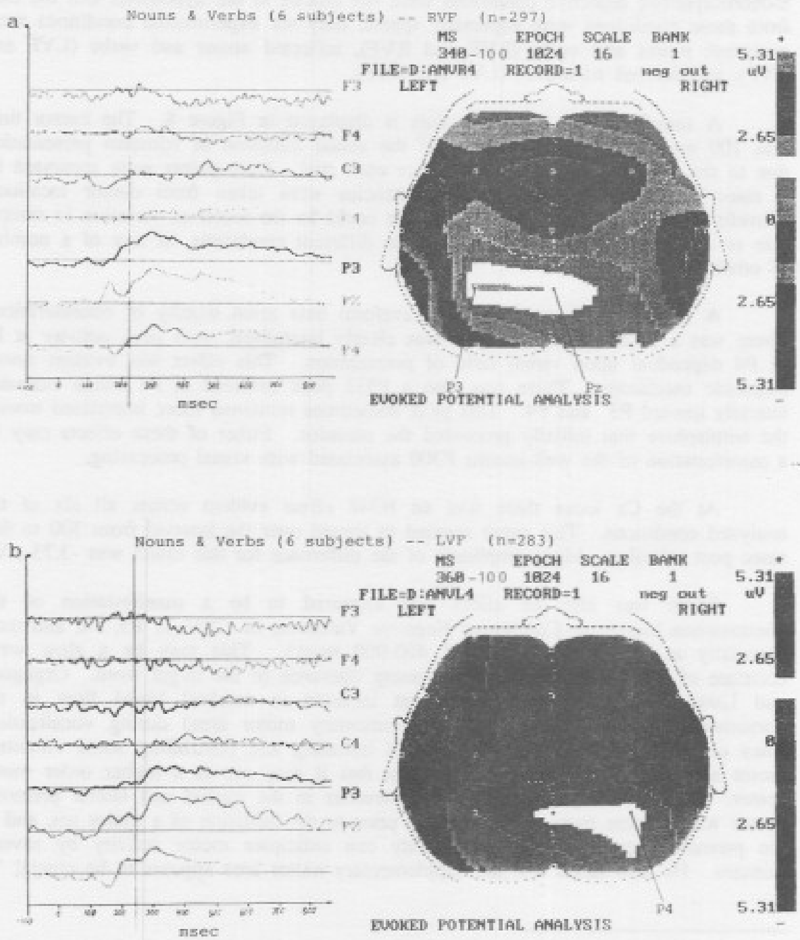


Fig. 8: Waveform data from the Nouns and Verbs condition for (a) the RVF (n=297) and (b) the LVF (n=283), for six subjects.

preparatory processes leading to initiation or suppression of movement in response to sensory input."

The latencies of certain peaks over parietal sites is of some interest. With RVF presentation, P3 mean latencies precede P4 latencies by 28 msec (n=773). LVF presentation, however, results in P4 mean peak latencies preceding P3 mean latencies by a mere 8 msec. (n=734). It is perhaps worth noting that at all electrode sites for RVF presentations the morphologically complex forms have longer latencies (especially for Pz and P4). Yet for LVF presentations, P3 and P4 have nearly identical peak latencies (except for the inflected forms). The greater delay for P4 and Pz peaks for RVF presentation may be a reflection of the initial processing by the hemisphere receiving the stimulus. The disproportionate delay at these electrode sites for the morphologically complex stimuli may be an indication of the increased difficulty of processing linguistically complex stimuli in the right hemisphere.

Analysis of peak amplitudes at the parietal sites can help assess some of the effects reported above. With RVF presentation, P3 mean peak amplitudes exceed P4 mean peak amplitudes by 1.12 μV (and Pz exceeds P3 by 0.42 μV). LVF presentation results in a P4 mean peak amplitude exceeding the P3 mean peak amplitude by 0.35 μV (and Pz exceeds P4 by 1.37 μV). The amplitude data across linguistic conditions are suggestive, but hardly unequivocal. It does appear that the visual field receiving initial stimulation has the higher peak amplitude. This is especially true of words presented to the RVF and thus processed by the left hemisphere. One might speculate that the higher amplitude of the RVF (left hemisphere) peaks results from the simultaneous activation of the greater number of neurons involved in processing linguistic stimuli in that hemisphere relative to the smaller homologous cortical areas in the right hemisphere.

5. Discussion

Though intriguing, the ERP data from this study is less reliable than the accuracy data, which seem best able to support firm judgments concerning hemispheric differences in morphological processing. This is because the ERP data is compromised by low and variable accuracy in several experimental conditions, by variation in sample sizes, and by the small number of subjects.

One observation that can be made from the ERP data, however, concerns the lateralized positive peaks that appear at P3 and P4 before the positive peak at Pz. These peaks are the strongest in the data apart from the negative peak at Cz and the positive peak at Pz. Interestingly, these peaks arise late and are lateralized to the initially stimulated hemisphere until well after initial lobe processing would be expected to be complete. Note that the visual evoked potential corresponding to processing by the striate cortex occurs around 100 msec after stimulus presentation, depending on several experimental variables (Regan 1989) - which is approximately 200 msec before these peaks observed here reach maximum. What cognitive process(es) these peaks represent cannot be ascertained from these data, but we might speculate that they represent processing in the parietal-occipital-temporal association cortex, or perhaps, in the plana temporale (the BEAM-Hjorth localization is consistent with these possibilities). Either of these possibilities would suggest linguistic processing of the stimuli. Garnsey (1985) and others both report peaks in this general area and latency in response to linguistic stimuli.

Overall, the accuracy results are consistent with the conclusion that the right hemisphere has a lexicon which can process short, high frequency monomorphemic nouns and verbs. As for its morphological capacities, however, the accuracy data suggests that the right hemisphere by itself is deficient in its processing of the same words when inflectional and derivational suffixes are attached.

The BEAM ERP results suggest that a larger sample is needed before the relevant neurolinguistic questions can be seriously addressed. The problem of low accuracy for the LVF conditions is more difficult. Differences between linguistic conditions may be obscured within ERP data by coexisting differences in accuracy, even with larger sample sizes. Thus, it may be that only techniques such as intracerebral microelectrode arrays, or perhaps some refined version of magnetoencephalography, will support definitive claims about the neurophysiology and neuroanatomy underlying the linguistic behaviors observed here.

Acknowledgments

Several individuals have been very helpful (and patient beyond measure) while I have been working on this thesis, and I would like to thank as many as my memory allows. Wayne Cowart and Michael Torello have been invaluable in keeping me on course, and they have both been sources of much needed advice and encouragement. Tracey has not only spent many hours helping me type, but also has patiently listened to my ideas and provided support throughout. My parents have provided financial support when I've run low. Phil Enny, Gail Nakamora, and Amy Flanary have provided much needed technical support and information for which I am deeply grateful. This work was also supported in part by a grant from the National Institutes of Health (1 R01 NS22606-01) to W. Cowart, and a research grant from the OSU Department of Linguistics.

References

Allison, T., C. C. Wood, and G. McCarthy. 1986. The Central Nervous System. In M. G. H. Coles, E. Donchin, and S. W. Porges (Eds.) *Psychophysiology: Systems, Processes, and Applications*. New York: The Guilford Press, 5-25.

Anderson, D. R. 1982. *Testing the Field of Vision*. St. Louis: The C. V. Mosby Co.

Beaton, A. 1985. *Left Side, Right Side: A Review of Laterality Research.* New Haven: Yale University Press.

Beaumont, J. G. 1982. Studies with Verbal Stimuli. In J. G. Beaumont (Ed.) *Divided Visual Field Studies of Cerebral Organization.* London: Academic Press, 57-86.

Borden, G. J. and K. S. Harris. 1984. *Speech Science Primer.* Baltimore: Williams & Wilkins.

Braun, D. and B. G. Breitmeyer. 1988. Relationship between Directed Visual Attention and Saccadic Reaction Times. *Experimental Brain Research* **73:** 546-552.

Caplan, D. 1987. *Neurolinguistics and Linguistic Aphasiology: An Introduction.* New York: Cambridge University Press.

Chiarello, C. 1988. More on Words, Hemifields, and Hemispheres: A Reply to Schwartz and Kirsner. *Brain and Cognition* **7:** 394-401.

Chiarello, C., J. Senehi and M. Soulier. 1986. Viewing Conditions and Hemisphere Asymmetry for the Lexical Decision. *Neuropsychologia* **24:** 521-529.

Ducati, A., E. Fava, and E. D. F. Motti. 1988. Neuronal Generators of the Visual Evoked Potentials: Intracerebral Recording in Awake Humans. *EEG and Clinical Neurophysiology* **71:** 89-99.

Durrant, J. D. and J. H. Lovrinic. 1984. *Bases of Hearing Science.* Baltimore: Williams & Wilkins.

Francis, W. N. and H. Kucera. 1982. *Frequency Analysis of English Usage: Lexicon and Grammar.* Boston: Houghton Mifflin Co.

Garnsey, S. M. 1985. Function Words and Content Words: Reaction Time and Evoked Potential Measures of Word Recognition. *Cognitive Science Technical Report.* University of Rochester **29**

Geschwind, N. and W. Levitsky. 1968. Human Brain: Left-Right Asymmetries in Temporal Speech Region. *Science* **161:** 186-187.

Goodall, G. 1984. Morphological Complexity and Cerebral Lateralization. *Neuropsychologia* **22:** 375-380.

Harner, P. F. and T. Sannit. 1974. *A Review of the International Ten-Twenty System of Electrode Placement.* Grass Instrument Co.

Hecaen, H. and M. L. Albert. 1978. *Human Neuropsychology*. New York: John Wiley & Sons.

Hjorth, B. 1980. Source Derivation Simplifies Topographical EEG Interpretation. *American Journal of EEG Technology* **20**: 121-132.

Kutas, M., C. Van Petten, and M. Besson. 1988. Event-Related Potential Asymmetries During the Reading of Sentences. *EEG and Clinical Neurophysiology* **69**: 218-233.

Leventhal, A. G., S. J. Ault and D. J. Vitek. 1988. The Nasotemporal Division in Primate Retina: The Neural Bases of Macular Sparing and Spilting. *Science* **240**: 66-67.

McKeever, W. F. 1986. Tachistoscopic Methods in Neuropsychology. In H. J. Hannay (Ed.) *Experimental Techniques in Human Neuropsychology*. New York: Oxford University Press, 167-211.

McMullen, P. A. and M. P. Bryden. 1987. The Effects of Word Imageability and Frequency on Hemispheric Asymmetry in Lexical Decisions. *Brain and Language* **31**: 11-25.

Miceli, G. and A. Caramazza. 1988. Dissociation of Inflectional and Derivational Morphology. *Brain and Language* **35**: 24-65.

Millar, J. M. and H. A. Whitaker. 1983. The Right Hemisphere's Contribution to Language: A Review of Evidence from Brain-Damaged Subjects. In S. J. Segalowitz (Ed.) *Language Functions and Brain Organization*. New York: Academic Press, 87-113.

Molfese, D. L. 1983. Event-Related Potentials and Language Processes. In A. W. K. Gaillard and W. Ritter (Eds.) *Tutorials in Event-Related Potential Research: Endogenous Components*. Amsterdam: North-Holland Publishing Co., 345-368.

Orgogozo, J. M. and B. Larsen. 1979. Activation of the Supplementary Motor Area During Voluntary Movement in Man Suggests It Works as a Supramotor Area. *Science* **206**: 847-850.

Picton, T. W. and D. T. Stuss. 1984. Event-Related Potentials in the Study of Speech and Language: A Critical Review. In D. Caplan, A. R. Lecours, and A. Smith (Eds.) *Biological Perspectives on Language*. Cambridge, MA: The MIT Press, 303-360.

Petersen, S. E., P. T. Fox, M. I. Posner, M. Mintun, and M. E. Raichle. 1988. Positron Emission Tomographic Studies of the Cortical Anatomy of Single-Word Processing. *Nature* **331**: 585-589.

Regan, D. 1989. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. New York: Elsevier.

Samar, V. J. and G. P. Berent. 1986. The Syntactic Priming Effect: Evoked Response Evidence for a Prelexical Locus. *Brain and Language* **28:** 250-272.

Scheible, A. B., I. Fried, L. Paul, A. Forsythe, U. Tomiyasu, A. Wechsler, A. Kao, and J. Slotnik. 1985. Differentiating Characteristics of the Human Speech Cortex: A Quantitive Golgi Study. In D. F. Benson and E. Zaidel (Eds.) *The Dual Brain: Hemispheric Specialization in Humans*. New York: The Guilford Press, 65-74.

Todorovic, D. 1988. Hemispheric Differences in Case Processing. *Brain and Language* **33:** 365-389.

Tomlinson-Keasey, C., A. Brewer, and K. Huffman. 1983. The Importance of Being First: An Analysis of Tachistoscopic Presentations of Words. *Cortex* **19:** 309-325.

Witelson, S. F. 1983. Bumps on the Brain: Right-Left Anatomic Asymmetry as a Key to Functional Lateralization. In S. J. Segalowitz (Ed.) *Language Functions and Brain Organization*. New York: Academic Press, 117-144.

Young, A. W. 1982. Methodological and Theoretical Bases of Visual Hemifield Studies. In J. G. Beaumont (Ed.) *Divided Visual Field Studies of Cerebral Organization*. London: Academic Press, 11-27.

Zaidel, E. 1978. Lexical Organization in the Right Hemisphere. In P. A. Buser and A. Rougel-Buser (Eds.) *Cerebral Correlates of Conscious Experience*. Amsterdam: Elsevier/North-Holland Biomedical Press, 177-197.

OHIO STATE UNIVERSITY WORKING PAPERS IN LINGUISTICS **38** (24-37)

A Visual Half Field Study of Sentence Processing

James Beale
Cornell University
and
Wayne Cowart
University of Southern Maine

Abstract: This study attempts to exploit visual half field presentations of words in sentence contexts as an aid in the analysis of a cognitive function related to anaphoric processing. The function in question assimilates several kinds of information to resolve certain syntactic ambiguities. The experiments address the question whether all aspects of this linguistically complex function are supported by the language dominant left hemisphere. The evidence suggests that the anaphoric function is bihemispheric, i.e., that subfunctions supported by both hemispheres play a crucial role.

1. Introduction

A traditional concern among neuropsychologists and neurolinguists has been to associate cognitive functions involved in language with specific regions of the brain defined in terms of gross anatomy, differences of cell structure and patterns of interconnection, or possibly differences in local neurochemistry. The project discussed here is **not** an attempt to further this important enterprise, though it borrows neuropsychological methods and, if successful, would contribute to neuropsychological theory.

The goal of the exploratory research discussed in this paper is to exploit a feature of the mammalian visual system to reveal something of the logical structure of the processes that underlie sentence comprehension.

It might be thought that for an enterprise of this sort to succeed, it is necessary to have an antecedently established theory of how the neuropsychological manipulation we intend to use is related to language processing. While this is surely desirable, we do not believe it is necessary. We will assume that for a neuropsychological manipulation to be useful in the analysis of language processing it is only necessary that it produce some stable, repeatable effects that can be interpreted in linguistic terms. For example, suppose we did not know that there is something special about the left posterior frontal lobe, with respect to language. Suppose further that by applying a paste to the left scalp just forward of and above the ear we could temporarily produce some of the effects typically associated with Broca's aphasia. Though such a result would tell us something important about the brain, it would also be interesting from a psycholinguistic point of view, completely independently of its impact on neuropsychology. That is, the difference between the language functions a subject was able to exercise with and without the strategically placed scalp paste would be informative not only about the role of the underlying cortical tissue, but also with respect to the logical organization of the processing system. In short, we hold that neuropsychological manipulations can, at least sometimes, be psycholinguistically interesting, even if their impact on the central nervous system is not well understood within neuropsychological theory.

Under these assumptions, it is of some interest that there is a standard technology in neuropsychology by which one may force one hemisphere to do all the initial processing of a linguistic stimulus. This technology is that of visual half-field presentations. The fundamental fact is that the mammalian visual system is organized in such a way that visual stimuli presented in the left visual field[1] are directly presented only to the visual areas of the right hemisphere while stimuli in the right visual field are presented only to the left hemisphere (Kelly 1985). This split in the visual system has been useful in discriminating the psychological functions of the two hemispheres (see Beaumont 1982).

There are, of course, structures, most notably the corpus callosum, that allow for rapid communication between the hemispheres. It appears, however, that even in the visual cortex the information that passes from one hemisphere to another is an interpreted, abstracted copy of the original stimulus, not a 'photographic' image of it. Thus, what one hemisphere is able to relay to another is limited to those properties of the stimulus that can be recovered by the resources that participate in the initial analysis of the stimulus, up to the level of abstraction relevant to interhemispheric communication. Furthermore, any communication between the hemispheres requires time, which may alter the way various kinds of information are integrated in a hemisphere that is applying internal resources to information received from the other hemisphere.

These facts suggest a case parallel to the paste experiment described above. Suppose that visual target words are presented following auditorily presented sentence contexts. Further suppose that these visual stimuli are presented exclusively to one or the other hemisphere for initial visual processing. If this manipulation has a linguistically interpretable result, then this technique can be useful for revealing the logical structure of the psychological processes of language regardless of what may or may not be independently known about linguistic differences between the two hemispheres.

Of course, there is perhaps more evidence for linguistic differences between the hemispheres than for any other difference (see Caplan 1987). This fund of information can be useful in attempts to interpret differences that might be detected, but, as we have argued, this need not be seen as a precondition for the successful exploitation of visual half-field technology. For present purposes, prior work on linguistic differences between the hemispheres contributes two things. First, the

---

[1] Roughly the half of the visual field that falls to the left of the point on which the eyes are currently focused.

elementary observation that in most adults the hemispheres are not equally able to support linguistic function, and that the left hemisphere plays the more central role in most adults. Secondly, this special role of the left hemisphere does not exclude participation of the right hemisphere, though the extent and nature of the contributions of the right hemisphere are not clearly established. This set of elementary observations opens the possibility that any result that exploits a rich complex of linguistic functions might in some way depend upon activity in both hemispheres. This in turn raises the possibility that procedures that in some fashion intervene in or disrupt normal modes of collaboration between the hemispheres may be able to reveal something of the way the logical problem the task presents is attacked by different functions. Looking at it another way, presenting a target to one hemisphere may have the effect of subtracting out, or limiting the participation of, resources resident in the other hemisphere. The effects of this subtraction should reveal something of how the resources that support the task are structured.

The work reported here constitutes a preliminary attempt to apply visual half-field technology to a question in sentence processing.

2. The Pronoun Bias Effect and its analysis

The background to the present investigation lies in previous work reported in Cowart and Cairns (1987) that argued for the presence of a notably asemantic antecedent-finding mechanism within the syntactic processing system. Cowart and Cairns reported two key results. On each trial the subject was required to quickly read aloud a verb form presented immediately at the offset of an auditorily presented sentence fragment. The evidence of several studies indicates that subjects spontaneously attempt to integrate the visual target word with the sentence fragment by the time they produce their voice response, though there is nothing in the procedure that requires this. The first result was that when subjects were required to read **is** at the end of fragments such as (1), response times were slower when the subject of the fragment was **they** rather than a plural NP such as **the birds**.

(1) As **the birds/they** soar gracefully over the field, flying kites. . .

This apparently reflects a tendency for subjects to interpret the ambiguous expression **flying kites** as a plural NP when **they** appears in prior context; this provides an antecedent for the otherwise uninterpreted pronoun, but yields an agreement anomaly when the subject is forced to read **is**. This effect of the pronoun subject of the clause on the subsequent reading of the verb was termed the Pronoun Bias Effect. The second result showed that in sentences such as (2), the Pronoun Bias Effect was not diminished when the grammatically optional interpretive link between the pronoun and the ambiguous expression produced an anomalous interpretation.

(2) If they eat a lot of oil, frying eggs. . .

Thus, it seemed that subjects made the referential link between the pronoun and the ambiguous expression on structural rather than semantic grounds, and even in spite of consequent semantic anomaly.

This phenomenon appears to be quite complex. It seems to involve solutions to at least three logically distinct problems, 1) the problem of finding an antecedent for the pronoun, 2) the problem of resolving the ambiguity of the ambiguous expression, and 3) the problem of assessing the agreement relation between the ambiguous expression and the following verb form (the target word).

On the evidence of Cowart and Cairns (1987) the mechanisms that deal with these three problems are able to operate and interact without allowing semantic or pragmatic concerns to block the coreference relation, at least as of the moment when subjects form their responses in that procedure[2]. What makes the result surprising is that general solutions to at least two of the problems enumerated above (antecedent finding and ambiguity resolution) would have to be sensitive to semantic and/or pragmatic information in many instances.

The interest of visual half field technology in relation to this problem is that it is not necessarily the case that all the linguistic cognitive resources the Pronoun Bias Effect engages are located in (or are most accessible via) the left hemisphere. As noted, both ambiguity resolution and antecedent finding commonly exploit semantic and pragmatic information relevant to the input, as well as syntactic analyses. On the general principle that the more complex a function is, the more of the brain's structure is engaged by its operation, these processes are good candidates for exploiting a broader range of brain structures. In particular, right hemisphere structures might be more heavily engaged than they would be if these processes were more narrowly syntactic in character. If the Pronoun Bias Effect is the product of some integration of resources across the hemispheres, then there is reason to hope that visual half field technology might be useful in beginning to tease apart those contributions.

There is also a somewhat more prosaic reason why the character of the Pronoun Bias Effect might change with lateralized presentations of the target words. In order for the effect to appear it is clearly essential that the subject identify the word presented and determine its morphological analysis (i.e., whether it is singular or plural). There is evidence that the two hemispheres differ in their ability to cope

---

[2] It should be noted that there is some evidence that semantic/pragmatic concerns may affect the subject's response. In particular, the magnitude of the Pronoun Bias Effect observed in the Anomalous Selectional condition in Cowart and Cairns' Experiment 2 is numerically much larger than in the Non-anomalous condition. The semantic/pragmatic anomaly thus seemed able to affect the subject's performance, but not in such a way as to block the coreference relation (without which there is no anomaly). Definitive evidence on this effect awaits further experimentation.

with these aspects of the task (see McAdams 1990, this volume, and references cited there). It may be that the Pronoun Bias Effect will differ by hemisphere of presentation because morphological information needed to implement it is extracted from the visual display less effectively by the right hemisphere than the left hemisphere.

One further aspect of the Pronoun Bias Effect should be mentioned. The discussion to here has viewed the Pronoun Bias Effect solely as a certain pattern of reaction time results. There is, however, a parallel phenomenon that has been observed with a judgment task. This phenomenon was first detected in a paper and pencil survey in which subjects were given printed forms containing sentence fragments such as those in (1) followed by two verb forms. The subject's task was to indicate which verb form made the better continuation of the fragment. When the subject of fragment was **you**, singular and plural forms were about equally acceptable, but there was a strong bias against singular verb forms when the subject was **they**. In Experiment 3 in Cowart and Cairns (1987) as part of each trial subjects indicated whether the verb form they had read aloud seemed to make a good continuation of the fragment. Judged acceptability of the verb form **is** declined dramatically with **they** subjects. Statistically, effects of these kinds have been markedly more robust than the reaction time results.

3. Experiment 1

The central problem addressed here is to discriminate the various cognitive resources that contribute to the Pronoun Bias Effect and to understand how their individual contributions are brought about and integrated. The specific prospect that motivates the present experiment, however, is the possibility that the effect might in some fashion depend upon higher order cognitive resources in both hemispheres. If both hemispheres do participate, then visual half field methods might be useful in addressing the more basic problem of analysis.

Even if the resources needed to implement the Pronoun Bias Effect are distributed across the two hemispheres, it is quite possible that the effect will persist in the face of lateralized presentations of the target word. There are many ways for this experiment to fail to detect bihemispheric participation even if that is the fact of the matter. Nonetheless, one possible outcome is that the pattern of reaction time effects that characterize the Pronoun Bias Effect[3] will not emerge at all with lateralized presentation. If this happens, and cannot be attributed to some uninteresting cause (e.g., low level visual system effects), then this will argue that the needed resources are bilaterally represented. There will then be grounds for using visual half field methods to try to understand how the component processes are differentiated, how they are distributed across the hemispheres, etc.

---

[3] That is, a slowing of naming responses to **is** when **they** rather than **you** (or another control NP) appears as subject of the context sentence fragment.

Another possible, perhaps more likely, outcome is that a pattern typical of the Pronoun Bias Effect will emerge when targets are presented to one hemisphere, but not the other. In this case it will matter a lot which hemisphere shows the effect. If the effect were apparent only with left visual field/right hemisphere presentations this would suggest that some crucial component of the processes supporting the effect is resident in the right hemisphere. Determining what the component was should lead to a better understanding of the processes that underlie the effect. The less interesting result would be a finding that the effect is apparent only with right visual field/left hemisphere presentations. This would neither add an interesting qualification to existing evidence locating language function primarily in the left hemisphere, nor would it encourage further work in this area with visual half field technology.

The basic methodological issue the experiments confront is the question whether interhemispheric communication will obliterate any hemisphere-specific effects. No attempts will be made here to correct for this problem. Further experiments could employ techniques designed to interfere with interhemispheric transfer of stimulus information.

3.1. Method

3.1.1 Design

The design of the present study is derived from that of Experiment 1 of Cowart and Cairns (1987), modified for visual half-field presentation. The design involves two within-subject factors: Pronoun (the context fragment contains either **you** or **they**), and Visual Field (Left vs. Right Visual Field). This design is then replicated with the target verbs **is** and **are**. Stimulus materials were the 72 sentence fragments used in Experiment 2 in Cowart and Cairns in the non-anomalous conditions. There are thus nine trials in each of four experimental conditions for two different target verbs. To reduce the likelihood of subjects anticipating the target verbs 156 fillers were also included. Fillers contained a variety of sentence structures and verbs; filler verbs were limited to high frequency 2- and 3-letter verbs (as found in Francis and Kucera 1982). Twelve of these fillers were used as training stimuli, leaving 216 trials with the ratio of fillers to experimentals at 2:1.

3.1.2 Procedure

Much current work on language-brain relationships makes some use of visual half-field presentations. However, the visual half field literature has little to say regarding the processing of words within a meaningful context, such as a sentence. The most relevant of this literature has focussed on lexical processing (e.g., Chiarello and Nuding 1987, Drews 1987, Howell and Bryden 1987, McAdams 1990, this volume). Apart from the addition of the auditory context, the procedures used here follow general practice in most respects.

All stimulus materials were recorded and digitized for computer presentation using OSU Linguistics Department computing facilities. Control of intonation was achieved by editing out the version of each VERBing NOUNs expression that was produced with **they** in the subject position of the initial subordinate clause. This same rendition of the ambiguous expression was then 'spliced' onto the end of both the **you** and **they** versions of the sentence fragment to insure that the two target verbs would be presented in the same intonational context regardless of what pronoun was used in the fragment.

An initial questionnaire was used to determine subjects' primary language, visual/auditory health, and handedness background (see below). Subjects were then seated in a booth approximately 30 to 45cm from a computer screen and given oral instructions describing the sequence of events on a typical trial. Subjects were asked to focus on a dot displayed at the center of the screen during the auditory presentation of the sentence fragment on each trial.

A computer presented auditory stimuli (sentence fragments) over a pair of earphones. Each sentence fragment was followed by lateral visual presentation of the target verb, 10 characters (26mm) to the left or right of the central dot; this placed the outer limits of the display at about 3 to 5 degrees of visual angle from the central dot. Exposure duration of the target verb was 100ms to insure that subjects did not have time to execute a saccade to bring the target word into foveal vision (Beaumont 1982). The computer timed the interval from the onset of the visual presentation of the target verb to the onset of the subject's phonation. Subjects then indicated whether or not the target verb provided a good continuation of the sentence fragment; this helped insure that subjects were attending to the sentence fragments. Response latency and judgements were recorded as dependent variables. This procedure required 45 minutes per subject.

3.1.3. Subjects

Recent work by Bever, Carrithers, Cowart and Townsend (1990) and Cowart (1988) suggests that, in addition to previously recognized effects of personal handedness, there may be language processing effects associated with differences in familial sinistrality. Familial sinistrality is present when one or more of an individual's blood relatives are left-handed. Possible effects of this sort have been discussed in the neurolinguistic literature for some time (see Caplan 1987, for a review). There have also been reports suggesting that sex affects aspects of lateralization (see Caplan 1987). To partly control for possible influences of these sorts, only right-handed males without familial sinistrality in their primary and secondary families were used for this study.

Nine subjects meeting these criteria were drawn from the OSU Psychology 100 Subject Pool, thus allowing one subject per experimental condition, with one duplication.

3.2 Results

Cowart and Cairns (1987) found that Pronoun Bias Effect was best viewed as an effect on naming responses to **is**, rather than as an effect on responses to **are** or an interaction across the two verbs. This analysis will, in the main, follow their practice of considering responses to the two verb forms separately.

The **is** data is summarized in Figure 1. On visual inspection it appears that a slowing of responses to **is** in the presence of **they** appears only with left visual field/right hemisphere presentations.

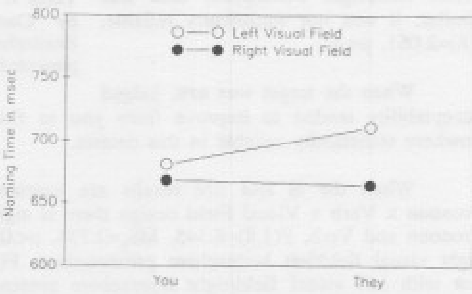


Fig. 1: Naming time to **is** by context pronoun and visual field in Experiment 1.

However, statistical analyses on these data reveal that they are highly variable and that the patterns they present are generally not reliable. In the **is** data there is neither a main effect of Pronoun nor an interaction between Pronoun and Visual Field. When the left visual field data is considered in isolation, the slowing of responses with **they** is significant, $F(1,8)=6.268$, $MS_e=646.5$, $p<.05$. There is also a main effect of Visual Field, $F(1,8)=8.828$, $MS_e=893.0$, $p<.02$, reflecting the overall slowing of responses to **is** with left visual field presentations. Little importance can be attached to these tests in the face of the nonsignificant interaction term.

In the **are** data there was a decline in naming times with **they** in context that seems consistent with earlier results. There is, however, no main effect of Pronoun, Visual Field, nor an interaction between these. The three way interaction among Pronoun, Verb, and Visual Field did not approach significance.

The judgment data reflecting subject's estimates of whether or not the verb form they read made a good continuation of the fragment they heard is summarized in Table I for both **is** and **are** targets. Visual inspection suggests that, much more than the naming responses, the judgment responses pattern very much like the results of earlier investigations. Relative to **you**, **they** reduces acceptability on trials with **is** targets and increases acceptability on trials with **are** targets.

The judgment data include a few outliers but no effort was made to correct for difficulties these may have introduced. The following tests must therefore be regarded with some caution.

When the target was **is** the decrease in acceptability from **you** to **they** contexts was reflected in a robust main effect of Pronoun, $F(1,8)=8.963$, $MS_e=3.375$, $p<.02$, which was mirrored by a significant effect of Pronoun in the data for the right visual field/left hemisphere, $t(8)=3.507$, $p<.01$. Though the pattern in the data for the left visual field/right hemisphere data was similar, it was not statistically reliable, $t(8)=2.061$, $p<.1$.

		you	they
Left	**is**	59	38
	are	22	36
Right	**is**	69	49
	are	32	40

Table I: Percent of cases judged acceptable by Context Pronoun, Target Verb and Hemisphere to which target was initially presented in Experiment 1.

When the target was **are**, judged acceptability tended to improve from **you** to **they** contexts, but these effects are nowhere statistically reliable in this dataset.

When the **is** and **are** results are examined together in the context of a Pronoun x Verb x Visual Field design there is significant overall interaction between Pronoun and Verb, $F(1,8)=6.345$, $MS_e=2.736$, $p<.05$. A similar result appeared with right visual field/left hemisphere presentations, $F(1,8)=6.066$, $MS_e=3.59$, $p<.05$, but not with left visual field/right hemisphere presentations, $F(1,8)=4.209$, $MS_e=3.194$, $p<.1$.

There was also an overall decrease in acceptability with all targets presented to the left visual field/right hemisphere relative to targets presented to right visual field/left hemisphere, $F(1,8)=7.84$, $MS_e=0.347$, $p<.05$.

3.3. Discussion

The best summary of the naming time results is simply that nothing happened. There is no replication of the Pronoun Bias Effect nor any convincing evidence of some other pattern of response. It is perhaps a little surprising that the data from the left visual field/right hemisphere comes closer to replicating the effect than does the data from the right visual field/left hemisphere, but this difference is buried in noise and unlikely to replicate.

The judgment data is markedly more stable and more intelligible as well. It shows a pattern that is straightforwardly analogous to the Pronoun Bias Effect. **Is** is judged a poorer completion in the presence of **they** than **you**. While there is a trend in the opposite direction with **are** targets, it is not significant (as in previous studies). The significant Verb x Pronoun interaction further reinforces the impression that these data replicate earlier findings.

The only evidence of interhemispheric differences is that some of the judgment effects are robust with targets presented to the left hemisphere, but not the right. But since the relevant interactions are not significant, there is no reliable evidence here of hemispheric differences.

4. Experiment 2

One possible account of the lack of hemispheric differences in Experiment 1 is that subjects were free to move their heads, which may have compromised the effectiveness of the visual half field manipulation. The second experiment replicates the first with improved procedure.

4.1. Method

4.1.1. Design and Procedure

The design of Experiment 2 is simply a replication of Experiment 1. To eliminate potentially confounding variables due to variations in the positioning of the subject's head, a viewing box was constructed for use in this study. This box places the screen at eye level and includes a headrest to minimize head movements. Distance to screen was thus held constant at 39cm. Target verbs were presented eight characters (21mm) to the left or right of the central dot, thus placing the outer limit of each target at 3 degrees of visual angle. Exposure duration of the target verb was increased to 150ms from 100ms because some subjects found the briefer displays of Experiment 1 hard to cope with.

4.1.2. Subjects

Eight subjects were drawn from the OSU Psychology 100 Subject Pool for this study. Subjects were once again right-handed males, without sinistral relatives.

4.2. Results

The **is** naming data are summarized in Figure 2. The patterns evident are unlike those of Experiment 1 and not like the Pronoun Bias Effect. But like Experiment 1, there is a great deal of variability in these data and the patterns are not reliable.

Tests on the **is** data show no significant main effects or interactions. The puzzling acceleration of naming time with

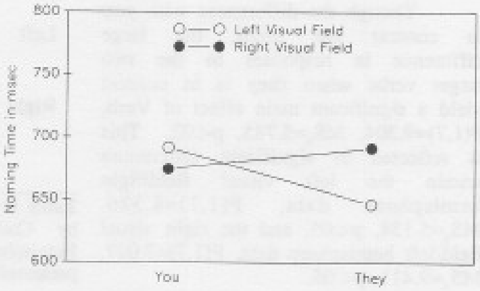


Fig. 2: Naming time to **is** by context pronoun and visual field in Experiment 2.

they in context in the left visual field/right hemisphere data yields a relatively large Pronoun x Visual Field interaction term, $F(1,7)=3.509$, $MS_e=2228$, $p=.1$. Within the left visual field/right hemisphere data the Pronoun effect is significant, $F(1,7)=6.915$, $MS_e=1190$, $p<.05$. This, however, is not a reliable measure in the absence of the interaction effect and is in any case exactly the opposite of what was observed in Experiment 1.

There were no significant effects or interactions in the **are** data or in analysis that combined **is** and **are** naming data.

The judgment data is again far more stable and intelligible. The results are summarized in Table 2. The pattern is like that of Experiment 1 except that the Pronoun effect seems larger with **are** targets and there is no evidence of the hemisphere differences seen in Experiment 1.

Within the **is** data there is a significant main effect of Pronoun, $F(1,7)=7.199$, $MS_e=1.567$, $p<.05$, indicating, as before, decreased acceptability with **they** in context rather than **you**.

Within the **are** data there is also a significant main effect of Pronoun, $F(1,7)=6.472$, $MS_e=3.786$, $p<.05$, reflecting an increase in acceptability with **they** rather than **you** in context. This effect is reliable in the data for the left visual field/right hemisphere, $t(7)=2.567$, $p<.05$, but not for the right visual field/left hemisphere, $t(7)=2.121$, $p<.1$.

An analysis that considers **is** and **are** data in the context of a Pronoun x Verb x Visual Field design shows a significant Pronoun x Verb interaction, $F(1,7)=14.121$, $MS_e=1.222$, $p<.01$, confirming that there is a reliable difference in the effects that the two pronouns exert on the two verbs. This effect is also independently significant in left visual field/right hemisphere data, $F(1,7)=10.343$, $MS_e=1.888$, $p<.05$, and the right visual field/left hemisphere data, $F(1,7)=7.631$, $MS_e=1.982$, $p<.05$.

Though the differences with **you** in context are small, the large difference in responses to the two target verbs when **they** is in context yield a significant main effect of Verb, $F(1,7)=9.304$, $MS_e=5.785$, $p<.02$. This is reflected in significant differences within the left visual field/right hemisphere data, $F(1,7)=8.326$, $MS_e=5.138$, $p<.05$, and the right visual field/left hemisphere data, $F(1,7)=7.027$, $MS_e=9.411$, $p<.05$.

		you	they
Left	**is**	64	50
	are	17	33
Right	**is**	61	49
	are	18	40

Table II: Percent of cases judged acceptable by Context Pronoun, Target Verb and Hemisphere to which target was initially presented in Experiment 2.

4.3. Discussion

The naming results for Experiment 2 are as unstable as those for Experiment 1 and contradict the faint hints of effects that were evident in the first experiment. The appropriate conclusion is that these experiments have produced no evidence of the Pronoun Bias Effect in the naming results. Perforce they have also not revealed any differences in naming performance with materials presented to the left vs. right hemisphere.

Like Experiment 1, the results of the judgment task in Experiment 2 were far more stable than those for the naming task. The judgment results produced a statistically significant reflection of the Pronoun Bias Effect, but no reliable evidence of hemispheric differences. There was a slight tendency for certain effects to be statistically more robust with left visual field/right hemisphere presentations, but this is the reverse of the slight hemispheric differences noted in the first experiment.

5. Conclusions

The default expectation for this set of experiments is that the Pronoun Bias Effect will be evident when the target words are presented to the right visual field/left hemisphere and that the effect will be diminished or absent with presentations to the left visual field/right hemisphere. This did not occur. Rather, the Pronoun Bias Effect, as a naming phenomenon, was completely extinguished.

The disappearance of the Pronoun Bias Effect might be attributed to methodological error of some kind were it not that the pattern of judgment effects reported in Cowart and Cairns (1987) was neatly and robustly replicated here. Thus a case can be made that the naming effects also would have been as they were previously if ordinary non-lateralized presentations of the target words had been used.

The next step is obvious. Further experiments must demonstrate within the context of a within-subjects design that the procedure used here yields the Pronoun Bias Effect when the target is simultaneously available to both hemispheres, but not when it is initially available to only one.

If this can be done, there will be a *prima facie* case that the Pronoun Bias Effect is the product of a fast interaction among resources resident in both cerebral hemispheres. This in turn will provide opportunities to differentiate those resources and to associate each with a hemisphere.

Acknowledgements

The experimental work reported in this paper was done as part of a B.A. Honors Thesis project by the first author under the direction of the second author.

The work was supported in part by a grant from the National Institutes of Health (1 R01 NS22606-01) to the second author. Additional support was provided to the first author as a grant from the Honors Program at Ohio State. This work benefitted substantially from a parallel project conducted by Brian McAdams, also reported in this volume, and from discussions with Gary Berntson and Ewen King-Smith. We are grateful to Phil Enny for invaluable technical assistance.

References

Beaumont, J. G. (Ed.). (1982). *Divided visual field studies of cerebral organization.* Academic Press.

Bever, T. G., C. Carrithers, W. Cowart, and D. Townsend. (1990). Right-handers without left-handed relatives are more sensitive to grammatical structure and less sensitive to referential information than right-handers with left-handed relatives. In A. Galaburda (Ed.), *From Reading to Neurons.* Cambridge: MIT Press.

Caplan, D. (1987). *Neurolinguistics and Linguistic Aphasiology: An Introduction.* Cambridge, MA: Cambridge University Press.

Chiarello, C. and S. Nuding. (1987). Visual field effects for processing content and function words. *Neuropsychologia* **25(3)**: 539-548.

Cowart, W. (1988). Familial sinistrality and syntactic processing. In J.M. Williams and C.J. Long (Eds.), *Cognitive Neuropsychology*, New York: Plenum.

Cowart, W. and H. S. Cairns. (1987). Evidence for an anaphoric mechanism within syntactic processing: Some reference relations defy semantic and pragmatic constraints. *Memory and Cognition* **15(4)**: 318-331.

Drews, E. (1987). Qualitatively different organizational structures of lexical knowledge in the left and right hemispheres. *Neuropsychologia* **25(2)**: 419-427.

Fodor, J. A. (1983). *The Modularity of Mind.* Cambridge, MA: MIT Press.

Francis, W. N. and H. Kucera. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston, MA: Houghton Mifflin Company.

Howell, J. R. and M. P. Bryden. (1987). The effects of word imageability on visual half-field presentations with a lexical decision task. *Neuropsychologia*, **25(3)**: 527-538.

Kelly, J. P. (1985). Anatomy of the central visual pathways. In Kandel, E.R. and Schwartz, J.H. (Eds.) *Principles of Neural Science*, New York: Elsevier, pp. 356-365.

McAdams, B. (1990). Asymmetries in naming accuracy and in event-related potentials for laterally presented words of variable morphological complexity. This volume.

Stillings, N. A. (Ed.). (1987). *Cognitive Science: An Introduction*. Cambridge, MA: MIT Press.

Ohio State University Working Papers in Linguistics **38** (38-53)

'Kriptenstein's' Skeptical Paradox and Chomsky's Reply

Barbara Scholz

University of Toledo

Abstract: Chomsky's Knowledge of Language addresses certain conceptual questions about the foundations of generative linguistics that center on a 'skeptical paradox' that Kripke attributes to Wittgenstein. Chomsky's discussion offers an extended defense of his psychological conception of grammar against this challenge. This essay argues that Chomsky's response to the skeptical paradox is inadequate, but instructively so. The inadequacies of Chomsky's reply surface as a destructive dilemma for the psycholinguist conceptually committed to the generative paradigm in such a way as to reveal a conceptual incoherence in that paradigm. Specifically, the essay exhibits the dilemma as it arises for the performance theory of Berwick and Weinberg (1986). While modification of the philosophical foundations of generative linguistics may show the working psycholinguist the way out of the dilemma, this essay leaves the dispute unresolved, making only the negative point.

1. Introduction

In Knowledge of Language, Noam Chomsky (1986) focuses attention on three questions that are fundamental to generative linguistics conceived as a branch of psychology. The questions are:

I. What constitutes knowledge of language?
II. How is knowledge of language acquired?
III. How is knowledge of language put to use? (1986:3)

Questions (I) and (III) are striking insofar as providing an adequate response to each is as much a philosophical as a linguistic project and touch on traditional philosophical issues concerning the nature of mind, language, and thought. In the context of articulating his response to (III), Chomsky discusses the so called 'skeptical paradox' that Saul Kripke attributes to Ludwig Wittgenstein in On Rules and Private Language (1984). Chomsky takes this paradox to pose a deep challenge to the philosophical foundations of psychological linguistics.

The objectives of this essay are twofold. First, it aims to make clear the foundational philosophical challenge the skeptical paradox poses for linguistics. Secondly, it attempts to evaluate the force and adequacy of Chomsky's replies to this challenge. I shall argue that Chomsky's response does not address the central charge that there are conceptual inadequacies in linguistics conceived as a branch of psychology. While I think that the prospects for finding a solution to the paradox are reasonably good, it is not within the scope of this essay to articulate what I think is the most promising solution. I leave the dispute unresolved, having made only the negative point.

2. The Skeptical Paradox

The skeptical paradox that Kripke attributes to Wittgenstein (hereafter Kriptenstein[1]) is a family of arguments in the form of a *reductio ad absurdum*. Throughout his exposition of the Wittgensteinian texts, Kripke focuses on the idea that word meaning and denotation is rule-governed. The paradox, however, is perfectly general in the sense that a member of the family of arguments applies to any behavior that is alleged to be rule-governed. The version of the paradox that will interest us is displayed below.

The Paradox

(1) Jones knows a language L by being in competence state S.

(2) If Jones knows L by being in competence state S, there must be some fact about Jones that constitutes his being in S that justifies claims that he is in that state.

(3) There is no neutrally specifiable[2] fact about Jones that constitutes his being in S and that justifies claims that he is in that state.

(4) Hence, it is not the case that Jones knows a language L by being in state S.

(5) (1) and (4) are incompatible.

---

[1] In ON RULES AND PRIVATE LANGUAGE, Saul Kripke presents a forceful interpretation of arguments found in Ludwig Wittgenstein's PHILOSOPHICAL INVESTIGATIONS and REMARKS ON THE FOUNDATIONS OF MATHEMATICS. In KNOWLEDGE OF LANGUAGE, Chomsky's primary concern is with the argument that Kripke attributes to Wittgenstein and eschews discussion of the exegetical question whether Kripke's Wittgenstein adequately represents Wittgenstein. In this paper, we shall follow Chomsky and concern ourselves only with the force of Kripke's reading of Wittgenstein for linguistics in the generative paradigm.

[2] The specification of the fact that constitutes the grasp of a rule must provide non-trivial necessary and sufficient conditions that do not assume that what constitutes using a rule has already been explicated.

All versions of the paradox assume (1) and (2). (3) is established by argument. The first premise attributes a psychologically real competence state to a speaker. (2) unpacks necessary conditions for that attribution. It is assumed that the facts that constitute S also justify attribution of S to a speaker.

Given the ruling philosophical realist idea that psychological state attributions are fact stating, the skeptical paradox challenges the philosopher *cum* linguist to specify or describe the kind of thing, at a high level of generality, that constitutes such a state. The demand seems fair, if the philosopher/linguist holds that there must be facts that constitute S, it is fair to ask him to say what kind of thing those facts are.

In its paradigmatic and most analytical form, the task of philosophical semantic theories has been understood as one of providing an analysis or informative explanation of what constitutes meaning and reference. The idea is that an adequate philosophical semantic theory will answer the question: in virtue of what does a token of 'A' mean A? The question is an ontological question about speaker meaning and understanding of meaning. An adequate ontological analysis or explanation will provide a non-trivial specification or description of the facts, states of affairs or states of mind that constitute internal representation and use of rules. It is assumed that an answer to the ontological question will also specify the epistemological ground of rule attributions. Of course, such a description will not describe specific experimental effects, but it will answer the general question of what state S is in a way that allows experimental data to be interpreted to warrant the attribution of S. That there is a neutral description of these rules and their use (i.e., a description that does not somehow assume the notion of rule-following) is the idea that is the target of the *reductio*.

Kripke makes the assumption (for *reductio*) explicit as follows:

> By means of my external symbolic representation and my internal mental representation, I 'grasp' the rule for addition. One point is crucial to my grasp of this rule. Although I myself have computed only finitely many sums in the past, the rule determines my answer for indefinitely many sums that I have never previously considered. This is the whole point of the notion that in learning to add I grasp a rule: my past intentions regarding addition determine a unique answer for indefinitely many new cases in the future. (1984:8)

The point of claiming that meaning and knowledge of language are constituted by the grasp of rules is to capture the idea that linguistic behavior is normative. This notion is specified by the following adequacy conditions.

Adequacy Conditions on descriptions of competence states:

Condition A: Internally represented rules determine (in some sense) future and as yet unconsidered linguistic behavior.

Condition B: These rules are uniquely[3] represented.

When a philosopher claims that linguistic behavior is normative, he typically means (at least) that it is behavior of which it makes sense to claim that it is correct or incorrect. Of a normative behavior it is intelligible to say that it was mistaken or in error. Inextricably bound with the idea of a normative phenomenon is the idea of a unique standard in virtue of which that phenomenon is judged permissible or not. Of course this idea of linguistic normativity contrasts with the linguist's standard conception of the prescriptivity of grammars. Introductory linguistics textbooks take pains to deny that linguists' grammars are prescriptive and deny that grammars (not behaviors) are normative. The import is to distinguish between the grammars of scientific linguistics and old fashioned grammarians' grammars. There is no conflict between what the philosopher claims is normative and what the linguist claims is not normative, since behaviors are not grammars.

Rule-governed behavior contrasts sharply with behavior that is merely rule-conforming. The description of rule-conforming behavior need satisfy neither Adequacy Condition A nor B. For example, the behavior of bodies conforms with Newton's Laws, which are rules of a sort, but bodies do not make mistakes if their behavior does not conform with those laws. Moreover, by continuing in its orbit, Neptune merely conforms with Newton's laws and does not apply an internal representation of them. Satisfaction of Adequacy Conditions A and B captures the notion that linguistic behavior is governed by internally represented rules. What is wanted in response to the paradox is an informative explanation of what it is to internally represent and apply a rule that does not assume that this notion has been satisfactorily explicated. A neutral description shows (3) is false.

Computational theories of mind can profitably be seen as attempts to provide neutral descriptions of what it is to follow a rule. For example, in the second chapter of The Language of Thought (1975) Jerry Fodor motivates his appeal to the computer metaphor with a discussion of Wittgenstein's skeptical paradox. The idea, of course, is that a computational theory of mind can neutrally specify facts in virtue of which a speaker internally represents and uses rules. Having a "language

---

[3] 'Unique' should not be read here to exclude the possibility of genuinely ambiguous syntactical tokens. Chomsky often states this adequacy condition by claiming that the rules must be correct. For example, he writes that the idea of an internal grammar or I-language is correct, while in the case of an E-language there is no issue of correctness or incorrectness (1986:26). Alternatively, he claims that a generative grammar "purports to depict exactly what one knows when one knows a language" (1986:24).

of thought" i.e., a language that the machine is built to use, and something like a compiler that translates from natural languages into a brain code, constitutes a neutral description of rules and their use.[4] While Fodor's suggested solution is well worth discussion, we shall not pursue it here. Our concern is specifically with Chomsky's views in KNOWLEDGE OF LANGUAGE where he is not concerned with computational theories of what it is to mentally represent and use a rule (1986:239)

So far, we have said nothing about the argument for (3), the claim that there is no fact specifiable in neutral terms that constitutes the grasp of a rule so that Adequacy Conditions A and B are satisfied. The argument for (3) is an argument by elimination. Kriptenstein considers and rejects a variety of candidates which, if adequate, would show that (3) is false. In general, the candidates fall into two categories based on the way they fail as solutions to the paradox. The first group of candidates (mental images, experiential states, dispositions) fail because they do not satisfy Adequacy Conditions A and B. In consequence, such candidate solutions fail to capture the relevant properties of linguistic behavior. Kripke's arguments against the idea that linguistic behavior can be explicated in terms of a speaker's dispositions are reminiscent of Chomsky's attack on Skinner. The second group of

---

[4] Too briefly, the debate with respect to whether computational or language of thought solutions to the paradox are adequate focuses on whether a causal description of mental content is adequate. For given a description of the content of mental representations and their use in purely causal (neutral terms), what Fodor calls the disjunction problem arises. In PSYCHOSEMANTICS (1987:102) Fodor describes the disjunction problem as follows.

> We can put it that a viable causal theory of content has to acknowledge two kinds of cases where there are disjoint causally sufficient conditions for the tokenings of a symbol: the case where the content of the symbol is disjunctive ('A' expresses the property of being (A v B)) and the case where the content of the symbol is not disjunctive and some tokenings are false ('A' expresses the property of being A, and B-caused 'A' tokenings misrepresent). ...
>
> The disjunction problem is extremely robust; so far as I know, it arises in one guise or another for every causal theory of content that has been thus far proposed.

The disjunction problem is the problem of distinguishing the case where 'A' correctly represents (A v B) and the case where 'A' misrepresents B, because it was caused by B. The problem arises because descriptions of mental representations (couched in purely causal terms) do not satisfy the adequacy conditions (much discussed in text above) on normative or rule-governed phenomena. Mental content like linguistic knowledge is both productive and unique in the relevant respect. The disjunction problem shows that a purely causal theory of content cannot capture the notion of a mistaken representation.

candidates, which includes linguists' descriptions of competence states, i.e., grammars, fails to show that (3) is false, because they do not neutrally describe internally represented rules and their use.

If my exposition of the paradox has been clear, then it should be obvious why the typical description of a competence state does not adequately neutrally describe the fact that constitutes an internally represented rule (or system of rules). What is wanted is an account of what it is for an individual to represent and apply rules in terms that make no appeal to the notion of a mentally represented rule. The explanandum must not appear in the explanans on pain of circularity. Our conception of a typical description of such competence states assumes the very idea that is to be explained -- that a speaker represents and uses rules. Hence, a competence state as typically conceived and described begs the very question at issue.

To be sure, Chomskian competence states satisfy Adequacy Conditions A and B. Such states are claimed to be explanatorily adequate, and not merely descriptively adequate and so satisfy Adequacy Condition B. Moreover, it is typically claimed that such states and their description account for the phenomena of linguistic productivity and so by their very nature satisfy Adequacy Condition A. But it is not a solution to the skeptical paradox insofar as it does not describe a fact in neutral terms that shows (3) of the paradox to be false.

Kriptenstein puts the point this way:

> . . .our understanding of 'competence' is dependent on our understanding of 'following a rule'. . . Only after the skeptical problem about rules has been resolved can we then define 'competence' in terms of rule following. . . Although the remarks in the text warn against the use of the 'competence' notion as a solution to our problem, in no way are they arguments against the notion itself. Nevertheless, given the skeptical nature of Wittgenstein's solution to his problem . . . , it is clear that if Wittgenstein's standpoint is accepted, the notion of 'competence' will be seen in a radically different way from the way it is implicitly seen in much of the linguistics literature. For if statements attributing rule-following are neither regarded as stating facts, nor to be thought of as explaining behavior, it would seem that the use of the ideas of rules and of competence in linguistics needs serious reconsideration. (1984:31)

If Kripke is right, then if the linguist is going to resist the contradiction at (5) -- that a speaker both knows L by being in S and does not know L by being in S, then he must do so by denying (2). It is with the rejection of (2) that the paradox becomes interesting.

(2) claims that a speaker is in S only if there is a fact (specifiable in neutral terms) that constitutes that state, satisfies Adequacy Conditions A and B, and justifies attribution of S to the speaker. The second premise will be false just in case speakers represent and use rules (in some sense), but one of the following four responses to the paradox is viable.

Four Possible Responses to the Paradox:

Response A: the facts that constitute S do not satisfy Adequacy Conditions A and B, or

Response B: the facts that constitute S do not justify or warrant competence state attributions, or

Response C: the facts that constitute S cannot be specified or described in other terms, or

Response D: there are no facts at all that constitute S.

Since competence states (for Chomsky) are described by appeal to rules, the linguist must claim that (2) not (3) is false if he is to avoid the antinomy at (5). In itself, rejecting (2) is not especially problematic *prima facie*. For there are four ways the linguist can avoid the conclusion that a speaker both is and is not in S. Some of these alternatives are more attractive than others, but the antinomy is well worth avoiding.

It will be useful to attend to exactly what the linguist accepts should he embrace one of the Responses A - D. To begin, consider Response A. Accepting that the facts that constitute competence states do not satisfy Adequacy Conditions A and B entails giving up either the idea that the rules that constitute competence states are unique or that linguistic knowledge is productive. Neither alternative is acceptable. Suppose, for example, that two rules of S, R and R', when applied to a string P assign a different status to P (e.g. noise and acceptable), so the rules of S are not unique. If both R and R' are claimed to be the standard that constitutes S in virtue of which P has the status it has, then S is no standard. Non-unique standards are not standards at all. Thus, the Adequacy Conditions seem essential for psychological linguistics.

Alternatively, the linguist might consider rejecting Adequacy Condition A as a means of accepting Response A. In this case he would reject the idea that knowledge of language is productive in the sense that what people know when they know a language applies to expressions they have never heard or previously considered. This option is not open to one who claims that language is rule-governed. Accepting Response A by rejecting Adequacy Conditions A or B would require a revision in the typical conception of competence.

Alternatively, consider Response B. The Chomskian idea is that facts about the speaker's competence states (e.g. psycholinguistic evidence) justify the attribution of such states to speakers. If the linguist accepts Response B, he denies that claim. If one wants to take psycholinguistic evidence to be evidence about competence states and the use of rules, then one cannot accept Response B.

Next, consider Response C. In accepting Response C the linguist accepts that facts that constitute S cannot be described without appeal to internally represented rules. The advocate of such a view might claim that the facts in virtue of which Jones knows L by being in S are brute linguistic facts that cannot be described in other terms. Such a view suggests that there is a wide and unbridgeable gulf between linguistic facts and physical and chemical facts, for example, which are specifiable without appeal to internally represented rules. *Prima facie* it also conflicts with naturalism, the view that linguistic facts are part of the natural biological order. Indeed, the view seems to entail a form of linguistic dualism on which the facts of linguistics are a *sui generis* kind of entity -- an internal-rule-fact.

Finally, accepting Response D entails that whatever represented rules are they are not psychologically real phenomena. In consequence of Response D, competence state attributions are not literally true or false, although they may be useful for some purpose. The linguist who accepts Response D adopts a form of instrumentalism about competence states.

At this point, the Chomskian linguist is faced with four possible responses to the skeptical paradox. From what we have said so far, Response C is the most attractive alternative --but much more on this below.

3. Chomsky's Response to Kriptenstein

In Knowledge of Language, Chomsky argues that the skeptical paradox does not show that "the notion of competence [must] be seen in a light radically different from the way that it is seen in much of the linguistics literature".[5] Chomsky's defense consists in accepting Response A and Response C but explicitly denying Response B and Response D.

[5] In Chapter 4 of Knowledge of Language Chomsky focuses a great deal of attention to arguing against the skeptical solution that Kripke attributes to Wittgenstein. Wittgenstein's solution to the paradox, like Chomsky's, is skeptical insofar as it accepts that (3) is true but (2) is false. Kripke claims that the thesis that there is such a thing as a private language follows as a corollary from Wittgenstein's own skeptical solution. Whether Wittgenstein's skeptical solution is adequate and whether the impossibility of a private language (whatever that is) does so follow is independent of the challenge that the paradox raises for the generative linguist.

We have already noted that as a matter of logic, the generative linguist must accept at least one of the above disjuncts. In this section, I shall argue that it is not open to the Chomskian linguist to accept Response A and that if Response C is accepted, the problem of saying what it is to use one rule rather than another resurfaces in the project of formulating performance models of competence theories.

When Kriptenstein claims that linguistic behavior is normative, what is claimed is that the Adequacy Conditions apply. In KNOWLEDGE OF LANGUAGE, Chomsky explicitly argues that linguistic phenomena are not normative (as *per* Response A) and urges that all issues of correct and incorrect performance can be 'dropped' by considering 'normal'[6] cases of attributing rules to native speakers. Chomsky illustrates the claim by observing that we do say that a child's internal rules are incorrect, but we are unlikely to say of adult (normal) native speakers that their rules are incorrect. So of children who overgeneralize and say "sleeped" Chomsky writes:

> we will say that their rules are "incorrect" meaning different from those of the adult community or a selected portion of it. Here we invoke the normative teleological aspect of the common sense notion of language. (1986:227)

By contrast, we do not say of the adult Irishman who says "There himself goes down the road," that his internal rules are incorrect. According to Chomsky, the generative linguist can embrace Response A, because the linguists' theory merely describes a speaker's internally represented rules.

Accepting Response A as a way to avoid the antinomy is fundamentally misguided. First, what is at issue (with respect to the paradox) is the normative status of linguistic behavior, not the normative status of the description of the competence state. For example, an anthropologist may claim to describe a system of moral rules in a particular community. While the anthropologist's description is not normative, what he describes, if accurate, will characterize morally permissible and impermissible (morally normative) behavior in that community. Similarly, the linguists' competence theory describes a competence state, nevertheless the hypothesized internally represented rules characterize the sentences of L which are linguistically (not morally) permissible or impermissible. Make no mistake, our concern is not with what one calls this interesting property of linguistic behavior. What it is important to see is that one of the primary reasons to posit internal rules that are used by speakers is to explicate what the philosopher (erroneously if you like) calls linguistic normativity. If one accepts Response A and so denies that the

[6] One wonders what Chomsky means by 'normal' here. One suspects that he means something like 'in cases where error in performance is not at issue', but then it is trivially and uninterestingly true that issues of distinguishing correct and incorrect rule attributions do not arise in those cases.

facts that constitute internal rules and their application must satisfactorily explain linguistic productivity or provide a standard of performance, then one rejects the idea that language is a rule-governed phenomenon. Accepting Response A entails denying distinctive characteristics of the phenomenon that one wants to explain.

Secondly, Chomsky's point that linguistic competence is not a normative notion (because we would say of the child that his rules are incorrect but would not say of the Irishman that his rules are incorrect) is moot. What is relevant is that if the child overgeneralizes the use of a rule, for example, that **Verb+PAST --> Verb+/d/**, then it is correct for the child to say "sleeped". What is correct or incorrect is behavior relative to internally represented rules not descriptions of those rules. In consequence, Chomsky's attempts to avoid the paradox by claiming the generative linguist can coherently accept Response A is not persuasive. Indeed, if successful, his arguments would undermine his own performance/competence distinctions[7] in the sense that a competence state provides a standard of performance and explicates linguistic productivity.

Chomsky explicitly advocates two routes out of the paradox, for he also endorses the idea that the linguist can avoid the antinomy at (5) by claiming that (2) is false in virtue of there being *sui generis* linguistic facts. I shall argue that if the generative linguist embraces this claim (Response C) the troubles of the skeptical paradox resurface, but now as a problem for the psycholinguist. In short, my thesis is that the challenge that the skeptical paradox presents for the linguist is a bump-under-the-rug phenomenon. If the linguist attempts to detoxify the paradox by claiming (2) is false in virtue of accepting Response C, then the paradox revisits itself on the working psycholinguist. But first, we need a clearer idea of what these *sui generis* rule-facts are supposed to be like.

Chomsky has recently taken to using the neologism 'I-language' to refer to what he previously called competence states. For Chomsky, an I-language is "some

---

[7] Space does not permit consideration of the many versions of the performance/competence distinction as made by Chomsky in the course of his long and illustrious career. Suffice it to say that at one time Chomsky seemed to think that a competence theory was to be distinguished from a performance theory only insofar as the competence theory required an idealization away from various interfering performance factors, e.g., memory limitations, background noise, etc. It is not clear that this notion of the competence/performance theory is normative in the sense I have been concerned with here. Linguistic normativity is at best a troublesome notion. Unfortunately, it is not at all clear that either the philosopher of language or the linguist can live without it. The intuitive distinction between rule-conforming and rule-guided behavior seems cogent. If the linguists' theory does not capture the relevant features of the phenomena, then it seems that his theory simply does not explain something that needs to be explained. A full scale study of linguistic normativity in conjunction with an examination of various versions of the performance/competence distinction would be useful.

element of the mind of a person who knows a language, acquired by the learner, and used by the speaker."[1986:22] In particular, it is a second-order property of speaker's (non-neurophysiological?) mind. It is a "distinct level of things in the world", not to be explicated in causal terms. That an I-language is a second-order property of native speakers is not sufficient to secure the claim that facts about such competence states are uniquely linguistic. A second-order property is simply a property of the first-order properties of objects. For example, dispositional properties like solubility are second-order properties of objects. A salt crystal has the property of being soluble, and being soluble is a property of the first-order properties of the salt crystal. More specifically, being soluble is a property of the relative electro-static charge on sodium and chloride ions (a first-order property) in the presence of water molecules. But clearly, there is nothing uniquely linguistic about the second-order property of solubility.

Claims about the psychological reality of grammars, for Chomsky, are formulated as claims about which grammar a speaker uses. In a characteristic passage he writes:

> Statements about the I-language are true or false, much the same way statements about the chemical structure of benzene . . . are true or false. The I-language L may be used by a speaker but not the I-language L', even if the two generate the same class of expressions. (1986:37)

The point that a grammar G is psychologically real in the sense that it is used by a speaker while the weakly equivalent G' is not is what makes the speaker of a language a rule-follower and rule-user and is what makes one description of a competence state psychologically real and the other not.

We shall explore how accepting Response C raises difficulties for the psycholinguist by considering the case of the Derivational Theory of Complexity [DTC] from the history of psycholinguistics. DTC was an early, if not the first, attempt to provide a performance model of the grammar outlined in ASPECTS OF THE THEORY OF SYNTAX with sufficient specificity and detail that Chomskian conjectures about the psychological reality of the so-called 'Standard Theory' could be tested (though some actual psycholinguistic uses of DTC preceded the publication of ASPECTS). DTC assumed the grammar (description of the I-language) of ASPECTS, and took on the assumption that the relation between that grammar and the parsing algorithm was transparent in the sense that the relation was isomorphic. The deep structure and surface structure of input strings was recovered by the parser which echoed the grammar, and the deep structure was derived from the surface structure by the application of 'inverse transformations'. DTC also assumed that each grammatical operation cost one unit of time and since the parser/grammar relation was one-to-one the temporal cost of constructing the deep structures from surface structures was the sum of the number of the applications of rules necessary for the derivation of the sentence. There are more subtle versions of DTC, but all versions share the notion of a relatively transparent relation between linguistic rules and psychological operations. The motto was one temporal unit for each rule application.

Given the assumptions of DTC, the theory predicts that the mapping from deep to surface structure for active sentences requires one fewer rule applications than passives, so it would take less time to parse actives than passives. Some very early evidence appeared to confirm DTC (and the grammar of Aspects). However, later experiments "found no correlation between sentence processing time and the length of transformational derivation".[8] There are three possible retrenched versions of DTC: reject the grammar, reject the assumption that the relation between the grammar and parser is isomorphic, or reject the computational complexity measure. Each alternative has subsequently been attempted, but it is the second avenue of retrenchment that shall be of particular interest to us here.

Fodor, Bever and Garrett (1974) were the first to suggest that the transparent, isomorphic relation between the grammar and parser be revised. In place of the isomorphic relation they substituted "heuristic strategies", with the effect that the subsequent performance model reduced the online computation involved in sentence comprehension. Now at one level, the rejection by Fodor et. al. of the transparency assumption is a standard piece of ordinary science. In the face of disconfirming evidence for the favored theory, hang onto the theory and reject an auxiliary hypothesis. However, the retrenchment in this case is self-defeating. If one permits the adoption of any "heuristic strategy" as the posited relation between the grammar and parser, then virtually any parser will model any set of rules. In this case, the parser/grammar relationship is completely unconstrained by the theory, and any sense in which the performance model can be used to test the psychological reality of the theory disappears.

Cognizant of the dangers heuristic grammar/parser relations pose for claims about the psychological reality of grammars, Berwick and Weinberg (1986) define a notion of grammatical covering in an attempt to respond to the problem. Informally characterized, they claim that one grammar G covers another G' if

> (1) both generate the same language L(G) = L(G'), that is the grammars are weakly equivalent; and (2) we can find parses of structural descriptions that G' assigns to sentences using G and then applying a "simple" or easily computed mapping to the resulting output. (1986:79-80)

---

[8] Quoted in Berwick and Weinberg (1986:42) from Joan Bresnan, "A Realistic Transformational Grammar" in M. Halle, J. Bresnan and G. Miller (Eds.) LINGUISTIC THEORY AND PSYCHOLOGICAL REALITY (Cambridge: MIT Press, 1978).

They cash out "simple" as follows:

> . . .[t]he usual definition of "simple" drawn from the formal literature is that of a string homomorphism. That is, if the parse of a sentence with respect to a grammar G is a string of numbers corresponding to the rules that were applied to generate the sentence under some arbitrary numbering of the rules of the grammar and some canonical mapping derivation sequence, the translation mapping that carries this string of numbers to a new string of numbers corresponding to the parse [under G'] must be a homomorphism under the concatenation. (1986:79-80)

If G covers G', the grammars (or grammar and parser) are not merely weakly equivalent. The structure of the parse, but not the number of rule applications, is preserved under string homomorphism. The covering relationship, however, is weaker than strong equivalence and the transparency relation. In consequence, if computational cost measures are held constant, predictions of total time cost will vary radically from parser to parser.

Recall that Chomsky defines psychological reality in terms of the use of a grammar -- applying internally represented rules. However, if we follow Berwick and Weinberg's elegant suggestion, we accept a notion of grammatical covering on which the description of the I-language, e.g., G, is not strongly equivalent to the parser, e.g., G', -- the used grammar. Once we do this, it is not easy to see what the idea of using the covering grammar comes to. By hypothesis, it is not the covering grammar, but G', the parser, that is used. Worse yet, there are indefinitely many parsers that are covered by the competence grammar. Are we to suppose that the speaker can use only one of these, or many? If we suppose speakers may implement more than one of the covered parsers, then temporal costs will vary for the same grammar (holding temporal cost measures constant) and strikingly different predictions follow from the same competence theory. [What empirical content would attach to the claim that the speaker uses one, rather than many of the covered parsers?] In this case the grammar no longer provides a standard of performance in terms of temporal cost and loses its ability to function as one needs a competence theory to function in the performance model. On the other hand, if we take the speaker to use just one of the parsers covered by the grammar, then there is no advantage in claiming that the competence grammar is used, although the parse is structurally homomorphic to the parse of the competence grammar had it been used. A further difficulty is that, assuming that the covering relation is symmetric, any given G' (the parser) might cover indefinitely many competence grammars, G.

Berwick and Weinberg claim that loosening the relation between (competence) grammar and parser (performance grammar) provides the computational linguist with a performance model that has clear methodological advantages. In response to the question "Why build a parser that is covered by a competence grammar?" they write:

> The answer is that by keeping the levels of grammar and algorithmic realization distinct, it is easier to determine just what is contributing to discrepancies between theory and surface facts. For instance, if levels are kept distinct, then one is able to hold the grammar constant and vary the machine architectures to explore the possibility of good fit between psycholinguistic evidence and model. Suppose these results came to naught. We can then try to vary machine architecture and covering mappings, still seeking model and data compatibility. . . . [M]odularity of explanation permits modularity of scientific investigation. (1986:80)

As a methodological claim, the thesis is unassailable. If I understand Berwick and Weinberg correctly, their idea is that a modular conception of performance models facilitates the manipulation of elements of computational simulations of those models and promotes ease of identifying various models of the competence grammar. However, such methodological boons would issue from taking the covering grammar to identify a class of parsers. Berwick and Weinberg's response does not bear on the issue of what one could now mean by the psychological reality of the covering grammar. That distinguishing between a competence grammar (class of parsers) and a used parser is useful does not explicate what it could now mean to use the competence grammar. The use of a grammar, by hypothesis involves specific and determinate temporal costs, not a class of temporal costs for the same (grammatical) phenomenon.

Let us rehearse where we have been. The initial idea was that internally representing rules was a *sui generis* fact about Jones (Response C) that could be embraced to support the falsity of (2) in the skeptical paradox. In order to develop a performance model of a competence theory which is taken to describe those brute linguistic rule-facts, one must posit a grammar parser relation. If the relation is supposed to be transparent and consistent with DTC, we have a clear (but non-neutral) description of what it is to use those rules -- the transparency relation. But the transparency relation has empirical difficulties, so a string homomorphism was postulated to gain its methodological advantages. But now conceptual problems arise and philosophical difficulties resurface concerning what it means to say that a speaker uses the competence grammar (not the parsing grammar) and that the competence grammar is psychologically real. If we say that a class of parsers are the facts in virtue of which the speaker is in S, then there is incompatible psycholinguistic evidence about those facts that will *prima facie* both support that a speaker is and is not in S. For each parser will provide a different standard with respect to time costs for the same phenomenon. Alternatively, if we say to use G is to use just one G' covered by G, then it seems that we would be just as well off to say that G' is the real fact of the matter, and G a useful means of identifying G', but is neither used by the speaker nor psychologically real. This latter suggestion has been made by a number of theorists and rejected by Chomsky (see, for example, Soames 1984).

I have argued that if the linguist attempts to avoid the skeptical paradox that Kripke attributes to Wittgenstein along the lines Chomsky suggests in KNOWLEDGE OF LANGUAGE, then either he ends up rejecting the performance/competence distinction or (given the methodological requirements on a performance model) the paradox resurfaces for the psycholinguist in terms of saying what it is to use the competence grammar. If what I have argued is correct, then the alternative strategies open to the linguist for avoiding the paradox are to either claim Response D, that there is nothing about a speaker in virtue of which he uses one grammar rather than an other (i.e., linguists' competence state attributions are not fact stating), or Response B, that facts about competence states do not justify grammar attributions. Neither of these alternatives is acceptable, for both require considerable revision in the typical conception of competence. At least this is the situation unless the linguist can show (or plausibly hope) that there is a description of the facts that constitute what it is to use a rule in neutral terms. I have not, of course, addressed the issue of whether a neutral description of rule-following can be articulated by computational theories of mind. That is a distinct and very long story.

While I do not think the skeptical paradox poses insuperable conceptual difficulties for generative linguistics (even if one does not embrace the language of thought hypothesis and the computational metaphor), I have attempted to show that there is a troublesome conceptual problem here. Surely one of the conceptual burdens of psycholinguistics is to say what constitutes the use of rules. The Wittgensteinian skeptical paradox makes explicit the conceptual difficulties inherent in that task.

Acknowledgements

This paper has been vastly improved by the comments and sage editorial advice of Wayne Cowart. I am deeply indebted to him. I would also like to thank David McCarty for helpful comments on a much earlier draft.

References

Berwick, R. and A. Weinberg. (1986). *The Grammatical Basis of Linguistic Performance*. Cambridge, MA: MIT Press.

Chomsky, N. (1986). *Knowledge of Language*. New York: Praeger.

Fodor, J. (1975). *The Language of Thought*. New York: Thomas Y. Crowell.

________. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Kripke, S. (1984). *Wittgenstein On Rules and Private Language*. Cambridge, MA: Harvard University Press.

Soames, S. (1984). "Linguistics and Psychology". *Linguistics and Philosophy* 7.2. MIT Press.

Ohio State University Working Papers in Linguistics 38 (54-66)

Illicit Acceptability in *picture* NPs[1]

Wayne Cowart
University of Southern Maine

Abstract: Four experiments examine the interaction between extraction and specificity in **picture** NPs. The results indicate that the acceptability judgements of naive speakers show highly robust patterns that do not conform well to widely held assumptions about the relative acceptability of several theoretically important kinds of sentence. There is also evidence that the difference between argument and non-argument extractions has a marked impact on acceptability (though no such acceptability difference has figured in linguistic theory). Further, the paper argues that there are circumstances in which ungrammatical sentences may be rendered acceptable via the intrusion of extragrammatical mechanisms in comprehension. Thus, the acceptability of these sentences is 'illicit'.

1. Introduction

The empirical generalization addressed by Chomsky's subjacency principle seems to apply to sentences such as (1) (see Chomsky 1973, 1981, 1986; also Riemsdijk and Williams 1986, Lasnik and Uriagereka 1988).

(1) Who did the Duchess sell a portrait of?

Nevertheless sentences such as this seem fully acceptable and intelligible to many speakers. In consequence, such cases have been regarded as fully grammatical in the linguistic literature.

Broadly speaking, there are two approaches by which a grammatical theory that incorporates some form of the subjacency constraint might accommodate these facts. First, the components of the grammatical theory that capture the subjacency constraint might be formulated in such a way that they do not apply to cases such as (1). Second, the grammar might incorporate principles that have the effect of shielding cases like (1) from the subjacency mechanism(s).

This study takes a different tack, looking more closely at the intuitions about that underlie the assumption that it is grammatical. The study begins from the

---

[1] This paper originally appeared in Wiltshire, Gracyk and Music, eds. (1989). *Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society, Part One: The General Session.* Chicago: Chicago Linguistic Society. Reprinted with permission of CLS.

tentative suggestion that the acceptability of (1) might be spurious, that (1) is in fact ungrammatical, despite its apparent high acceptability. This requires a more systematic approach to assessing intuitions about (1) and close consideration of various alternative cognitive mechanisms that might account for the acceptability of the cases in the face of possible ungrammaticality.

The work reported below assumes that, all other things being equal, it is reasonable to hold the grammar accountable for patterns of judged acceptability. Where observed patterns do not coincide with those predicted by some grammar, it is appropriate to seek an account of the discrepancy in the larger ensemble of mental resources that speaker/hearers bring to language comprehension, in the cognitive processes that realize specifically grammatical knowledge, in the grammar itself, or in some combination of these. These issues are discussed further at the end of the paper.

2. Experiment 1: Possible Subjacency Effects

The first experiment was designed to determine whether the patterns of judged acceptability obtained with sentences similar to (1) are in reasonable accord with the patterns of grammaticality commonly assumed in the formulation of grammatical theory.

2.1. Materials

The target cases are presented in Table I. The Control cases are taken to be uncontroversially acceptable and grammatical. The Specified Subject cases are equally uncontroversially unacceptable and ungrammatical. The status of the other two cases is unclear. The definite cases have sometimes been taken to be acceptable and sometimes unacceptable, with corresponding assumptions about grammaticality. The Indefinite cases are patterned on (1), discussed above. Prevailing assumptions about grammaticality in these cases predict that the Indefinite cases will pattern with the Controls, that the Specified Subject cases will be distinctly less acceptable, and that the Definite cases will pattern with either the first set or the second, depending upon whether they are in fact grammatical and acceptable.

Control:	Why did the Duchess sell Turner's portrait of her father?
Indefinite:	Who did the Duchess sell a portrait of?
Definite:	Who did the Duchess sell the portrait of?
Specified Subject:	Who did the Duchess sell Turner's portrait of?

Table I: Materials for Experiment 1.

There were 24 sets of materials modeled on those in Table I. There were also 96 filler sentences of diverse kinds. Four presentation lists of sentences were constructed so that only one member of each set appeared in each list and so that six items of each of the four types shown in Table I appeared in each list. Thus each subject saw equal numbers of items of each type distributed throughout a much larger list of fillers and no subject saw more than one member of any set.

2.2. Methods

For this and all subsequent experiments, the materials were presented to subjects as printed lists. Subjects were asked to read each sentences and to

> "...indicate whether the item seems like a fully normal, understandable sentence to you. If it does, please check the box on the far right. If, on the other hand, the sentence seems very odd, awkward or difficult to understand, please check the box on the far left. If your feelings about the sentence are somewhere between these extremes, check one of the middle boxes. THERE ARE NO 'RIGHT' OR 'WRONG' ANSWERS. Please base your responses solely on your personal judgments, not on rules you may have learned about what is 'proper' or 'correct' English." (emphasis in original)

For some experiments a separate machine-scored answer sheet was used. Subjects responded by way of a four point scale whose extremes were marked "Odd" and "OK". Subjects were encouraged to respond rapidly and typically finished the list of 120 sentences, plus additional background questions, in less than 15 minutes.

Subjects were undergraduate students at The Ohio State University.

2.3. Results and Discussion

The results of two separate runs of Experiment 1 with a total of 228 subjects are combined and summarized in Figure 1. The most important result is that the pairwise difference in acceptability between the Indefinite and Control cases is highly significant, $F(1,227)=459.27$, $p<.001$, and indeed is the single largest difference observed. The pairwise differences between the Indefinite and Definite cases and between the Definite and Specified Subject cases, though numerically smaller, are also highly robust, $p<.001$. The overall decline in acceptability across the four cases is significant, $F(3,681)=619.69$, $p<.001$, as is the decline across the Indefinite, Definite and Specified Subject cases, $p<.001$.

The pattern of results seen in Figure 1 does not fit well with typical assumptions about the relative acceptability of these kinds of cases. The Indefinite cases are worse than expected and the difference between them and the Definite and Specified Subject cases is smaller than expected. One possible view of the pattern in Figure 1 posits that the Indefinite cases are ungrammatical in virtue of being covered by the subjacency generalization and that the decline in acceptability across the Indefinite, Definite and Specified Subject cases reflects increasing specificity in the determiners of the **picture** NPs (Fiengo and Higgenbotham 1981, Fiengo 1987). Two alternative hypotheses must be considered. First, the reduced acceptability of the three most impaired cases might result simply from the presence of a preposition at the end of the sentence, regardless of other aspects of the structure of the sentence. Perhaps the common prescriptivist ban on sentences ending in prepositions exerted some influence (despite the instructions to subjects to ignore such considerations). Second, somewhat similarly, it might be that the apparent effect of specificity is not sensitive to the structure in which the determiner occurs but is somehow induced by mere surface variation in the form of the determiner. Neither hypothesis is especially interesting from a linguistic point of view, but they cannot be dismissed out of hand. The second experiment addresses these issues.

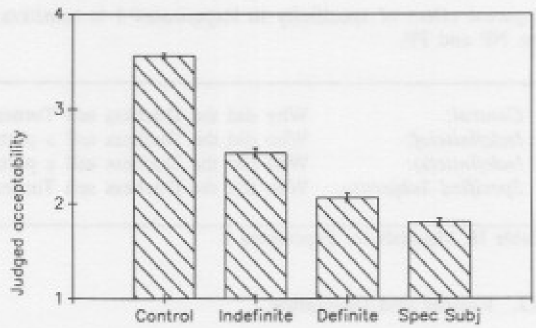


Figure 1: Mean judged acceptability for the four materials conditions of Experiment 1.

3. Experiment 2: *of* vs. *to*

The strategy of Experiment 2 is to compare cases similar to those used above to others where the prepositional phrase containing the extraction site has *to* as its head and is a sister of the NP to its left, rather than embedded within it. Such cases test the claim about sentence-final prepositions and provide an alternative control case against which to compare the Indefinite cases of Experiment 1.

3.1. Materials

A sample set of materials is displayed in Table II. The Control and Indefinite/of cases are drawn from Experiment 1. The Indefinite/to case should be identical to the Indefinite/of case if the depressed results in the former case results merely from the presence of a preposition at the end of the sentence. The contrast between the Indefinite/to and Specified Subject/to cases likewise tests whether the apparent effect of specificity in Experiment 1 is sensitive to the structural relation of the NP and PP.

Control:	Why did the Duchess sell Turner's portrait of her father?
Indefinite/of:	Who did the Duchess sell a portrait of?
Indefinite/to:	Who did the Duchess sell a portrait to?
Specified Subject/to:	Who did the Duchess sell Turner's portrait to?

Table II: Materials for Experiment 2.

3.2. Results and Discussion

The most critical result of Experiment 2 is that there is a statistically robust difference between the acceptability of the Indefinite/of and Indefinite/to cases, $F(1,40)=12.93$, $p<.001$. Furthermore, the decline in acceptability from the Control case to the Indefinite/of case is reliably larger than the decline from the Control to the Indefinite/to case, $F(1,40)=12.93$, $p<.001$. This suggests that the unacceptability of the Indefinite/of cases is not to be explained solely by the presence of some preposition at the end of the sentence. This picture is somewhat clouded, however, by the finding that the Indefinite/to cases are also significantly less acceptable than the Control, $F(1,40)=10.20$, $p<.005$. This fact is not explained by either the subjacency or specificity proposals discussed above.

Some support for the specificity proposal is evident in the fact that the Specified Subject/to cases are more, rather than less, acceptable than the Indefinite/to cases. This is opposite to the pattern found in Experiment 1, though the difference here is not robust, $F(1,40)=1.04$, NS. When the PP is outside of the NP, increasing specificity does not compromise acceptability.

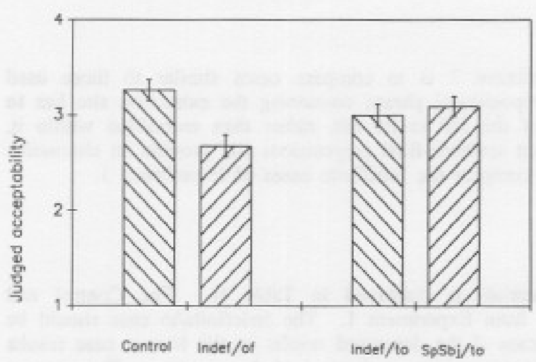


Figure 2: Mean judged acceptability for the four cases of Experiment 2.

The difference between the Control and Indefinite/of cases is robust, $F(1,40)=35.23$, $p<.001$, which replicates Experiment 1. The observed acceptability of the Control cases in Experiment 2 is somewhat lower than that obtained in Experiment 1. Such interexperiment differences and, in general, the absolute numerical values of the acceptability means will not be considered here. There were small differences in the filler sentences used in different experiments, and other minor differences of technique, that may account for any such differences.

In sum, these results are consistent with the claim that the Indefinite cases of Experiment 1 are affected by subjacency and that the acceptability of similar extraction cases is sensitive to the specificity of the NP within which the PP is embedded.

4. Experiment 3: Depth of Embedding

Another possible confound relevant to the results of Experiment 1 is depth of embedding. Though there is no very clear and generally accepted metric of depth of embedding, nevertheless, it seems clear that the Indefinite/of cases involve extraction from a site that is more deeply embedded in the hierarchical structure of the sentence than is the comparable site in the Control sentences or the Indefinite/to cases. Counting only S and NP boundaries, the of cases involve extraction from two levels down, while the others require extraction from no more than one level down, as illustrated in Table III.

Why did [the Duchess sell [Turner's portrait of her father] *t*]
Who did [the Duchess sell [a portrait of *t*]]
Who did [the Duchess sell [a portrait] to *t*]
Who did [the Duchess say [Max likes *t*]]

Table III: Depth of embedding, bracketing only relevant NPs and Ss.

As a partial control for the possible influence of depth of embedding, Experiment 3 compared the Control/Indefinite contrast of Experiment 1 with pairs contrasting two degrees of embedding where the more deeply embedded case is uncontroversially regarded as acceptable in the linguistic literature, as in the last example in Table III.

4.1. Materials

One complete set of materials is illustrated in Table IV. The Subjacency Cases are drawn from Experiment 1. The Shallow Depth Cases involve extraction from the subject of the higher clause, while the Deep cases involve extraction of the object of the lower clause.

Subjacency Cases

Shallow: When did the Duchess sell Max's portrait of Bill?
Deep: Who did the Duchess sell a portrait of?

Depth Cases

Shallow: Who said Max likes George?
Deep: Who did the Duchess say Max likes?

Table IV: Materials for Experiment 3.

4.2. Results and Discussion

The results of Experiment 3 are displayed in Figure 3. Most importantly, there is a significant interaction between Depth (Deep vs. Shallow) and Sentence Type (Subjacency Cases vs. Depth Cases), $F(1,27)=5.69$, $p<.05$. This indicates that the difference between the two Subjacency cases is reliably larger than that between the two Depth cases. There was also a reliable main effect of Depth, $F(1,27)=46.69$, $p<.001$.

This pattern of results is not consistent with the claim that the Control vs. Indefinite contrast of Experiment 1 arose because of a depth of embedding difference between the two extraction sites.

Two notable features of these results are that the Shallow Depth cases are significantly less acceptable than the Shallow Subjacency cases (i.e., the Control cases of Experiment 1), $F(1,27)=27.94$, $p<.001$, and that there is no appreciable difference in acceptability between the two Deep cases. These observations suggest a partial alternative account of the pattern obtained in the first experiment. This account draws a sharp distinction between adjunct extractions (questions with **why**, **when**, or **where**) and those from argument positions (questions with **who**, **what**, or **which**), with the latter associated

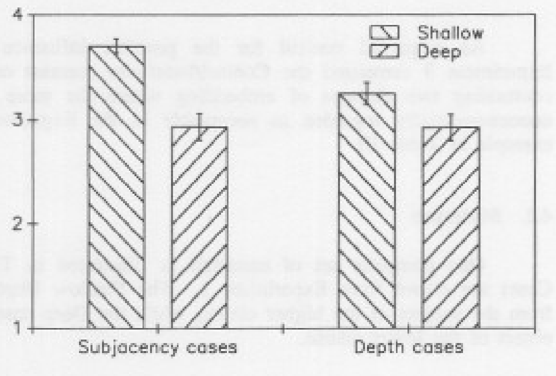


Figure 3: Mean judged acceptability for the four cases of Experiment 3.

with reduced acceptability and, presumably, greater difficulty. On this account, it was inappropriate in Experiment 1 to use non-argument extractions as controls against which to assess the acceptability of the three cases involving extraction from an argument position. It is, in other words, the distinction between argument and adjunct extraction that best explains the difference between the Control and Indefinite cases of Experiment 1, not the engagement of any effect related to subjacency. This, of course, does not explain the differences among the three cases of argument extraction, nor why the scale of those differences was less than that between the Control and Indefinite conditions.

Further experimental work that will help determine the source of the Control/Indefinite difference in Experiment 1, and which will control for extraction-type, is planned.

5. Illicit acceptability and its analysis

There are two puzzles implicit in the various findings discussed above. One is why some cases that are regarded as acceptable in the linguistic literature get such poor acceptability ratings with the methods used here.

Another puzzle is that, on either analysis of why the Indefinite cases of Experiment 1 are rated so poorly, some sentence types seem much better than might be expected from one or another point of view. If the results of Experiment 1 are interpreted as evidence that the Indefinite cases are covered by the subjacency generalization, it is surprising that these sentences have been seen as entirely acceptable in the literature. On the other hand, the low ratings of the Indefinites may be due simply to the fact that they involve extraction from an argument position, not to ungrammaticality. On this view, it seems surprising that the Definite and Specified Subject cases were not judged more negatively in Experiment 1. A proponent of this extraction-type analysis must somehow explain why there should be greater differences in judged acceptability when grammatical sentence types are compared to other grammatical types than when grammatical types are compared to ungrammatical types. Thus, from either point of view there seem to be sentences that enjoy a degree of illicit acceptability, i.e., surprisingly high degrees of acceptability associated with ungrammaticality.

Findings such as these suggest as yet undescribed complexities in the relation between strictly grammatical phenomena and the judgments of acceptability on which this study is based. Experiment 4 constitutes a preliminary attempt to identify one potential source of that complexity.

There is a widely held view of the relation between performance phenomena, such as acceptability judgments, and matters of competence, as represented by some grammatical theory. On this view, the standard approach to accounting for many apparent discrepancies between the two is to advert to features of the psychological mechanisms that implement and deploy grammatical knowledge in support of speech behavior and language comprehension. A classic example of this approach appears in the account of doubly center-embedded sentences in Chomsky and Miller (1963). Here it is maintained that the sentences are grammatical despite their manifest unacceptability on the grounds that their unacceptability arises from characteristic limitations of the memory structures used by the parser. Any discrepancy that can be analyzed in these terms is properly seen as of little or no relevance to grammatical theory. For example, in so far as there might be evidence of informal heuristic mechanisms playing some role in sentence comprehension, this is seen as arising in some way within the mechanisms that implement the grammar.

Another much less widely discussed view situates the linguistic system as but one of several cognitive resources that might be engaged in the course of language comprehension. On this view, the presentation of an utterance will typically engage several discriminable mental competencies and the utterance's effect on the listener might best be viewed as a negotiated result that integrates effects arising from two or more of the participating systems. On this view there are several competing and collaborating kinds of competence and an associated performance theory for each. For example, there are clearly extralinguistic phenomena of deixis. If the competence theory of general deixis turns out to subsume the theory of deixis in language, then the deictic system could be seen as another competency, closely linked to, but also partly independent of, the linguistic system. A similar account might be given of a mechanism that exploits pragmatic knowledge to infer what roles a given set of nouns might play relative to some verb. In that humans clearly have some ability to discern possible sentential interpretations where lexical material is presented with few or no syntactic cues, this too might be seen as another associated but independent competency.

Part of the importance of this second view is that it can lead to a quite different treatment of apparent discrepancies between observed performance and grammatical theory. It becomes possible in this context to ask whether an utterance that is, say, ruled ungrammatical by the grammar-based linguistic component, might nevertheless become acceptable through the intervention of some extralinguistic mechanism (see Bever 1974 for a discussion of some possible instances where ungrammatical forms may nevertheless be acceptable). On this view the grammatical implications of acceptability are more difficult to discern. The grammar ought to be held accountable only for those utterances whose acceptability does not arise through extra-grammatical means. Thus studies of performance interpreted in this frame have a potential to bear on linguistic theory somewhat more directly than can results interpreted in terms of the more common frame. Experimental observations of performance that argue that a given sentence type comes to be acceptable by way of the involvement of some extragrammatical competency allow grammatical theory to set aside certain sentence types that it would otherwise have to cope with.

6. Experiment 4: *Which*-effects

The role of Experiment 4 is to examine one suggestion as to how an extragrammatical mechanism might be involved in some cases similar to those in Experiment 1. If the Indefinite cases of Experiment 1 are ungrammatical, this is presumably because grammatically based mechanisms for linking the wh-element and the gap are somehow impaired by the hierarchical configuration of the sentence. If such sentences were to have their interpretability and thus acceptability restored by some extragrammatical mechanism, it is apparently the filler-gap relation that this other mechanism must address. It does not seem far fetched to suggest that a relevant mechanism could be defined along the following lines: it would maintain only a flat (i.e., non-hierarchical) lexical representation of the utterance and would simply look for overt cues to fillers and gaps. On finding a filler and a gap, it would associate them in some way that would facilitate recovery of an analysis for the entire utterance. If this mechanism were to operate in an informal and heuristic fashion, its performance would likely improve with surface features that somehow made the elements of the filler-gap more salient or conspicuous. Thus, the essential idea of Experiment 4 is simply to manipulate the saliency of the wh-element in sentences like those used in Experiment 1 to determine whether more salient wh-elements are associated with higher acceptability.

6.2. Materials

The materials for Experiment 4 were in part similar to those of Experiment 1. Pairs of sentences of the Indefinite and Specified Subject types of Experiment 1 were matched to other pairs that were identical in every respect except that in the second pair the **who** or **what** was replaced by a **which** phrase. The **which** phrase was identical for the two members of the pair. A sample set of items appears in Table V. The expectation is that if the **which** phrase makes the filler more conspicuous, acceptability will improve in the cases including this structure.

What Cases

What did Sue resent a comment about?

What did Sue resent Tom's comment about?

Which Cases

Which of the new pledges did Sue resent a comment about?

Which of the new pledges did Sue resent Tom's comment about?

Table V: Sample Materials for Experiment 4.

6.3. Results and Discussion

The results of Experiment 4 are displayed in Figure 4. The most important result is that there is no significant improvement in acceptability in the presence of the **which** phrases, $F(1,66)=2.18$, NS, though the experiment did replicate the difference between the Indefinite and Specified Subject cases of Experiment 1, $F(1,66)=140.90$, $p<.001$.

These results of course do not support the suggestion that more salient wh-elements are associated with improved acceptability. Close informal examination of the by-sentence results of Experiment 4 suggests, however, a somewhat different picture. Though further experimentation will be required to assess the generality of these effects, it appears that there were numbers of specific materials sets within which the sentences with **which** were more acceptable. The aim of further investigation will be to identify factors that distinguish such sentences from others that did not show a **which** effect.

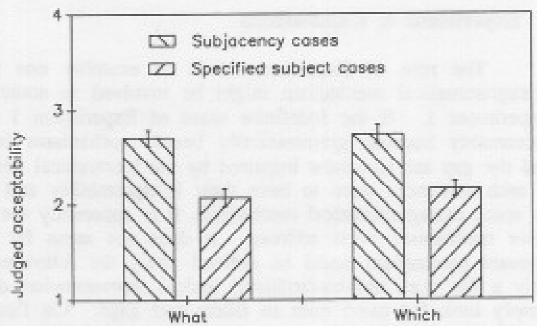


Figure 4: Mean judged acceptability for the four cases of Experiment 4.

8. A multimodal view of language comprehension

The theoretical frame of this study is somewhat different than that conventionally assumed in linguistic and psycholinguistic work that touches on these issues. In the standard view, linguistic competence is imperfectly represented in language behavior due to a variety of phenomena that arise in the psychological and neurological mechanisms that implement the speaker/hearer's knowledge of language. Usually, such phenomena do not motivate changes in the theory of competence any more than the observation that people often make mathematical errors, even systematic ones, would motivate a change in the principles of mathematics. The underlying model of language comprehension might be termed unimodal; it assumes that all utterances are interpreted via the grammar and that all departures from grammatically defined norms must be accounted for by reference to the character and limitations of the mechanisms that apply the grammar.

There is, however, another potential source of discrepancies between the form of language behavior and the principles underlying that behavior, as captured by linguistic theory. It seems quite plausible, especially in language comprehension, that there may be more than one cognitive system that can participate. Thus, while there surely is a parser that implements a grammar, there may also be one or more other cognitive systems that can be involved in comprehension, either in collaboration with the parser/grammar or more independently in cases where the parser/grammar fails to provide an analysis. The availability of such a mechanism is suggested by the seeming facility with which humans can sometimes cope with ill-formed utterances in the speech of immigrants, very young children, and individuals with anatomical or neurological impairments of speech.

Thus, an alternative view is available whenever some discrepancy between grammar and behavior is apparent. On this view, the discrepancy may arise because of some interaction between the core linguistic system and some other system, viewed as an alternative kind of competence, i.e., not as some manifestation of the mechanisms that implement grammatical knowledge. This might be termed a multimodal account of comprehension.

For any given problem, the matter can be put as a question: What is it about this particular discrepancy between grammar and behavior that shows that it is best attributed to implementational aspects of the linguistic system rather than to some interaction between the linguistic system and some other component of the cognitive system?

Where there is no compelling reason to take the implementational view, it seems necessary to consider whether the result in question may bear on linguistic theory, in particular, competence theory. That is, where discrepancies can be attributed to some other mechanism, then they need not be addressed by the theory of grammar. But where there are persistent conflicts between observed performance and what a grammar implies, and no convincing basis for attributing the difference to some other cognitive mechanism, it may be reasonable to hold the grammar responsible for those discrepancies and to consider whether it might be appropriate to modify the grammar to account for the obtained results.

As noted above, one of the arguments for a multimodal account of comprehension is that it offers some hope of accounting for the apparent resiliency of the language comprehension system. It seems plausible to suggest that a collaborating ensemble of mechanisms, each exploiting a different aspect of the information available in the utterance and its context, should, in general, be able to cope with anomalies in the input better than a more unified system. It is worth emphasizing, however, that any resiliency gained in this fashion is available only where the several mechanisms enjoy a high degree of autonomy. The more tightly the work of a given module is linked to that of another, the more vulnerable it will be to anomalies in the input that disrupt the work of the module on which it depends. It is just in so far as each module can make a useful contribution to the analysis of an utterance in the face of failures elsewhere that the multimodal model of language comprehension offers a better account of resiliency.

Acknowledgements

The work reported in this paper was supported in part by a grant from the National Institutes of Health (1 R01 NS22606-01). I am grateful to Tom Bever, Helen Cairns, Bob Fiengo, Bob Levine, Dana McDaniels, Craige Roberts, and

Michael Studdert-Kennedy for much wise counsel relevant to these matters. All responsibility for errors of fact or judgment that may be found here of course remains entirely with the author. Diana Smetters, Ken Williams, Rick Neighbarger, Andy Gudgel, and Grace Kuo have provided invaluable assistance with various aspects of the experimental work reported here.

References

Bever, T. (1974) The ascent of the specious, or, There's a lot we don't know about mirrors. In Cohen, D. (Ed.) *Explaining Linguistic Phenomena.* Washington, D.C.: Hemisphere Publishing. Pp. 173-200.

Chomsky, N. (1973) Conditions on transformations. In S.R. Anderson and P. Kiparsky (Eds.) *A Festschrift for Morris Halle.* New York: Holt, Rinehart and Winston. Pp. 232-286.

__________. (1981) *Lectures on Government and Binding.* Dordrecht: Foris.

_______. (1986) *Barriers.* Cambridge: MIT Press.

Chomsky, N. and G. Miller. (1963) Introduction to the formal analysis of natural languages. In Luce, Bush and Galanter (Eds.), *Handbook of Mathematical Psychology*, Vol. 2, Ch. 11. New York: John Wiley & Sons.

Fiengo, R. (1987) Definiteness, specificity, and familiarity. *Linguistic Inquiry* **18**: 163-166.

Fiengo, R. and J. Higgenbotham. (1981) Opacity in NP. *Linguistic Analysis* **7**: 395-421.

Lasnik, H. and J. Uriagereka. (1988) *A Course in GB Syntax.* Cambridge, MA: MIT Press.

Riemsdijk, H. and E. Williams. (1986) *Introduction to the Theory of Grammar.* Cambridge, MA: MIT Press.

OHIO STATE UNIVERSITY WORKING PAPERS IN LINGUISTICS 38 (67-87)

Interarticulator Timing and Single-Articulator Velocity-Displacement in English Stress Pairs

Kenneth deJong

Ohio State University

Abstract: Models of speech production utilize mentalist accounts of speech phenomena to varying degrees. Especially noted in this paper are accounts of the timing of speech which have sought to eliminate altogether the temporal dimension from mental control. Two major parts of the theory--that in Harris et al. (1986), and that in Kelso et al. (1985)--are explained and tested using a body of X-ray microbeam tracings of articulatory movement in English stress pairs. The first study is a replication of Harris et al. Correlations between jaw movement periods and some variables indicative of the relative timing of lip and tongue-blade movement within the period duplicate those found in Harris et al. However, eliminating the effect of vocalic expandability and the effect of a part-whole relationship between the variables renders the results ineffective in showing the tested timing relationship. The second part replicates Kelso et al. (1985). The present study finds strong correlations between velocity and displacement in jaw movement, both in upward and downward movement, in tokens in normal and frame conditions. In addition, as in Kelso et al., downward movement in stressed syllables shows a shallower regression slope than does downward movement in unstressed syllables. However, in upward movement, the relationship is reversed.

1. Introduction

Much of the work done in linguistics assumes a complex and abstract mental linguistic capability, a mental grammar. Although the actual form of this mental grammar is a matter of heated and continuous debate, the actual existence of some sort of abstract mental linguistic capability is relatively uncontroversial. Anyone interested in understanding what language is at the level of actual production must determine which speech phenomena (or which aspect of a speech phenomenon) are to be seen as indicative of part of this mental grammar, and which are more profitably described in terms of some other set of principles.

Timing is one aspect of speech production which has presented itself as a rather controversial topic in light of this question of the extent of grammatical explanation. The timing of speech events is theoretically cogent to a large range of linguistic notions besides rate, stress, and other prosodic features which come immediately to mind. The treatment of timing in any theory shapes that theory's representation of the segment. Its treatment of timing also impinges upon its account of co-articulation, since the notion "co-articulation" assumes some sort of segment whose production overlaps temporally with neighboring segments.

Approaches to timing have varied considerably in their reliance on mental grammar and their appeal to extra-grammatical principles in explanation. At one extreme are those approaches which treat the timing of articulatory events as almost exclusively the fall-out of mental processes executed upon abstract mental entities. Two examples of such mentalist approaches to handling the temporal overlap of movements associated with adjacent segments--i.e. co-articulation--are those of Wickelgren (1969), and Henke (1966) (explicated in Bell-Berti and Harris 1981, and cf., for a more general discussion of mentalist explanations of speech data, Hammarberg 1976). Wickelgren, for example, assumed a multiplication of mental categories which, in some formal sense, more closely mirror real speech sounds than, say, phonemes would. The result is a detailed, segment-by-segment speech plan which is to be converted into real speech by means of a relatively simple mechanical execution. The actual temporal dimension of such speech plans are often introduced by reference to a mental timing device such as a metronome. The rhythmic pulses generated by such a device are grammatically specified to have some temporal relationship to the various parts of the speech plan. Examples of this type of device are found in Kozhevnikov and Chistovich (1965) and, more recently, MacKay (1985), as well as in Lindblom (1963) and many others.

Non-mental approaches to timing, however, are also attested in the literature. Lindblom's early study of vowel centralization (1963) had a non-mentalist aspect to its mentalist explanation. Vowel centralization was taken to be resultant from articulatory undershoot. Too little time is alloted in unstressed syllables for the tongue body to reach the grammatically specified target for the vowel. Thus, unstressed vowels tend to be centralized. Although this account has not withstood the scrutiny of later study (cf. Gay 1981 and Harris 1978), his methodological insight--that there are articulatory principles which account for co-articulatory phenomena--has found manifestation in other more general theories of speech production, theories which posit a complex and more central physiological component in the study of articulatory timing.

These theories attempt to reduce the role of mental grammar in the explanation of articulatory timing much further than researchers such as Lindblom did. To this end, two differing approaches to timing facts have been taken. An earlier approach (Harris 1978, and Bell-Berti and Harris 1981) involved a linear organization in which the details of production are pre-programmed movements with an internal time dimension which are executed with a fixed temporal overlap with neighboring movements.

The model of speech production examined by this inquiry is a later one presented in Harris, Tuller and Kelso (1986) and various other similar works by these authors. Recently, Kelso et al. (1985) have done work on explicating a crucial aspect of this theory's account of 'timing facts' in speech events. The time specification, it is noted, can be eliminated altogether from the mental input into speech production by modeling the production mechanism as some sort of spring-mass system. During activity, this system undergoes an oscillatory process in which kinetic and potential energy are transferred back and forth. The readily-observable result of this process when mapped over time is a sine wave. (This movement, loosely speaking, is similar to that observed of articulators in real speech.) The temporal dimension then is mathematically dependent on the properties of the spring-mass system--linear stiffness, total mass, etc.--and the perturbing forces acting upon it.

Thus, the clearest evidence cited by Kelso et al. for their approach are strong correlations between the peak velocity of jaw movement and the distance the jaw traverses. Such correlations would be expected within the model, since peak velocity and displacement are both indicative of the energy stored in the spring-mass system. The system in oscillation transfers the energy in one form--potential energy, indicated by displacement from rest position--into energy of the other form--kinetic energy, indicated by peak velocity. Inertia then causes the transfer to be reversed. All three of the variables--peak displacement, peak velocity, and the overall time which the transfer takes (indicated by one quarter of the period of the movement)--are all dependent on the properties of the spring-mass system, as is shown in Figure 1.

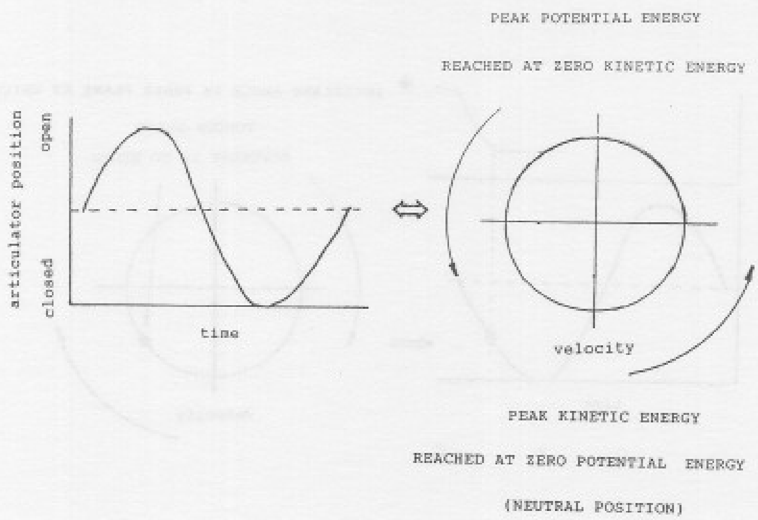


Fig. 1: An illustration of the relationship between a representation of articulatory movement over an explicit time dimension and one utilizing articulator velocity and displacement.

The second part of the theory is explicated in Harris et al. (1986). Normal speech, of course, involves the use of not just one articulator, but many. These articulators have been noted to work co-operatively in the production of various speech segments. For example, in sounds involving a bi-labial stop closure, the jaw and both of the lips co-operate to close off the vocal tract. Due to the immediacy of the co-operation, and the speaker's common unawareness of it, researchers such as Fowler (see especially, Fowler et al. 1980, and as a classic example of the research on which Fowler's position is based, Abbs and Netsell 1973) and others, have positted a low level yoking of the muscles controlling the articulators involved into what is termed a 'co-ordinative structure.'

Harris et al. suggest that the temporal relationship of movements associated with neighboring segments are resultant from the yoking together of muscles (and articulators) at a higher level in a phase relationship. Such phase relationships are exhibited, for example, in the various movements implemented in walking or running. Each of the individual movements comes at a predictable time within the overall stride. If the rate of the stride is increased, rate of each movement within the stride is proportionally increased. Thus, the temporal position of one movement is mathematically predictable from the temporal position of other movements. An illustration of this relationship is given in Figure 2.

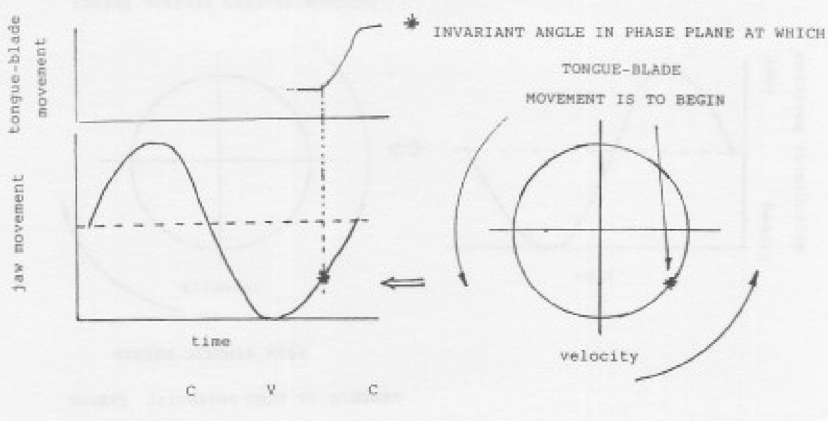


Fig. 2: An illustration of phase plane interarticulatory timing.

The timing of local movement associated with the production of a consonant--shown here as tongue-blade movement--is to occur at an invariant angle in the phase plane portrait depicting the cyclic movement associated with vowels--shown here being indexed by the upward and downward movement of the jaw.

Harris et al. have taken the vowel period to be a global movement similar in kind to the stride in walking. Various other articulator movements come at predictable times relative to this movement. For example, the production of intervocalic consonants is automatically timed in reference to the production of the vowels. The faster the rate of the movement associated with the vowel's production, the faster the rate of the production of the consonants. Thus, to demonstrate this behavior, Harris and others have shown correlations to exist between size of the vowel period--measured as the time between the onset of jaw movement associated with two successive vowels--and the latency of the onset of consonantal (lip or tongue blade) movement from the beginning of the vowel period in the production of various nonsense syllables (e.g. in Harris et al. 1986 and Tuller et al. 1983).

The interpretation of these correlations as evidence for some fixed timing relationship between consonant and vowel production has been criticized on several grounds. Barry (1986) points out several ways in which the correlations could be the artifact of some other fairly well-attested facts. One worrisome criticism has to do with the relative temporal expandability of the vowel, as opposed to the temporal insensitivity of the consonants (cf. Gay 1980). Over variations in rate and stress conditions, a disproportionate share of the temporal expansion or contraction will occur in the vocalic portion of an utterance. Thus, any portion of the vocalic period which is included in both the latency and the period will tend to produce higher correlations between the latency and period.

Another criticism has to do with the interpretation of the obtained correlations in light of a part-whole relationship between the two correlated variables (see Munhall 1986). If one variable is a part of the other variable, a significant correlation between them is expected. A third criticism of these studies points out that both Harris et al. (1986) and Kelso et al. (1985) used a form of non-speech data, either reiterant speech or nonsense words. It is, thus, not clear that their results can be taken as valid for normal speech.

This study will seek to replicate these earlier experiments using real English words of varied stress patterns, placed in two contexts, natural and frame context. It also attempts to take into account the two criticisms noted above. A body of X-ray microbeam data was analyzed. For a replication of Harris et al. (1986), correlations were taken between consonant and vowel periods in jaw movement, and the relative positions of various events in tongue-blade and lip movement. Effort is made to factor out the contribution of vocalic expandability by comparing the predictive power of vowel and consonant periods on the timing of the production of consonants. A second study was also performed, a replication of Kelso et al.'s experiment concerning the relationship between the velocity of jaw movement and the distance it traverses.

2. The Data-base

The data involved in this study consist of the articulatory trajectories of a male speaker of an East Coast dialect of American English. The trajectories were obtained with a computer-controlled X-ray microbeam system at the University of Tokyo. This system utilizes X-ray microbeams which alternately track various metal pellets attached to various articulators during the production of an utterance (see Kiritani et al. 1975 for a more detailed description of this system). The data-base, then, consists of the relative vertical and horizontal (longitudinal) position of each of these pellets--placed on the velum, tongue blade, tongue dorsum, lower lip, mandible, and a reference pellet placed on the nose--after successive time increments. In this study, only three of the pellet traces were used--lip, jaw, and tongue blade.

The analysis of the data-base was performed using the X-ray Database Display Program developed by Joan Miller, implemented on a PC6300 under MS-DOS. This system produces both a one-dimensional display of articulator movement over time (examples of this kind of display are Figures 3 - 5), and a two-dimensional display. The system also includes algorithms for producing velocity functions and combining trajectories in various ways. These algorithms were used to subtract out the contribution of one articulator's movement from that of an other.

The corpus, designed by Mary Beckman, includes eight bi-syllabic English stress pairs involving intervocalic alveolar or palatal consonants. Two pairs involving labial consonants were also studied. Some of the target words occurred in normal English sentences designed to hold the environment of the target as constant as possible; others occurred in the frame sentence, 'Say __________ again.' The targets and sentences are given in Table I. Roughly half of the targets were placed in each context. Most of the natural sentences were repeated twice within each run, while the short frame sentences were repeated from 3 to 5 times in a given run. The number of cases of each pair ranged from 14 to 16. 148 cases were analyzed altogether.

3. Study 1

3. 1. Methods

The temporal distance between the onset of jaw lowering for the first vowel and the onset of movement associated with the second vowel was measured. This measure is taken to be the length of the vowel period. As a matter of comparison, temporal measurements were taken at the zero-velocity point in the negative position

digest (N)	But THIS is how the Reader's Digest talks.
digest (V)	But THIS is how the readers digest talks.
contract (N)	Will you use the older pipe contract as a model for it?
contract (V)	Will even the older pipe contract as much as that?
antics	Their antics are intolerable.
antiques	Their antiques are adorable.
content (N)	You DO know about the content of it.
content (ADJ)	You DO know how to be content about it.
retake (N)	The director wanted a retake the second he saw the developed footage.
retake (V)	The general wanted to retake the second city he had lost.
suspect (N)	The police also think that the FBI-suspect did the crime.
suspect (V)	This is the man that the FBI suspect did the crime.
insight	The poet's insight hurts.
incite	The poets incite hurts.
insult (N)	The soldier's insult counts.
insult (V)	The soldiers insult counts.
deepened	The mountain lakes deepened upon the melting of the winter snow.
depend	The mountain lakes depend upon the melting of the winter snow.
defects (N)	The Russian's major defects and their minor virtues are the same.
defects (V)	The Russian major defects and the corporals under him are punished.

Table I: The corpus.

for the vowels. The difference between these was taken to be indicative of some sort of consonantal period. Example of such marks are shown in Figure 3.

Events in the vertical movement of the tongue blade--or lip, in tokens with labial intervocalic consonants--were also studied. Blade movement was ascertained, as was suggested in Edwards (1985), by finding the difference between the jaw position at a given moment and the mean position of the jaw over the entire utterance. This figure was multiplied by a coefficient reflective of the difference in distance of the jaw pellet and the blade pellet from the jaw hinge--0.8 in the case of the pellet which was attached to the tongue blade. Figure 4 is an example of this spatial relationship. Since the jaw pellet is placed further out on the mandible than the tongue-blade pellet is, it traverses a greater distance for a given swing of the mandible than the tongue-blade. The relationship between the jaw pellet and the lip pellet is, however, the reverse. The lip pellet moves a greater distance than the jaw pellet for a given swing of the jaw. Thus, the coefficient used in subtracting out the jaw's contribution to the observed lip movement is 1.1.

Three events in the movement of the tongue blade were measured--the point at which movement toward the consonant closure began, the point at which that same movement ended, and, finally, the point at which movement began toward the position assumed for the second vowel. Figure 3 is a relatively clear example of the two traces with the appropriate timing marks.

It should be noted that not all of the pairs presented such clearly defined movements. For example, although movement associated with alveolar consonants is usually registered in the movement of the tongue blade, this movement is sometimes obscured by movement associated with neighboring high vowels. The occurrence of /r/, e.g. in **retake**, causes a major upward shift in the movement of the tongue blade. Difficulties are also encountered in radically reduced, unstressed vowels. Often in such words as **depend** and **insult** (V), no vertical jaw-movement is observed during the first vowel. In such unclear cases, the alternative method of display, plotting the vertical and horizontal position of the pellets on an x/y grid,

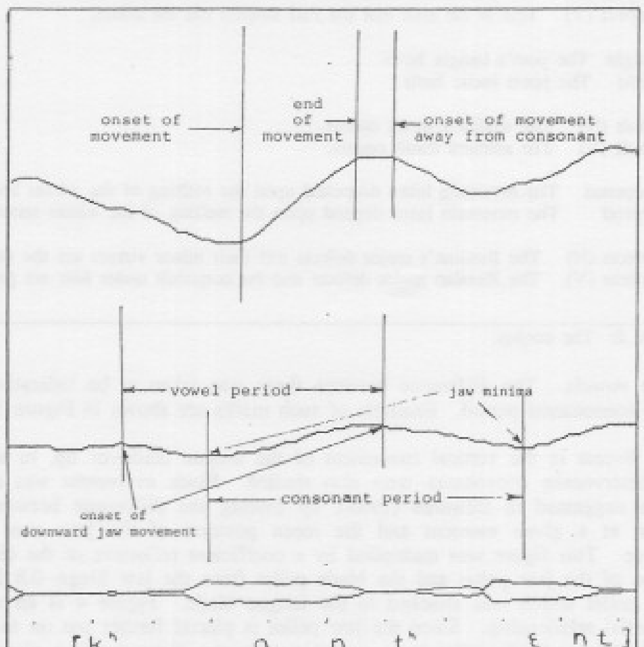


Fig. 3: A tracing of the vertical movement of the tongue blade and jaw during the utterance "Say CONtent again." The timing marks in a latency-period correlation are indicated by vertical lines.

will often reveal subtle changes in the rate and direction of articulatory movement not apparent in the vertical dimension traces. In the few remaining cases, acoustic cues were used, e.g. the alveolar release in **depend** was taken as the beginning of the consonantal period, since there is no jaw-movement associated with the first vowel, as well as no periodic acoustic component.

The reader is referred to Harris et al. (1986) for a short discussion of the validity of these trajectory events as indicators of the events within their model. Unlike that study, only one mark was taken for each case before any of the correlations were calculated.

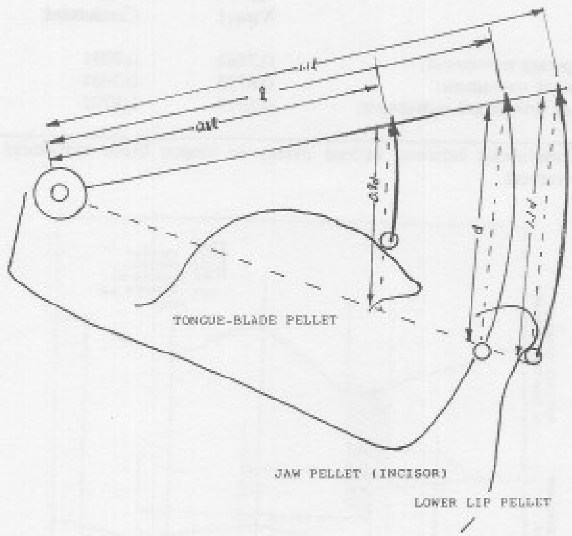


Fig. 4: An abstract depiction of the spatial relationship of the tongue blade and lip pellets to the jaw pellet (see Edwards 1985).

3. 2. Results and discussion

Correlations between latencies calculated with respect to the onset of consonant-articulator movement proved consistently to be the largest of the three events measured as is shown in the figures in Table II. This result, consistent with the findings in Harris et al., suggests that the beginning of the execution of a movement is more salient for the temporal structure of speech than an ending point--target for the consonant--or the onset of tongue-blade or lip movement for the production of the following vowel. These results fit in well with both their contention that articulatory targets are not important in the temporal organization of speech, and that the movement of the smaller articulators--lip and tongue blade, as opposed to jaw--is superimposed over movements associated with the production of the vowels. The differences in r shown here are rather small, but are consistent over the various methods of calculating latencies. The rest of the discussion will, thus, center upon the timing of the onset of movement as opposed to the other two events.

Event	Vowel to Vowel	Consonant to Consonant
Onset of upward movement:	0.7866	0.7031
End of upward movement:	0.6723	0.5924
Beginning of downward movement:	0.6519	0.6792

Table II: Correlations between various events in tongue blade movement and larger periods of movement.

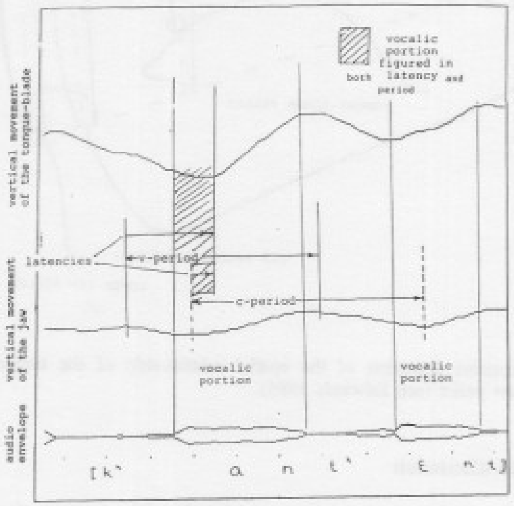


Fig. 5: The amount of the vocalic portion (shaded portion) included in vowel period (upper part) and consonant period (lower part) correlations with the latency of the onset of tongue blade movement.

Correlations were calculated between the latency of the onset of consonant-articulator movement from the beginning of the two periods, and the length of the vowel period and the length of the consonant period as is shown in Figure 5. To investigate the effects of vocalic expandability brought up by Barry, correlations were also calculated between the latency of the end of the periods from the onset of movement associated with the consonant, as is illustrated in Figure 6, and the length of the periods. Correlations are expected to be greater when calculated in the first manner for the vowel period than for the consonant period, since more of the vocalic portion of the utterance is figured into the vowel correlations than into the consonantal correlations. As is shown by the shaded portions of Figure 5, the shaded material in the vowel period latency is larger than that in the consonantal period latencies.

Calculating correlations in the second manner should preserve any indicators of the posited phase relationship, while switching the contribution of the vocalic portion to the vowel period correlation to the consonantal period correlations. In Figure 6, the shaded portion for the consonantal latency is larger than that for the vowel period.

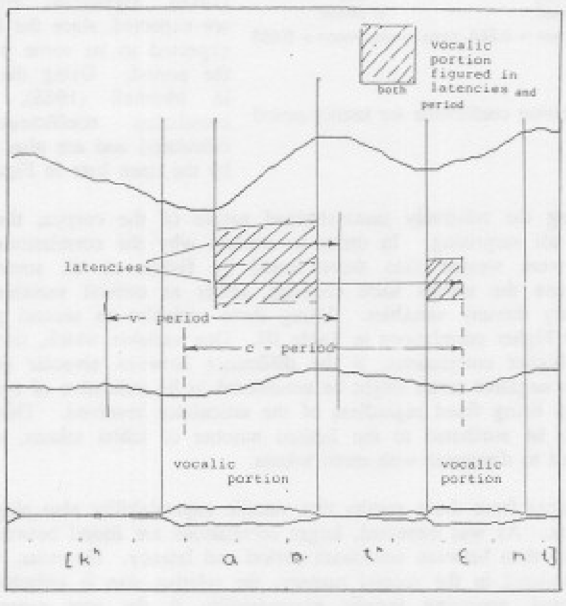


Fig. 6: As in Figure 5. However, the shaded portion indicates the amount of the vocalic portion included when the timing of consonant related movement is ascertained in an alternative method.

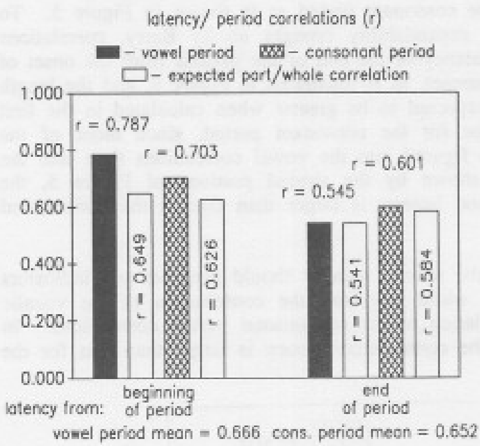


Fig. 7: The regression coefficients for latency-period correlations.

The results taken across all word and stress conditions are shown in Figure 7. The group shown on the left are for the method used in Harris et al. (1986). The group on the right are for an alternative method of ascertaining the relative timing of consonantal movement. Dark boxes are for the vowel period. Light boxes are for the consonantal period. Correlations do exist in the English utterances tested. These results, again, are in keeping with those found in Harris et al., but are slightly weaker. As was pointed out in Munhall (1986) and Barry (1986), significant correlations are expected, since the latency is expected to be some portion of the period. Using the formula in Munhall (1986), expected correlation coefficients were calculated and are also indicated by the open bars in Figure 7.

Considering the relatively unconstrained nature of the corpus, the weaker correlations are not surprising. In order to explain why the correlations for the present corpus were weaker than those found by Harris et. al, some of the differences between the tokens were encoded, either as ordinal variables where possible, or binary dummy variables. Using these variables as second predictors yield the slightly higher correlations in Table III. One variable which, surprisingly, does not yield higher correlations, is the difference between alveolar and labial consonants. This negative result might be considered to be indicative of the internal phase relationship being fixed regardless of the articulator involved. This lack of result might also be attributed to the limited number of labial tokens, and thus, would be expected to disappear with more tokens.

It is apparent from these results that vocalic expandability also plays a part in the correlations. As was expected, larger correlations are found between vowel period and latency than between consonant period and latency. However, when the latencies are calculated in the second manner, the relative size is switched. This pattern is expected, assuming vocalic expandability is the root cause of the difference between the correlations of the two periods. Assuming that the amount of contribution of vocalic expandability to the differences found in the two versions of the calculations are equal, one can neutralize this contribution by taking the mean

Predictors	Vowel to Vowel	Consonant to Consonant
latency X period	0.7866	0.7031
period, stress	0.8007	0.7044
period, articulator	0.7867	0.7034
period, # of cons.	0.7855	

Table III: Latency-Period correlations with multiple predictors.

of the two conditions. The difference between these means is remarkably small, suggesting that the vowel period and the consonant period are equally well-suited to predict the relative temporal position of the onset of tongue-blade movement.

One final result of note here is the large main difference between the correlations figured in the first manner and those figured in the second manner. This difference might be resultant from the figuring of the latencies across the intervocalic syllable boundary. It is a fairly well-attested fact that syllables differing in stress show rather global articulatory and acoustic differences (e.g. Summers 1987). The differences between the two stress-patterns would introduce unaccounted-for variation in the latencies, and, thus, would yield lower correlations. If this is the explanation for the lower correlations, it suggests that both the vowel period and the consonant period might be too small to account for relative timing of articulatory events across stress conditions. Also lending credence to this approach is the slight effect of adding the stress pattern as a predictor of latency. Of the differences between the types that were encoded, stress is the variable which made the greatest increment in r-squared. Indeed, stress alone correlates significantly with latency, accounting for about 17 percent of the variance in latency ($r = 0.412$).

A direction for future research would be to investigate internal timing relationships in terms of a larger unit of production. It might be more useful, for example, to predict the relative position of the various internal movements within a stress foot by making reference to the overall length of the stress foot. The global unit of production, then, would not be a single cycle of jaw movement, but a more abstract unit of two, or maybe even three cycles. Each cycle would have differing but predictable dynamic properties. For example, trochaic stress feet might consist of a large articulatory cycle, followed by a more restrained cycle. The details of production, according to the phase relationship theory, should then be predictable on the basis of the size of the stress foot.

4. Study 2

4. 1. Methods

The second experiment is an attempt to replicate the velocity-displacement correlations found in Kelso et al. It utilized the same trajectory data-base that was used in the first experiment. Four events in jaw movement were marked and the relative jaw position taken for each event. First, the jaw position at the onset of movement associated with the second vowel as well as the upward position associated with the first medial consonant were marked. The jaw minima associated with each vowel were also taken as is shown in Figure 8. Two displacements were then calculated--downward displacement for the second vowel, and upward displacement for the offset of the first vowel. A velocity function was calculated from the raw vertical trace of jaw movement. Peak upward velocity of the jaw as well as peak downward velocity associated with the production of the medial consonants were then taken.

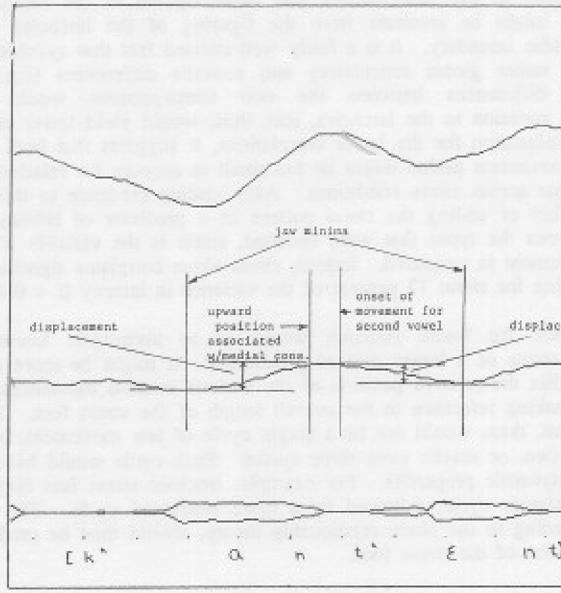


Fig. 8: The same tracings as in Figure 3. Shown here are the (vertical) timing marks and the (horizontal) displacement marks necessary for computing velocity displacement correlations.

4. 2. Results and discussion

Figures 9 - 12 are scatter plots showing every token. As can be seen in these figures, in both the upward movement and the downward movement, and in both natural sentence and frame contexts, peak velocity and amplitude of movement are strongly correlated. These correlations are of the same magnitude as those found in Kelso et al. of reiterant speech. At the outset, I wondered if the correlations might be indicative of a energy-minimizing strategy induced by the repetitive nature of the reiterant speech task. This seems unlikely, since strong correlations are found both in the natural sentences and in often-repeated frame sentences.

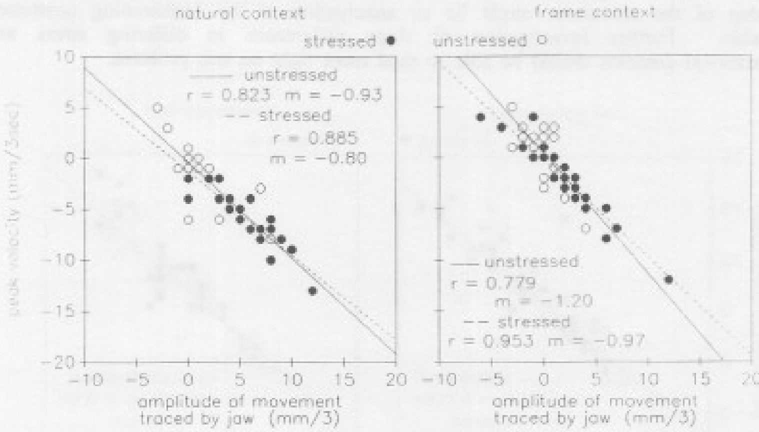


Fig. 9: Scatter plots showing all of the tokens. The velocity (vertical axis) and displacement (horizontal axis) plotted here are for the downward jaw movement into the first vowel of the target words.

One further similarity between the results found here and those in Kelso et al. is shown in Figures 9, 11 and 12. The left graph shows the tokens placed in natural sentences; the right graph shows the tokens placed in an often repeated frame sentence. Solid circles and dashed regression lines indicate movements associated with a stressed syllable, while empty circles and solid regression lines indicate movements associated with unstressed syllables. There is a difference in the slope of the regression lines for stressed and unstressed syllables. Stressed syllables show a shallower regression curve than unstressed. This difference would be indicative of a spring-mass system of less linear stiffness. Thus, a difference noted here between stressed and unstressed syllables is in the rigidity of the articulators, unstressed syllables being marked by a relatively greater resistance to movement. Thus, the jaw traces a period of shorter duration, as well as opening less during unstressed syllables. Both of these features are commonly noted of unstressed syllables--lower amplitude, and shorter duration.

These results, however, are not consistent. Figure 10 shows exactly the opposite patterns in the closing gesture. Stressed syllables show a steeper regression line--indicative of a stiffer spring-mass system. Two explanations for this anomalous result come to mind. One explanation might assume the tensing of the jaw to be an expending of greater effort in the production of the stressed syllable coda. Another explanation would see the difference as a dynamic effect. The tensing of the articulator might be an anticipation of the forthcoming unstressed syllable. Further investigation of slope differences in differing stress and intonational contexts should be able to shed more light on this problem.

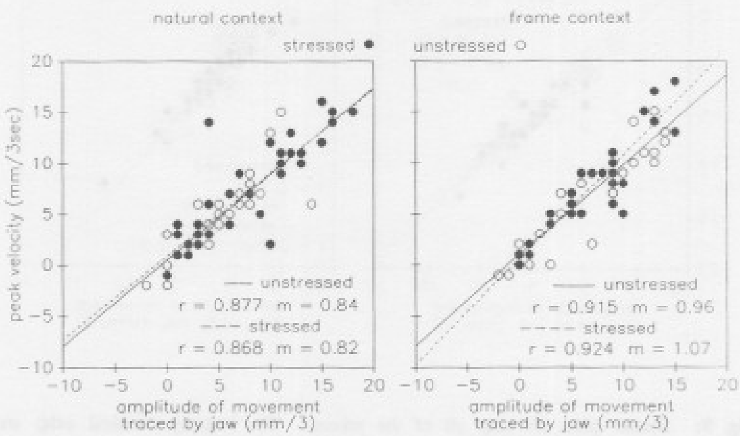


Fig. 10: As in Figure 9. The values plotted here are for the upward jaw movement into the medial consonants in each token.

5. Conclusion

The overall results of the present study duplicate those found in Harris et al. and Kelso et al., suggesting that the models supported in these earlier studies by reiterant speech are extendable to non-reiterant speech and the production of real lexical items. Although there are differences between the data used in this experiment and the natural speech of naive subjects--the presence of metal pellets in the oral cavity, the repetition of well-rehearsed sentences--and although the data is limited to the speech of one speaker, this study is a step toward the application of these models to natural speech.

One of the speakers in Kelso et al. yielded notably smaller velocity-displacement correlations than the other did. Whether this poor correlation is an artifact of the reiterant speech paradigm or indicative of a more general pattern to be found of naive speakers in natural settings is a question for further study. This study shows that the correlation found in the other speaker's production is probably not an artifact of the reiterant speech paradigm.

The velocity-displacement correlations found here are evidence for the validity of the spring-mass model in describing articulatory movement. Its relatively simple account of two of the multiple cues associated with stress--duration and intensity--is rather elegant in that it allows the reduction of the two features to one root cause in languages such as English, where intensity and duration are inextricably connected to stress. But the differing regression slopes found in differing conditions suggest, if the spring-mass model is right, that either there are interactions between neighboring syllables, or that higher level prosodic effects are registered in the stiffness of the spring-mass system.

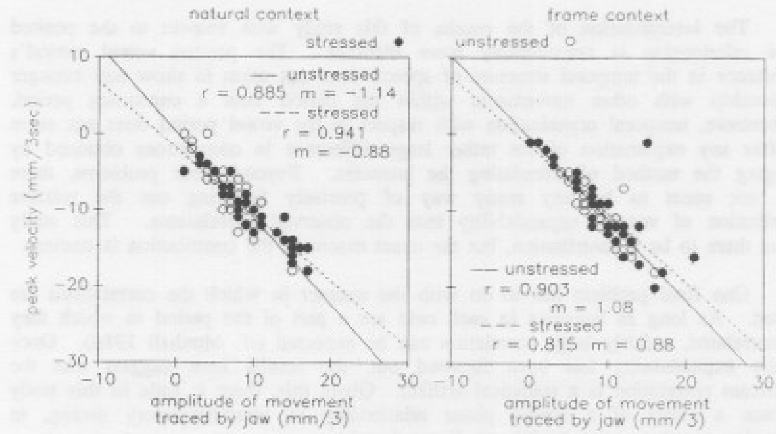


Fig. 11: As in Figure 9. The values plotted here are for the downward jaw movement out of the medial consonants into the second vowel.

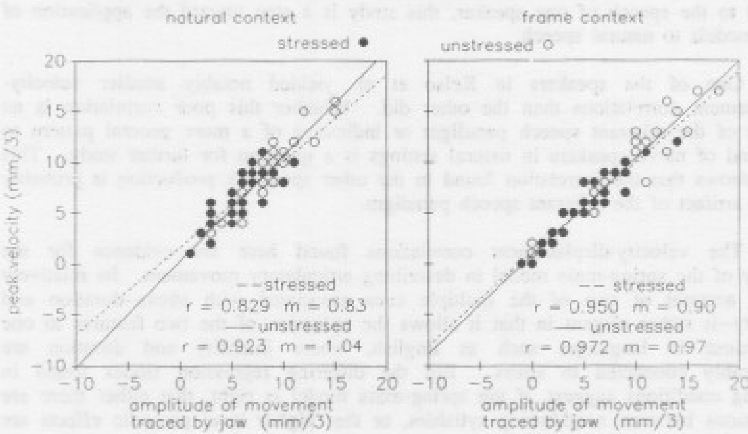


Fig. 12: As in Figure 10. The values plotted here are for the upward movement out of the second vowel.

The interpretation of the results of this study with respect to the posited phase relationship is considerably more difficult. The posited vowel period's importance in the temporal structure of speech does not seem to show any stronger relationship with other movements within the period than a consonant period. Furthermore, temporal organization with respect to the vowel period does not seem to offer any explanation of the rather large difference in correlations obtained by changing the method of calculating the latencies. Beyond these problems, there does not seem to be any ready way of precisely factoring out the relative contribution of vocalic expandability into the observed correlations. This study shows there to be a contribution, but the exact extent of the contribution is unclear.

One final problem has to do with the manner in which the correlations are figured. As long as latencies in each case are a part of the period to which they are correlated, a fairly large correlation can be expected (cf. Munhall 1986). Once vocalic expandability has been factored out, the results here suggest that the significant correlation is a statistical artifact. Given this, there is little in this study to base a claim of a scaled phase relationship in interarticulatory timing, in opposition to other, more traditional, linear theories.

Acknowledgements

The author is indebted to Jan Edwards, Keith Johnson, Susan Nittrouer, and Eric Vatikiotis-Bateson for their useful comments during various stages of this research. I especially want to thank Mary Beckman for valuable commentary and guidance, for encouragement, and for taking care of a considerable portion of the large amount of overhead entailed by getting into articulatory research.

References

Abbs, J. H. 1986. 'Invariance and Variability in Speech Production: A Distinction Between Linguistic Intent and its Neuromotor Implementation.' In Perkell and Klatt, eds., *Invariance and Variability in Speech Processes*.

Abbs, J. H. and R. Netsell. 1973. 'A Dynamic Analysis of Two-dimensional Muscle Force Contributions to Lower Lip Movement.' *JASA (Journal of the Acoustical Society of America)* **53(1)**: 295.

Barry, William J. 1983. 'Some Problems of Interarticulator Phasing as an Index of Temporal Regularity in Speech.' *Journal of Experimental Psychology* **9(5)**: 826-828.

Bell-Berti, Fredericka and Katherine Harris. 1981. 'A Temporal Model of Speech Production.' *Phonetica* **38**: 9-20.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Edwards, Jan. 1985. 'Contextual Effects on Lingual-Mandibular Coordination.' *JASA* **78(6)**: 1944-1948.

Fowler, Carol. 1980. 'Coarticulation and Theories of Extrinsic Timing.' *Journal of Phonetics* **8**: 113-133.

-----------. 1986. 'An Event Approach to the Study of Speech Perception from a Direct Realist Perspective.' *Journal of Phonetics* **14**: 1-27.

Fowler, C. A., P. Rubin, R. E. Remez and M. T. Turvey. 1980. 'Implications for Speech Production of a General Theory of Action.' In B. Butterworth, ed., *Language Production*, Vol. I, p. 373-420. New York: Academic Press.

Gay, Thomas. 1981. 'Mechanisms in the Control of Speech Rate.' *Phonetica* **38**: 148-158.

Hammarberg, Robert J. 1976. 'The Metaphysics of Coarticulation.' *Journal of Phonetics* **4**: 353-363.

Harris, Katherine. 1978. 'Vowel Duration Changes and its Underlying Physiological Mechanisms.' *Language and Speech* **21/4**: 356-361.

Harris, Katherine, Betty Tuller and J. A. S. Kelso. 1986. 'Temporal Invariance in the Production of Speech.' In Perkell and Klatt, eds., *Invariance and Variability in Speech Processes.*

Keating, Patricia. 1985. 'CV Phonology, Experimental Phonetics, and Coarticulation.' *UCLA Working Papers in Phonetics* **62**: 1-13.

Kelso, J.A.S., Eric Vatikiotis-Bateson, Elliot L. Salzman and Bruce Kay. 1985. 'A Qualitative Dynamic Analysis of Reiterant Speech Production: Phase Portraits, Kinematics, and Dynamic Modeling.' *JASA* **77(1)**: 266-280.

Kiritani, S., K. Itoh, and O. Fujimura. 1975. 'Tongue-pellet Tracking by a Computer-controlled X-ray Microbeam System.' *JASA* **57(6)** Part II: 1516-1520.

Kozhevnikov, N. A. and L. A. Chistovich. 1965. *Speech: Articulation and Perception.* U.S. Dept. of Commerce, JRPS 30-543.

Lindblom, Bjorn. 1963. 'Spectrographic Study of Vowel Reduction.' *JASA* **35(11)**: 1773-1781.

Lindblom, Bjorn E. F. and Johan E. F. Sundberg. 1971. 'Acoustical Consequences of Lip, Tongue, Jaw and Larynx Movement.' *JASA* **50(4)** Part 2: 1166-1179.

MacKay, Donald G. 1982. 'The Problems of Flexibility, Fluency, and Speed-Accuracy Trade-off in Skilled Behavior.' *Psychological Review* **89(5)**: 483-506.

Munhall, Kevin G. 1985. 'An Examination of Intra-articulator Relative Timing.' *JASA* **78(5)**: 1548-1553.

Nelson, Winston. 1983. 'Physical Principles for Economics of Skilled Movements.' *Biological Cybernetics* **46**: 135-147.

Nelson, Winston L., Joseph S. Perkell and John R. Westbury. 1984. 'Mandible Movements During Increasingly Rapid Articulations of Single Syllables: Preliminary Observations.' *JASA* **75(3)**: 945-951.

Nittrouer, S., K. Munhall, J.A.S. Kelso, B. Tuller, and K.S. Harris. 1988. 'Patterns of Interarticulator Phasing and Their Relation to Linguistic Structure.' *JASA* **84(5)**: 1653-1661.

Ohala, John. 1975. 'The Temporal Regulation of Speech.' In Fant and Tatham, eds., *Auditory Analysis and Perception of Speech*. London: Academic Press.

Öhman, S. E. G. 1966. 'Coarticulation in VCV Utterances.' *JASA* **39(1)**: 151-168.

Summers, Van. 1987. 'Effects of Stress and Final-Consonant Voicing on Vowel Production: Articulatory and Acoustic Analyses.' *JASA* **82(3)**: 847-863.

Tuller, Betty, J. A. S. Kelso and Katherine Harris. 1983. 'Converging Evidence for the Role of Relative Timing in Speech.' *Journal of Experimental Psychology* **9**: 829-833.

Wickelgren, Wayne. 1969. 'Context-sensitive Coding, Associative Memory, and Serial Order in (Speech) Behavior.' *Psychological Review* **76**: 1-15.

Wood, Sidney. 1986. 'The Acoustical Significance of Tongue, Lip, and Larynx Maneuvers in Rounded Palatal Vowels.' *JASA* **80(2)**: 391-401.

Ohio State University Working Papers in Linguistics **38** (88-104)

VCV Coarticulation in Arabic

Lutfi Hussein

Ohio State University

Abstract: Vowel-to-vowel coarticulation in VCV utterances has been the subject of several studies. Öhman (1966) found that vowels in VCV utterances in English and Swedish have trans-consonantal effects on one another. He also found some evidence suggesting that secondary articulation features like palatalization in Russian block coarticulation. Action theorists, such as Fowler (1983), explain V-to-V coarticulation in terms of universal principles of speech timing; that is, they claim that vowels in speech production are underlyingly overlapping and consonants ride on top of the vowels. This suggestion implies that intervocalic consonants, regardless of whether they have secondary articulation features, do not block coarticulation. Keating (1985), on the other hand, explains it in terms of autosegmental phonology. She places the features for vowels and consonants on two separate tiers, and leaves consonant features unspecified for vowel features, so that V-to-V coarticulation is an interpolation between vowel targets. Keating's model implies that consonants that have secondary articulation (i.e. vowel features) must block coarticulation. 72 VCV utterances which include combinations of all vowels in Standard Arabic and a set of four pharyngealized consonants and their nonpharyngealized counterparts have been acoustically analyzed to assess the validity of the two models. The final analysis of the data indicates that V-to-V coarticulation is not as simple as either of the two models claims it to be. Several other factors such as the identity of the vowel included in the sequence, the speaker, and the direction of coarticulation (anticipatory versus carryover) have proven to be crucially important in accounting for V-to-V coarticulation.

1. Introduction

Vowel-to-vowel coarticulation in VCV utterances has been the subject of several acoustical and perceptual studies during the last two decades (Öhman 1966; Purcell 1979; Fowler 1983; Keating 1985; Recasens 1985, 1986). Öhman (1966), a pioneer in investigating vowel-to-vowel coarticulation, found that vowels in Swedish and English exhibit systematic coarticulatory effects on one another across intervocalic stops in VCV utterances. He also found that neither fricatives in Swedish and English nor palatalized stops in Russian permit similar systematic coarticulatory effects in VCV utterances. This outcome had been a primary incentive for the many studies and models that followed, among which the most prominent are those of Fowler (1983) and Keating (1985).

Fowler (1983) explains coarticulation in terms of timing. She claims that vowels are underlyingly overlapping in speech production and consonants are superimposed on the vowels. Thus, speech is primarily a continuous production of vowels. The following figures, where 1A is taken directly from Fowler (1983), elaborate Fowler's model.

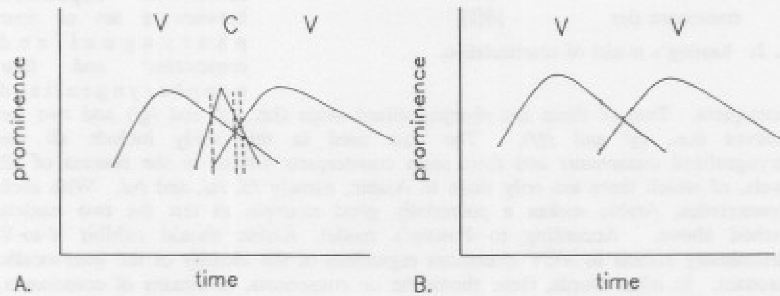


Fig. 1: Fowler's model of coarticulation.

According to this model, V-to-V coarticulation occurs because there is no temporal separation of vowels in either 1A, where the intervocalic consonant is "superimposed" on the vowels, or 1B where there is no intervocalic consonant. This model implies that intervocalic consonants, regardless of what phonological features they may have, do not block V-to-V coarticulation. Such a claim obviously contradicts Öhman's findings which indicate that some consonant classes such as fricatives in English and Swedish and palatalized consonants in Russian block coarticulation, and Purcell's (1979) finding, which is derived from an extensive study on palatalization, that palatalized consonants in Russian indeed block coarticulation.

Keating, by contrast, explains coarticulation in terms of feature association patterns in autosegmental phonology. See Figure 2 below for illustration. She places the features for vowels and consonants on two separate tiers, and leaves consonants unspecified for vowel features, so that V-to-V coarticulation is an interpolation between vowel targets. This model claims that intervocalic consonants that use the vowel tiers in the production of consonants with secondary articulations such as palatalized consonants in Russian, velarized /ɫ/ in Catalan, pharyngealized consonants in Arabic, etc. should block V-to-V coarticulation in VCV sequences. It also implies that consonants with no vowel features should not block coarticulation. This model, though it elegantly accounts for the behavior of palatalized consonants in Russian, lacks the explanatory power to account for the behavior of intervocalic fricatives in English and Swedish as reported by Öhman.

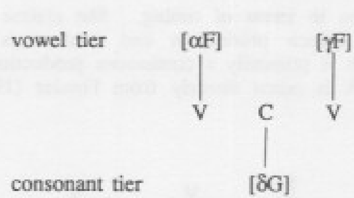


Fig. 2: Keating's model of coarticulation.

In this study I acoustically analyze a set of VCV utterances from Arabic to further assess the validity of the two models. Arabic has a distinctive opposition between a set of four pharyngealized consonants[1] and four nonpharyngealized counterparts. Two of these are pharyngealized stops (i.e., /t̲/[2] and /d̲/) and two are fricatives (i.e., /s̲/ and /ð̲/). The data used in this study include all the pharyngealized consonants and their plain counterparts uttered in the context of all vowels, of which there are only three in Arabic, namely /i/, /a/, and /u/. With such characteristics, Arabic makes a potentially good example to test the two models sketched above. According to Fowler's model, Arabic should exhibit V-to-V coarticulatory effects in VCV utterances regardless of the identity of the intervocalic consonant. In other words, there should be no consonants, or classes of consonants, that block V-to-V coarticulation in Arabic. Keating's model, on the other hand, predicts that Arabic should only exhibit V-to-V coarticulation when the intervocalic consonant is nonpharyngealized. It places no constraints on whether any nonpharyngealized consonants are likely to block coarticulation, as is the case with fricatives in English and Swedish.

Another related issue that will be discussed in this paper is whether coarticulatory effects are restricted to formant transitions or extend into the steady state. Early acoustic studies on coarticulation reported that coarticulatory effects in VCV utterances are limited to the transitions of the vowels (Öhman 1966). More recent studies have shown that transconsonantal coarticulatory effects can extend into the steady state of the adjacent vowels as well as the transitions (Manuel and Krakow 1984).

---

[1] Several terms such as *emphatic* (Jakobson 1957; Ali and Daniloff 1972; Bonnot 1977 and 1979), *mufaxxama* (Jakobson 1957), *velarization* (Obrecht 1961), and *pharyngealization* (Ali and Daniloff 1972; Ghazeli 1977; Card 1983) have been used to describe the distinctive opposition among a set of consonants in Arabic. According to cinefluorographic studies, none of these terms is exclusively accurate in describing the distinction. All these terms, however, have been used interchangeably by many linguists. In this study I will, following some tradition, be using the term *pharyngealization* to denote this distinction though the production process for these sounds involves more than just pharyngealization.

[2] Underlining is used in this study to denote pharyngealization. Thus /s/ is a plain dental fricative in Arabic, but /s̲/ is a pharyngealized dental fricative.

The two models mentioned above present different views as to whether the steady state should show coarticulatory effects or not. According to Keating, coarticulation is an interpolation between two targets; a claim which implies that coarticulatory effects should be manifested somewhere on the transitions, not on the steady states. Fowler describes coarticulation as extending on the whole unstressed vowel (1981). The data from Arabic used in this study will be analyzed and tested for significance in relation to these issues. In particular, the effects of V_1 and V_2 on the steady state of one another in V_1CV_2 utterances and the effects of pharyngealization will be emphasized. It is expected that pharyngealization, a feature which is a combination of raising the tongue dorsum towards the velum and retracting it towards the posterior pharyngeal wall, will lower the frequency of the second and third formant steady states for all adjacent vowels.

2. Experiment

2.1. Stimuli and Materials: A list of 72 utterances was constructed. A few of these utterances are actual words in Arabic and the majority are nonsensical sequences. The list contained the long vowels /i/, /a/ and /u/ in initial and final position of the sequences and the consonants /t/, /d/, /s/, /ð/ and their pharyngealized counterparts. The total of sequences were calculated in the form three preceding vowels X eight consonants X three following vowels = 72. This list includes all possible combinations of vowels and eight consonants included in the study.

2.2. Subjects: The subjects are two graduate students at The Ohio State University who are native speakers of Levantine Arabic. One of them will be referred to as LH, who is also the author of this paper, and the other as MA.

2.3. Procedure and Design: All VCV sequences were written in Arabic script on 3x5 inch cards. The cards were numbered and randomized in an attempt to eliminate any possible practice effects. Each of the 72 types was read five times by each subject with the following constraint on ordering: no token could be repeated until each of the other 71 tokens had been repeated at least once for any given cycle. The data were recorded in an anechoic chamber in the Linguistics Laboratory at The Ohio State University on a four track reel-to-reel recorder. The recording was then transferred to a cassette in a way that preserved the quality of the original recording. Subjects were instructed to read in a monotone and with phonemically long vowels. Wide-band spectrograms for all the tokens were then projected on the screen of a Voice ID RT 1000 machine used for spectrographic analysis in the Speech and Hearing Department at The Ohio State University. Zooming was carried out to enlarge spectrograms for the best possible projection for all the tokens. Measurements of frequency for the second formant were taken at four points in the V_1CV_2 sequences:

(1) The steady state of V_1
(2) The end of the transition of V_1 (into C)
(3) The beginning of the transition of V_2 (out of C)
(4) The steady state of V_2

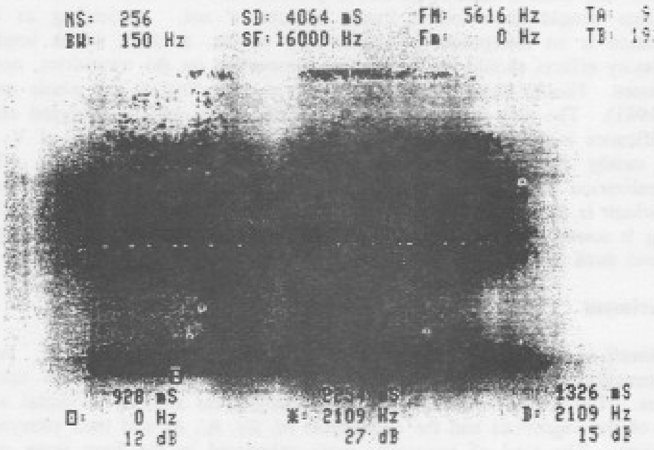


Fig. 3: Cursor placement at the center of the steady state of V_2 in the sequence /idi/ (speaker MA).

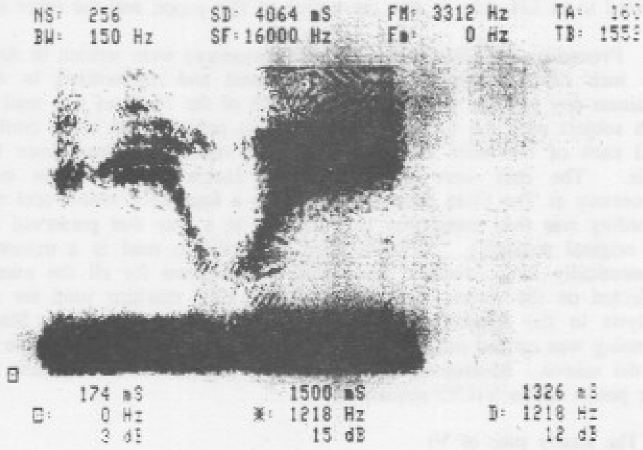


Fig. 4: Cursor placement at the beginning of the transition of V_2 in the sequence /iðˤi/ (speaker LH).

The measurements were taken by moving a cursor to the center of the formant and recording the frequency given automatically by the machine at the bottom of the screen. Figure 3 shows an example of placing the cursor at the steady state of V_2 and Figure 4 shows it at the beginning of the transition for V_2.

It is worth noting that the final decision of writing the utterances in isolation rather than having them in a frame sentence was based on three pilot studies. The outcomes of these studies have shown that other intervening factors in speech production such as adjacent phones, rate of speech, and stress placement resulting from using the frame sentence contribute uncontrollably to coarticulation. Therefore, the researcher became convinced that the best method of conducting the experiment is to have the utterances recorded in isolation as did Öhman.

3. Results

Figures 5 to 10 show mean and standard deviation values of steady state (top) and transition endpoint (bottom) for each vowel when the preceding or following vowel is /i/ versus /u/, divided by speaker (dashed lines for LH versus solid lines for MA), by intervening consonant feature (squares for pharyngealized versus circles for nonpharyngealized), and by consonant class (fricative to the left and stops to the right). The four graphs in each figure have been labelled A, B, C, and D to facilitate the process of reference in the explanation section.

The six figures have been grouped into two sets depending on whether V-to-V coarticulation is anticipatory or carryover. The first section shows the anticipatory effects and the second section shows the carryover effects. V-to-V coarticulation would be evident in the graphs as a significant difference between the connected /i/-context and the /u/-context means, with /u/-context means having lower frequencies than /i/-context means.

3.1. Anticipatory V-to-V Coarticulation

The first three figures show the effects of V_2 on V_1 in V_1CV_2 sequences. These effects are presented according to the preceding vowels in the order of /i/, /a/, and /u/.

3.1.1. Anticipatory coarticulatory effects on /i/

Figure 5 displays the anticipatory effects on /i/. None of the four graphs shows any significant V-to-V coarticulatory effects across pharyngealized consonants. Indeed, the top two graphs (A and B) show a tendency for the second formant to systematically have higher frequencies when the following vowel is /u/ than when it is /i/. The top two graphs also show a tendency for the second formant frequency to decrease across nonpharyngealized intervocalic fricatives for both speakers and across stops for MA when the following vowel is /u/. This decrease, however, is not statistically significant ($F = .80$, $P < .3778$ for fricatives) and ($F = .32$, $P < .5765$ for stops), so we cannot claim that there are coarticulatory effects on the steady state of /i/. The lower two graphs (C and D) show that pharyngealized consonants block coarticulation. Nonpharyngealized consonants, on the other hand, manifest two types of behavior that depend on the intervocalic consonant class. Stops block coarticulation while fricatives show significant effects ($F = 12.51$, $P < .01$). This conclusion conflicts with that of Öhman who reported that stops in Swedish and English permit V-to-V coarticulation and fricatives in both languages block it. One explanation for this contradiction is to assume that V-to-V coarticulation can be a language-particular phenomenon. It is not necessary to believe that what applies to one or two languages should apply to the rest of the languages in the world.

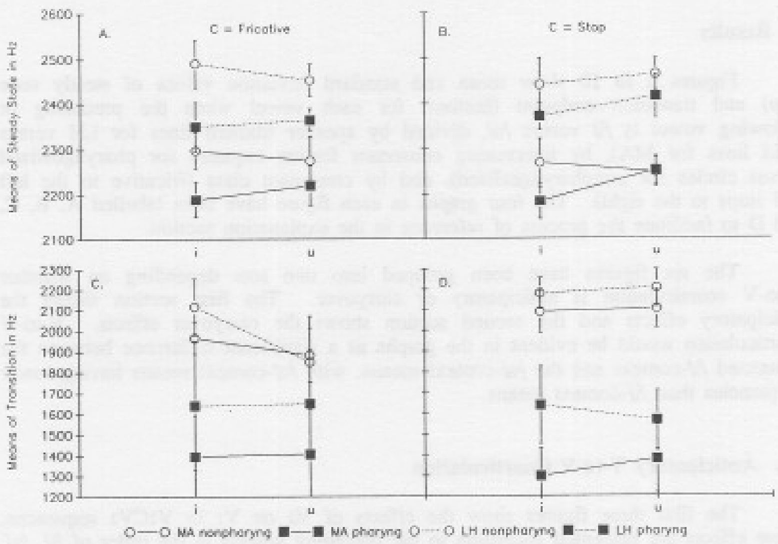


Fig. 5: Anticipatory effects of V_2 on /i/ as V_1.

3.1.2. Anticipatory coarticulatory effects on /a/

Figure 6 displays the anticipatory coarticulatory effects on /a/. The figure shows significant coarticulatory effects on the transitions of /a/ across the two speakers when the intervocalic consonant is nonpharyngealized ($F = 38.60$, $P < 0.001$ for fricatives and $F = 57.82$, $P < 0.001$ for stops). By contrast, there are no coarticulatory effects on the transition of /a/ across pharyngealized consonants (see

Graphs C and D). The feature pharyngealization plays a similar role in blocking coarticulation between the steady states. Pharyngealization, however, does not seem to be the only factor that determines whether coarticulation should or should not occur. Other factors, especially the speaker, become important for coarticulation in this case. The speaker MA, for example, shows significant coarticulatory effects across intervocalic fricatives and stops ($F = 40.69$, $P < 0.001$). LH, on the other hand, shows coarticulatory effects across intervocalic stops only.

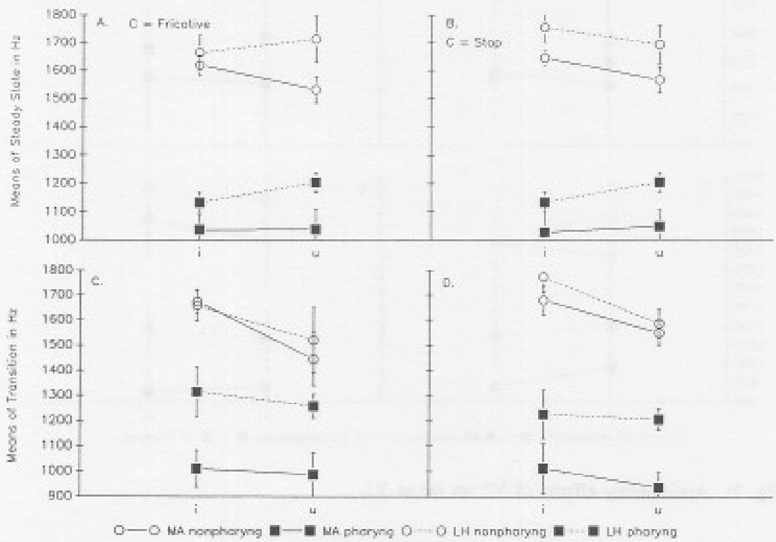


Fig. 6: Anticipatory effects of V_2 on /a/ as V_1.

3.1.3. Anticipatory coarticulatory effects on /u/

Figure 7 displays the anticipatory effects on /u/. No systematic coarticulatory patterning can be elicited from Figure 7; none of the four graphs included in the figure shows significant coarticulatory effects. This outcome supports neither Keating nor Fowler. It shows that V-to-V coarticulation does not always exist in VCV utterances as Fowler claims. It also shows that coarticulation cannot always be attributed to the existence or absence of the secondary articulation feature pharyngealization as Keating claims; rather it can be the intrinsic feature of the vowel being studied that determines whether coarticulation should occur or not. Some vowels have shown to be more resistant than others to V-to-V coarticulation. /a/ seems to be the vowel most tolerant of V-to-V coarticulation, /i/ is less tolerant, and /u/ is the least tolerant of coarticulation. This might be explained by the fact that /a/ is the only low vowel in Arabic. Therefore, it has more room for allophonic variation than either /i/ or /u/.

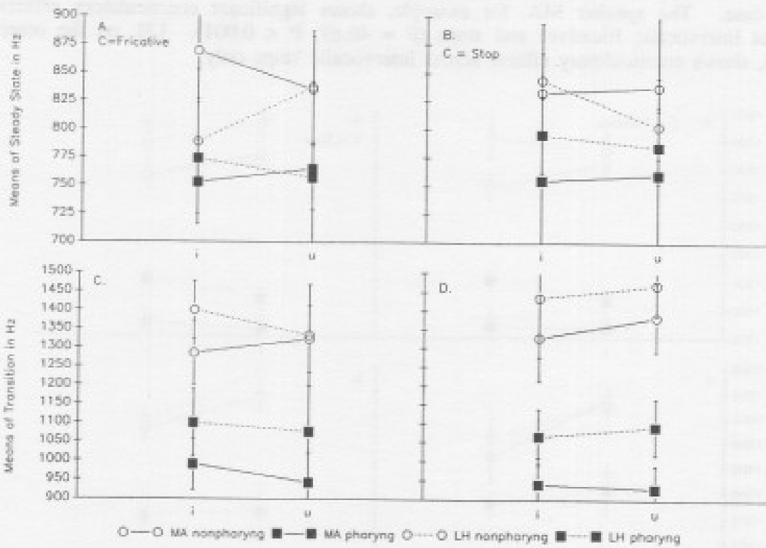


Fig. 7: Anticipatory effects of V_2 on /u/ as V_1.

3.2. Carryover V-to-V Coarticulation

Figures 8-10 display the effects of V_1 on V_2 in V_1CV_2 sequences. The effects are arranged according to the quality of the following vowels in the order of /i/, /a/, and /u/.

3.2.1. CARRYOVER COARTICULATORY EFFECTS ON /i/

Figure 8 displays the carryover coarticulatory effects on /i/. Figure 8 confirms Keating's prediction in two aspects: (1) only consonants with no vowel feature [+back], nonpharyngealized consonants, allow V-to-V coarticulation, and (2) the coarticulatory effects can be restricted to the transitions of vowels for some speakers. The top two graphs (A and B), which display the carryover coarticulatory effects on the steady state of /i/, show significant effects for the speaker LH ($F = 7.37$, $P < 0.001$), but no significant effects for the speaker MA. The lower two graphs (C and D), which display the carryover coarticulatory effects on the transition of /i/, show significant coarticulatory effects (F = 20.38, P <.001 for fricatives and F = 25.18, P= <.001 for stops) on /i/ when the intervocalic consonants are nonpharyngealized. By contrast, their pharyngealized counterparts do not show significant effects. This conclusion argues against Fowler who claims that intervocalic consonants, regardless of what features they may have, should not block coarticulation. It also shows that plain (nonpharyngealized) fricatives permit V-to-V coarticulation in Arabic, unlike their "counterparts" in English and Swedish, as it has been reported by Öhman.

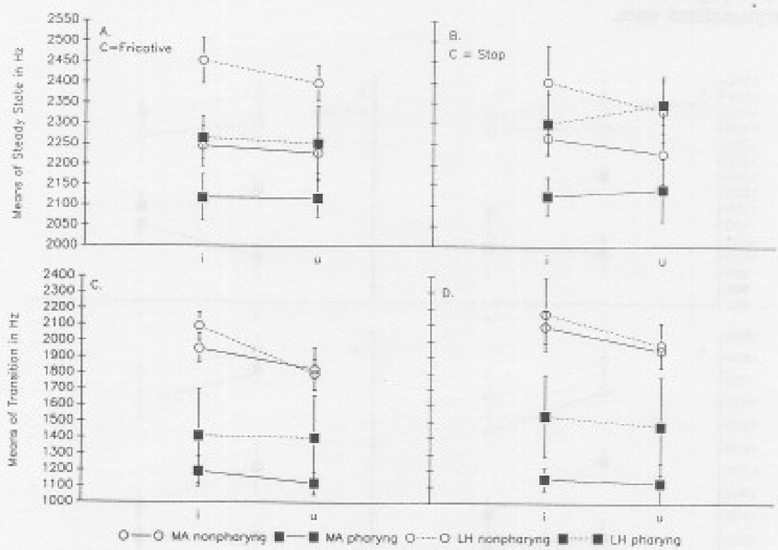


Fig. 8: Carryover effects of V_1 on /i/ as V_2.

3.2.2. CARRYOVER COARTICULATORY EFFECTS ON /a/

Figure 9 summarizes the carryover coarticulatory effects on /a/. Similar to Figure 6, coarticulatory effects occur systematically on the transition of /a/ so long as the intervocalic consonant is nonpharyngealized. Both speakers exhibit significant coarticulatory effects across nonpharyngealized intervocalic stops and fricatives at F = 21.74, P < 0.001 for stops and F = 15.82, P < 0.001 for fricatives (see Graphs C and D). In the meantime, and unlike any other figure that has been discussed so far, Figure 9 shows a case of V-to-V coarticulation across pharyngealized consonants. One speaker, LH, shows significant coarticulatory effects on both transitions (F = 63.60, P < 0.001) and steady state (F = 34.01, P < 0.001) of /a/.

The other speaker, MA, by contrast, does not show similar coarticulatory effects on either.

Coarticulatory effects on the steady state seem to be dependent on the speaker. LH, on the one hand, shows significant and systematic coarticulatory effects across all intervocalic consonants; namely, across nonpharyngealized stops, nonpharyngealized fricatives, pharyngealized stops, and pharyngealized fricatives. MA, on the other hand, shows a tendency for coarticulation to occur across nonpharyngealized consonants, but no coarticulation whatsoever across pharyngealized ones.

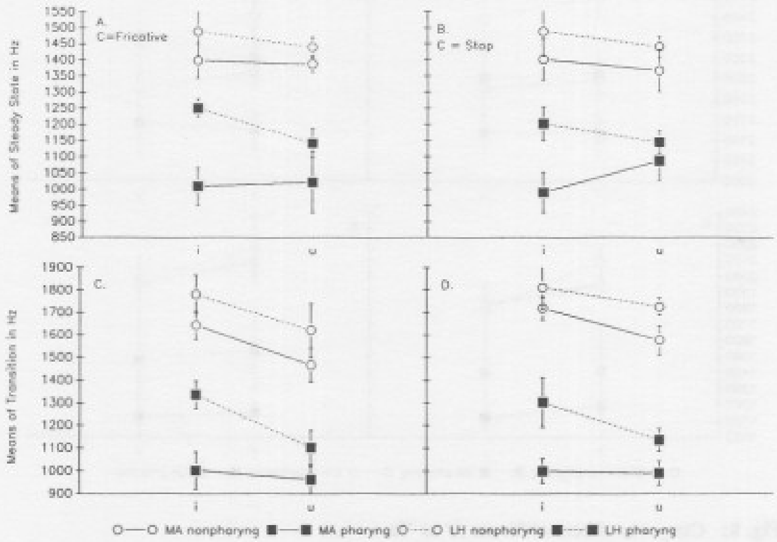


Fig. 9: Carryover effects of V_1 on /a/ as V_2.

Several conclusions can be drawn from this figure:

(1) Neither of the two models sketched above (Keating's and Fowler's), which were the main incentive to this study, adequately accounts for all cases of V-to-V coarticulation. Keating's model falls short of accounting for the coarticulatory effects across pharyngealized consonants as is the case with LH, and Fowler's model cannot account for the absence of coarticulation across pharyngealized consonants in the case of MA.

(2) The speaker has proven to be an important factor in V-to-V coarticulation. LH shows coarticulatory effects in all environments given in Figure 9. MA, by contrast, shows coarticulatory effects on the transitions of /a/ across nonpharyngealized consonants only.

(3) The direction of coarticulation (carryover versus anticipatory) seems to induce different coarticulatory patterns. The comparison between Figure 2, which summarizes anticipatory coarticulatory effects on /a/, and Figure 9, which summarizes carryover coarticulatory effects on /a/, shows that V-to-V coarticulation is likely to occur across pharyngealized consonants as a carryover case but not as an anticipatory one.

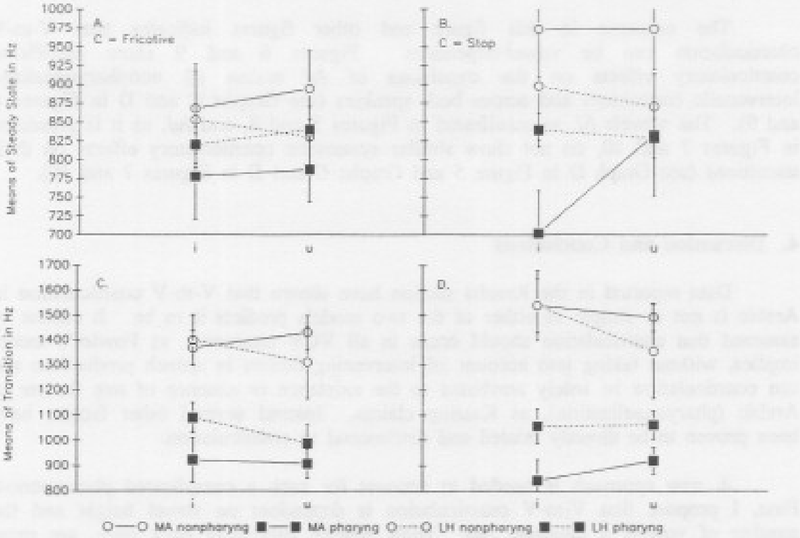


Fig. 10: Carryover effects of V_1 on /u/ as V_2.

3.2.3. CARRYOVER COARTICULATORY EFFECTS ON /u/

Figure 10 summarizes the carryover coarticulatory effects on /u/. Similar to Figure 7, Figure 10 does not show any systematic coarticulatory effects across the two speakers. It confirms the conclusion that /u/ does not tolerate V-to-V coarticulation and consequently acts as the vowel in Arabic most resistant to coarticulation. One speaker, LH, however, exhibits anomalously significant coarticulatory effects ($F = 8.59$, $P < 0.0057$) on the transition of /u/ in nonpharyngealized environments. The same speaker exhibits coarticulatory carryover effects on the transition and the steady state of the other two vowels in nonpharyngealized environments (see Figures 8, 9). He even sometimes exhibits carryover effects in pharyngealized environments as is the case with /a/ (see Figure 9). He does not, however, show similar systematic anticipatory coarticulatory effects. This outcome implies that there might be some association between some speakers or dialects (see **Discussion and Conclusions** for an explanation on dialectal differences) and the direction of coarticulation.

In general, the speakers in this study, especially LH, tend to show more carryover coarticulatory effects than anticipatory ones. Similar conclusions where greater carryover than anticipatory coarticulation occurs have been reported for English (Fowler 1981).

The outcome in this figure and other figures indicates that V-to-V coarticulation can be vowel-dependent. Figures 6 and 9 show significant coarticulatory effects on the transitions of /a/ across all nonpharyngealized intervocalic consonants and across both speakers (see Graphs C and D in Figures 6 and 9). The vowels /i/, as manifested in Figures 5 and 8, and /u/, as it is presented in Figures 7 and 10, do not show similar systematic coarticulatory effects on their transitions (see Graph D in Figure 5 and Graphs C and D in Figures 7 and 10).

4. Discussion and Conclusions

Data reported in the Results section have shown that V-to-V coarticulation in Arabic is not as simple as either of the two models predicts it to be. It cannot be assumed that coarticulation should occur in all VCV utterances, as Fowler's model implies, without taking into account all intervening factors in speech production; nor can coarticulation be solely attributed to the existence or absence of one feature in Arabic (pharyngealization), as Keating claims. Instead several other factors have been proven to be directly related and detrimental to coarticulation.

A new approach is needed to account for such a complicated phenomenon. First, I propose that V-to-V coarticulation is dependent on vowel height and the number of vowels a language has. High vowels, especially back ones, are more resistant to V-to-V coarticulation than low ones. Also the degree of V-to-V coarticulation increases as the number of vowels in a language decreases and vice versa. This assumption accounts for the fact that V-to-V coarticulation occurs more frequently across non-pharyngealized consonants than across pharyngealized ones. Pharyngealization, a secondary articulation vocalic feature, does not only retract the tongue root towards the posterior pharyngeal wall, but it also raises it towards the velum. Figure 11, adopted from Ali and Daniloff (1972), shows the tongue movement during the production of pharyngealized /ṭ/ and non-pharyngealized /t/. The figure presents a sample that is based on a cinefluorographic investigation of pharyngealized and non-pharyngealized consonants' articulation. With pharyngealized consonants being [+back] and [+high], it becomes more likely that V-to-V coarticulation will be blocked. Therefore, none of the figures 5 through 10 shows systematic V-to-V coarticulation across pharyngealized consonants.

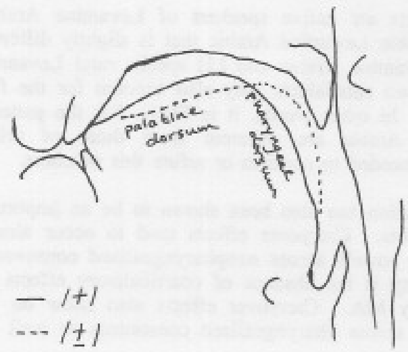


Fig. 11: A sample cine frame showing differences (in mm) in tongue position for contrasting pharyngealized /ṭ/ with non-pharyngealized /t/.

This assumption also accounts for the fact that the three vowels of Arabic have shown various degrees of tolerance towards V-to-V coarticulation. /a/, being non-high and the only low vowel in Arabic, creates the perfect environment for coarticulation. There is enough room for allophonic variation and there is no vowel height to resist coarticulation. Thus, Figures 6 and 9 show the highest degree of tolerance for V-to-V coarticulation. There are systematic coarticulatory effects on all transitions and most steady states across non-pharyngealized consonants. There are also carryover coarticulatory effects for LH across pharyngealized consonants as well. /u/, being high and back vowel, is the most resistant to coarticulation. No systematic coarticulatory effects across the two speakers can be inferred from Figures 7 and 10. /i/, being also a high front vowel, has more tolerance than /u/ and less than /a/.

Second, the speaker has been proven to be an essential factor in coarticulation. The two speakers participating in this experiment do not always show similar coarticulatory effects. In particular, speaker LH has shown coarticulatory effects in a wider range of environments than MA. This observation can be clearly seen in Figure 9 where LH coarticulates across pharyngealized and nonpharyngealized consonants while MA coarticulates across nonpharyngealized consonants only, in Figure 10 where LH shows significant coarticulatory effects on the transitions of /u/ while MA does not show similar effects, and in Figure 8 in which LH shows coarticulatory effects on the steady state of /i/ while MA does not. The only anomalous case exists in Figure 6 where the speaker MA shows coarticulatory effects on the steady state of /a/ and LH does not.

This difference between the two speakers can be attributed to the rate of speech production that has been developed by each of the speakers and the dialectal differences that exist between the two speakers. People who have developed a habit of speaking fairly fast are likely to coarticulate more than those who speak at a slower rate. Both speakers participating in this experiment have been noticed to have developed two different rates of speech production. The speaker LH has been noticed to speak faster than MA. This observation may account for the fact that he coarticulates in a wider range of environments than MA.

Meanwhile, though both subjects are native speakers of Levantine Arabic, each of them speaks a "subdialect" within Levantine Arabic that is slightly different from the other. MA speaks urban Levantine Arabic and LH speaks rural Levantine Arabic. The difference between the two subdialects may also account for the fact that LH coarticulates more than MA. In other words, it is likely that the patterns of coarticulation in rural Levantine Arabic are different from those of urban Levantine Arabic. Further studies are needed to confirm or refute this outcome.

Third, the direction of coarticulation has also been shown to be an important factor in V-to-V coarticulation in Arabic. Carryover effects tend to occur almost systematically on the transitions of all vowels across nonpharyngealized consonants. The only exception to this systematicity is the absence of coarticulatory effects on the transitions of /u/ as manifested by MA. Carryover effects also show on the transitions and the steady state of /a/ across pharyngealized consonants as well for the speaker LH.

Anticipatory effects, by contrast, are systematic across nonpharyngealized consonants on the transitions of /a/ only. No significant effects are manifested on the transitions of /u/ for either speaker in the same environment nor are there effects on the transitions of /i/ across nonpharyngealized stops. No significant effects are shown in any pharyngealized environment.

The reasons for the systematicity in the carryover direction versus the sporadicity in the anticipatory direction can be attributed to the stress patterns in two-long-vowel utterances and the phonological system of Arabic. Arabic places the primary stress on the first long vowel in these utterances and shortens the long vowel word finally. Thus, in actual production, V_1 in V_1CV_2 utterances is generally longer and more emphasized than V_2. Therefore, the likelihood for V_2 to coarticulate due to lack of emphasis and shortening is greater than that of V_1. Indeed, most non-native learners of Arabic perceive long vowels in word final position as /ə/. For example, the word /læ:/ meaning "no" is usually perceived as /lə/ (personal observation).

Like most studies in speech production, the results of this experiment are based on data obtained from a relatively small subject population. The findings are therefore far from conclusive. They, however, can serve as good starting points for examining a number of questions concerning V-to-V coarticulation in Arabic. Among these questions are:

(1) Do Arabic speakers show more systematic carryover coarticulation than anticipatory ones, as the present study reveals?

(2) Could there be some kind of priority scale which would classify some factors as more important than others in their effects on coarticulation? For example, the study reveals to me that vowel height is the most important factor in determining whether coarticulation in VCV utterances should occur or not; there are always coarticulatory effects on the transitions of /a/, less so on /i/ and /u/. The feature pharyngealization comes in second place; its presence blocks coarticulation in almost all cases, but its absence does not imply that coarticulation should occur. The speaker and possibly dialectal differences come in third place. Finally comes the intervocalic consonant class, though it has played a minimal role in this study, especially if we compare it to the role it played in Öhman's study (1966).

Acknowledgement

I wish to thank Mary Beckman for her assistance at all stages in writing this paper. Mary provided invaluable help in the fields of phonetics, computer use, and statistics. I would like also to express my thanks to Robert Fox who allowed me to use the RT 1000 spectrographic machine in the Speech and Hearing Department and Diane Hambrick who helped in editing the final draft of this paper.

Bibliography

Ali, Hassan Latif and R. G. Daniloff. (1972). "A contrastive cinefluorographic investigation of the articulation of emphatic-nonemphatic cognate consonants", *Studia Linguistica* **25**: 81-105.

Bonnot, J. F. (1977). "Recherche experimentale sur la nature des consonnes emphatiques de l'Arabe classique", *Travaux de l'Institut de Phonetique de Strasbourg* **9**: 47-88.

Card, Elizabeth Anne. (1983). *A phonetic and phonological study of Arabic emphasis.* Ph.D. dissertation, Cornell Univ. University Microfilms International.

Fowler, Carol A. (1981). "A relationship between coarticulation and compensatory shortening", *Phonetica* **38**: 35-50.

________. (1983). "Converging sources of evidence on spoken and perceived rhythms of speech: cyclic production of vowels in monosyllabic stress feet." *Journal of Experimental Psychology*, General. **112**: 386-412.

Gay, T. (1977). "Articulatory movement in VCV sequences", *JASA (Journal of the Acoustical Society of America)* **62**: 183-193.

Ghazeli, S. (1977). *Back consonants and back coarticulation in Arabic*, Ph. D. dissertation, The University of Texas at Austin.

Jakobson, R. (1962). "Mufaxxama, the 'emphatic' phonemes in Arabic", in *Selected Readings, Roman Jakobson* (Mouton and Co., The Hague), pp. 510-522.

Keating, Patricia A. (1985). "CV phonology, experimental phonetics, and coarticulation", *UCLA Working Papers in Phonetics* **62**: 1-13.

Manuel, S. Y. and R. A. Krakow. (1984). "Universal and language particular aspects of V-to-V coarticulation", Haskins Laboratories: *Status Reports on Speech Research* **SR-77/78**, 69-78.

Obrecht, D. H. (1961). *Effects of the second formant on the perception of velarization consonants in Lebanese Arabic.* Ph.D. dissertation, Univ. of Pennsylvania. On microfilm (from Penn.).

Öhman, S. E. G. (1966). "Coarticulation in VCV utterances: spectrographic measurement", *JASA* **39**: 151-168.

Purcell, Edward T. (1979). "Formant frequency patterns in Russian VCV utterances", *JASA* **66**: 1691-1702.

Recasens, Daniel. (1985). "Vowel-to-vowel coarticulation in Catalan VCV sequences", Haskins Laboratories: *Status Reports on Speech Research* **SR-81**: 181-204.

________. (1986). "An acoustic analysis of V-to-C and V-to-V: coarticulatory effects in Catalan and Spanish VCV sequences", Haskins Laboratories: *Status Report on Speech Research* **SR-86/87**: 71-86.

OHIO STATE UNIVERSITY WORKING PAPERS IN LINGUISTICS **38** (105-120)

Korean Lenis and Fortis Stops: Synthesis and Categorical Speech Perception Task

Sook-hyang Lee

Ohio State University

Abstract: This paper reports on an attempt to synthesize two different Korean syllables based on the previous articulatory and acoustic studies about Korean stops: /t'a/ which consists of an alveolar fortis stop and /ta/ which consists of an alveolar lenis stop. Using these two synthesized syllables as endpoints of the continuum, five more syllables between two endpoints were created for identification and discrimination tasks. Korean native speakers did not perceive the stimuli categorically; their identification boundary was not very sharp and the discrimination peak was not very high, either. Two possible reasons were proposed for this unexpected result.

1. Introduction

The stop consonants in Korean have a three-way distinction in place of articulation. For the same place of articulation there is a three-way distinction in manner of articulation in initial position. Linguists have used the following terms to describe the manner distinction. Category 1 is characterized as voiceless, tense, long, strong, forced and/or glottalized; Category 2 is voiceless, lax, slightly aspirated, and/or weak; Category 3 is voiceless, heavily aspirated. An example of this three-way distinction is shown in (1).

(1)	1		2		3	
	t'al	'daughter'	tal	'moon'	t^hal	'mask'
	p'aŋ	'bread'	paŋ	'room'	p^haŋ	'bang'

(Henceforth, I will call Category 1 a fortis stop, Category 2 a lenis stop, and Category 3 an aspirated stop.)

Abramson and Lisker (1964) claimed that they could be classified into three groups simply in terms of VOT differences. However, others have noted that fortis and lenis stops have overlapping VOT values.[1] Kim (1965) found this overlap made by the same speaker on the same day. A perceptual experiment done by Han and Weitzman (1970), where subjects were asked to identify stimuli which were

[1] Abramson and Lisker (1964: 403) also observed the overlap in VOT, but attached little importance to it, saying merely that the distribution of values is "somewhat anomalous."

made by cutting off portions of the aspirated part of an aspirated stop, showed that at least 75% of the responses were for a lenis stop even when the VOT was reduced to 1 ms or less, which is within the range of VOT for fortis stops. The result of Abramson and Lisker's (1971) perception test also indicated that there must be another dimension that works with VOT in distinguishing the categories.

Since the VOT difference alone is not enough to distinguish these two stops, there must be other characteristics which differentiate them significantly from each other. The present study consists of three parts. First, I will review the existing experimental literature on the production of the distinction. Second, I will report on an attempt to synthesize these two sounds using synthesizer parameter settings based on the articulatory and acoustic findings made in the previous studies. By synthesizing the stops we should be able to see what kinds of differences contribute most to distinguish them from each other. Third, I produced a lenis-fortis continuum along the dimensions of contrast used in the synthesis of the two types and used it in categorical perception tasks. I will report on how Korean native speakers perceive that continuum.

2. Summary of existing production studies

Previous production studies can be classified into four approaches: acoustic records, fiberscopic and EGG observation of laryngeal vibration and glottal area, EMG records of muscle activity, and measurements of air pressure and air flow. Results generally agree with an interpretation of the fortis stop as having ejective characteristics and the lenis stop as having breathy voiced characteristics. For example, Kim (1965) found that right after the release of fortis stops, the acoustic intensity was greater and the fundamental frequency was higher than after corresponding lenis stops. Han and Weitzman (1970) similarly found that at the voice onset following fortis stops, the amplitude rise time is shorter and intensity is greater than that for the corresponding lenis stops. Higher fundamental frequency at the voice onset of the vowel following the fortis stops was also reported in studies such as Han and Weitzman (1967, 1970), Hardcastle (1973) and Kagaya (1974). These results are suggestive of high f0 after fortis stops and low f0 around breathy consonants in other languages as cited in Hombert (1979). Kagaya (1974) observed longer duration of formant of transition for fortis stops than for the corresponding lenis stops.

In addition to these acoustic characteristics of Korean lenis and fortis stops, Kim (1965) observed weaker intensity of aspiration noise for the lenis stops than for the corresponding aspirated stops. He and others (e.g. Han and Weitzman 1970) also observed weaker higher formants (f3 and/or f4) which is believed to be indicative of breathy voice. Kim noted that Korean lenis stops were heard with a lot of breathiness by Westerners. This observation is supported by some spectrographic studies such as Han and Weitzman (1967). They found a mixture of voicing and noise at the onset of the vowel following the lenis stop. This can be seen in the spectrogram of my voice given in Fig. 1. Noise appears at the beginning of the vowel /a/ following the lenis stop /p/.

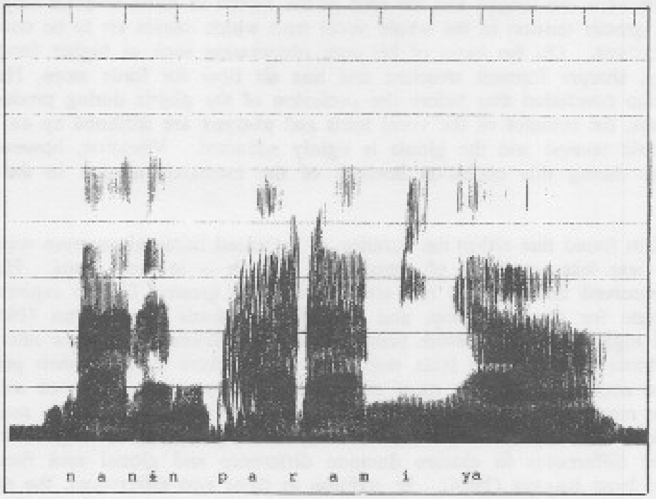


Fig. 1: A spectrogram of the Korean sentence /nanɨn paramiya/.

According to Kagaya's (1971, 1974) fiberscopic studies and Dart's (1987) summary, the peak value of glottal width during articulatory closure was largest for the aspirated stop, intermediate for the lenis stop, and smallest for the fortis stop. The timing of the narrowing gestures relative to articulatory release is also different in the various kinds of stops. For the aspirated stop, release generally occurred near the moment of maximal glottal opening. With lenis stops, the glottis began to close gradually through the occlusion, although the glottis was still open at release.[2] During the fortis occlusion, on the other hand, the vocal folds were in complete contact from 80 to 100 ms before release. Abberton's (1972) laryngographic study found that waveforms of the fortis stop showed the long closed phase and slow opening characteristics of creaky voice.

EMG studies of the intrinsic laryngeal muscles such as Hirose et al. (1974) showed a marked increase in lateral cricoarytenoid and vocalis muscle activity just

---

[2] In addition, a complete contact of the vocal folds is observed all through the utterance with the vocal folds vibrating for the lenis stop in the intervocalic position.

before release in the fortis stop. This means that the glottis is tightly closed before release by a tensing of these adductor muscles. Kim's (1965) EMG and palatograph study reported greater muscle activity at the lips during fortis bilabials and more contact between the tongue and the roof of the mouth in fortis alveolar stops. This indicates greater tension in the whole vocal tract which causes air to be compressed as in ejectives. On the basis of his own observation such as higher fundamental frequency, sharper formant structure and less air flow for fortis stops, Hardcastle (1973) also concluded that before the occlusion of the glottis during production of fortis stops, the muscles of the vocal folds and pharynx are stiffened by an increase in isometric tension and the glottis is tightly adducted. Vibration, however, does not occur during this occlusion because of the increased tension in the vocalis muscle.

Kim found that either the duration of increased intraoral pressure was shorter or there was less amplitude of pressure -- or both -- in lenis stops. Hardcastle (1973) observed that air flow rate after release was greatest for the aspirated stop, intermediate for the lenis stop, and least for the fortis stop. Dart (1987) also observed higher oral air pressure before release and lower oral airflow after release for the fortis stop than the lenis stop, although speakers differ in their production strategies: some seem to put more emphasis on air pressure difference and others put more emphasis on oral air flow difference. Dart explored possible reasons for this difference using Keating's (1984) computer implemented aerodynamic model, given the differences in closure duration difference and glottal area function as estimated from Kagaya (1974). In addition to these two differences, the results of the modeling experiment led her to infer that fortis stops are produced with greater vocal tract wall tension than in lenis stops. Furthermore, this simulation experiment suggested the factors that could be involved in the speaker-specific production strategies observed. Speaker who depend more on the oral flow difference may show larynx lowering or other supraglottal cavity expansion just before release of the fortis stop, while speakers who depend more on the air pressure difference may show a more rapid increase in respiratory muscle force during the closure and a smaller VOT difference between the stops.

To summarize the findings thus far, we can assume that the articulation of fortis stops differs from that of lenis stops in the following ways: first, greater tension in the laryngeal muscles (especially vocalis and LCA, which are adductor muscles) is involved in the production of fortis stops, causing a very small opening followed by a complete tight contact. Lower oral airflow right after release of the fortis stops could be a consequence of a complete contact of vocal folds at the points of release, while higher oral airflow right after the release of the lenis stops could be a result of a slight opening of the glottis at the point of release which might be one of the reasons for the observed breathiness at the onset of vocal fold vibration following lenis stops. Furthermore, the greater tension in the glottis adduction explains not only why creaky voice at the onset of voicing following fortis stops is observed but also why the fortis stop is not voiced intervocalically while the lenis stop is voiced.

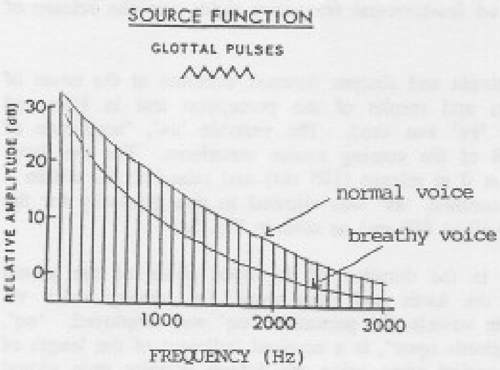


Fig. 2: The spectrum of the glottal source for a breathy voice and a normal voice (adapted from Borden and Harris 1984).

Second, as Hardcastle (1973) pointed out, fortis stops may involve an increase in isometric tension in the vocalis muscle to stiffen the vocal folds, which I believe causes them to be thinner, and thus to vibrate at a faster rate when they are closer together for voicing. Third, the tensing during the fortis stop is perhaps not confined to the glottis but is present in the rest of the vocal tract, offering an acoustically less damped formant structure. Fourth, as Kagaya (1974) suggested, the waveform of the volume velocity at the glottis for lenis stops could be characterized as a steeply falling slope from the fact that for all three types of stops, f3 and f4 are almost steady and each formant frequency is almost the same at the voice onset; but these formants of the vowel segment following the lenis stops were weaker. In other words, the amplitude of the higher harmonics for a breathy voice in the glottis would be lower relative to that for normal voice as can be seen in Fig. 2. This hypothetical steeply falling slope of the glottis waveform for Korean lenis stops acccompanied by breathiness is compatible with Fant's (1983) study about voice source by inverse filtering.

In short, the fortis vs. lenis contrast in Korean can be characterized not only by the presence vs. absence of tension in the vocal folds and the whole vocal tract but also by steepness vs. gradualness of the slope of the glottis waveform.

3. Synthesis

This study used the KLPC synthesizer program which is a version of the Klatt software synthesizer (Klatt 1980) rewritten to run on an IBM compatible PC by K. Johnson and Y. Qi. A list of the constant and variable parameter time functions that control the synthesizer is shown in the appendix.

I started with the English syllable /ta/. The whole length of the syllable is 700 ms. The value of the parameters, 'av', 'oq', and 'f0' were varied to reflect the findings of the previous studies such as differences with respect to degree of prominence in formant structure, the duration of open phase of the glottal pulse for the vowel following the stops, and fundamental frequency right after the release of the stops, respectively.

To reflect the more prominent and sharper formant structure at the onset of the vowel following fortis stops and results of the perception test in Han and Weitzman (1970), the parameter 'av' was used. The variable 'av', "amplitude of voicing", is the amplitude in dB of the voicing source waveform. For the fortis stop, the value of 'av' was set at 0 at release (195 ms) and raised to 60 within 5 ms, i.e., a very sharp rise. In contrast, 'av' was allowed to rise gradually for the lenis stops (from 0 at 195 ms to 60 at 260 ms) as seen in Fig. 3a.

To reflect the difference in the duration of the open phase of the glottal pulse for the vowel following the fortis and lenis stops, i.e., creaky voice vs. breathy voice, at the onset of the vowels, the parameter 'oq' was employed. 'oq', "percent of voicing period with glottis open", is a nominal indicator of the length of the glottal pulse relative to the period when using the default impulse train glottal source. Generally, three phases are evident in the cycle of vibration in normal voice: a rapid closing, a slower opening and an open phase. Breathy voice is characterized by a long duration of the open phase and a slow closing phase; creaky voice by a long closed phase and slow opening as can be seen in Fig. 4. For the creakiness in the fortis stop, the value of 'oq' was set at 10 from 0 ms to 195 ms, i.e., during closure for the fortis stop, and it goes to 50 during the first 50 ms of the following vowel. This means that the closure phase of one glottal pulse is much longer relative to an open phase. For the breathiness in the lenis stop, the value of 'oq' was set at 70 during the stop closure and it goes down to 50 during the first 50 ms of the following vowel as in Fig. 3b.

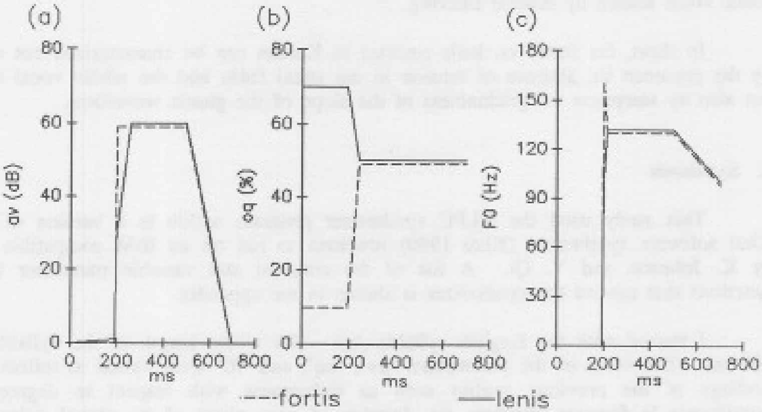


Fig. 3: The values of the parameters for Korean fortis and lenis stop syllables.

To reflect the higher fundamental frequency right after the release of the fortis stop and the lower one for the lenis stop, the parameter 'f0', "fundamental frequency", was varied as seen in Fig. 3c.

With the parameters used so far, I was able to synthesize an acceptable fortis stop. However, more parameters had to be changed for the lenis stop. The parameter 'tl' was used to reflect Kagaya's (1974) observation mentioned in section 2 that the higher formants, f3 and f4, are weaker at the onset of the vowel following the lenis stop, which was assumed to be due to the slope of the glottal waveform. The variable 'tl', "spectral tilt of voicing", is the additional downward tilt of the spectrum of the voicing source, in dB, as realized by a soft one-pole low-pass filter. This parameter is an attempt to simulate the spectral effect of a "rounding of the corner" at the time of closure in the glottal volume velocity waveform of breathiness, which is due either to an incomplete closure, or an asynchronous closure such that the anterior portion of the vocal folds meet at the midline before the posterior portions come together (Klatt 1980). The value of 'tl' was set at 0, the default value, for the fortis stop during the closure of the fortis stop but at 12 for the lenis stop. After the lenis closure, it went down to 0 during the first 50 ms of the following vowel as in Fig. 5a.

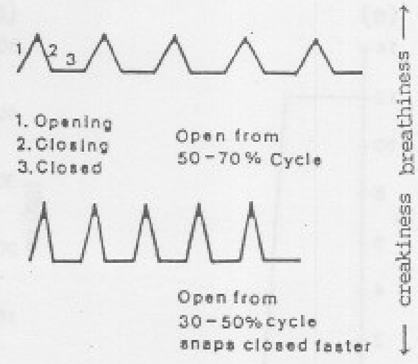


Fig. 4: Movement of the vocal folds during voicing for breathiness and creakiness (adapted from Borden and Harris 1984).

When changing the parameter 'tl', still I did not produce a good lenis stop. I tried turning on the 'ah' parameter to reflect some degree of intensity of the aspiration of the breathy release of the lenis stop. The variable 'ah' "amplitude of aspiration" is the amplitude in dB that is combined with periodic voicing, if present ('av'>0), to constitute the glottal sound source that is sent to the cascade vocal tract. I varied the value of 'ah' as in Fig. 5b.

In synthesizing these syllables, I only concentrated on getting those two sounds, not on determining what is the most important or primary cue. One thing that was observed by accident was that the fortis stop could not be synthesized with low f0 value with other parameters remaining unchanged. It would be interesting to see if there is a hierarchy among parameters used in this study in terms of the participation, or if they work equally in cueing those two types of stops.

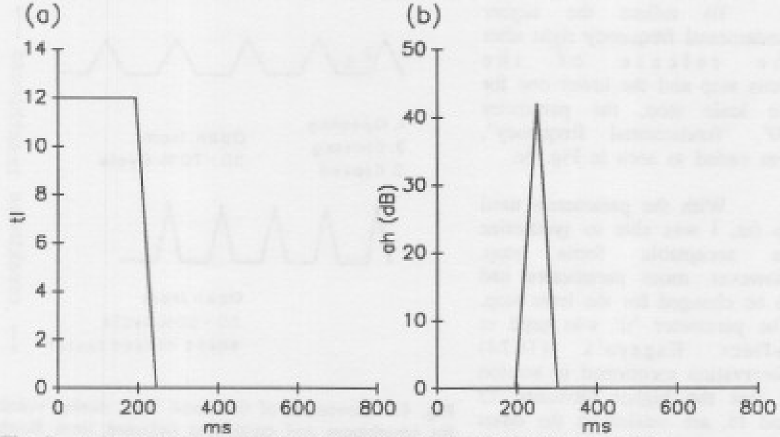


Fig. 5: Values of the parameters of 'tl' and 'ah' for the Korean lenis stop.

4. Experiment (Speech perception tasks)

To test the synthesized fortis and lenis stops, I did a preliminary perception study.

4.1 Methods

4.1.1 MATERIALS

A Korean lenis-fortis continuum was employed as the experimental stimuli. I used the two syllables which I had synthesized earlier as endpoints of the continuum. The basic pattern for each stimulus item consisted of 4 steady-state formants for the vowel [a] to which were added the appropriate release burst and formant transitions to produce alveolar fortis and lenis stops. The continuum consisted of 7 stimulus syllables which were varied in equal steps with respect to the same parameters used earlier, 'av', 'oq', 'f0', 'tl', and 'ah'.

4.1.2. SUBJECTS

Four Korean native speakers served as subjects for this experiment. They were graduate students with ages ranging from 24 to 30. Of these subjects, 3 were female and 1 was male, and all had normal hearing as reported subjectively. Dialect variance between subjects was not considered because I thought that it would not affect perception of Korean stops at all.

4.1.3 PROCEDURE

Seven syllables were randomized by K. Johnson's (1985) ERS program (Experiment Running System) for identification tasks and a 2 step AX discrimination task. For the identification task, each token was represented 20 times and for the discrimination task, each pair of tokens were represented 10 times, with 15 practice items for each task.

Subjects were tested individually in the recording room at the OSU Linguistics Lab. They were instructed to press one button out of two to respond to the stimuli depending on whether they heard fortis or lenis. The ERS program tabulated responses of each subject.

4.2 Results

When we look at the average results given in Fig. 6, it seems that Korean native speakers do not perceive the fortis-lenis continuum categorically. The identification boundary between token 4 and token 5 is not very sharp and the discrimination peak between token 4 and token 6 is not high, either.

The individual subject's responses given in Fig. 7 agree with this average result. No subject identified the lenis and fortis in a strongly categorical manner, and especially subjects KB and KC identify even the endpoint stimulus with about 15 or 20% responses for the other category. Furthermore, the peak in the

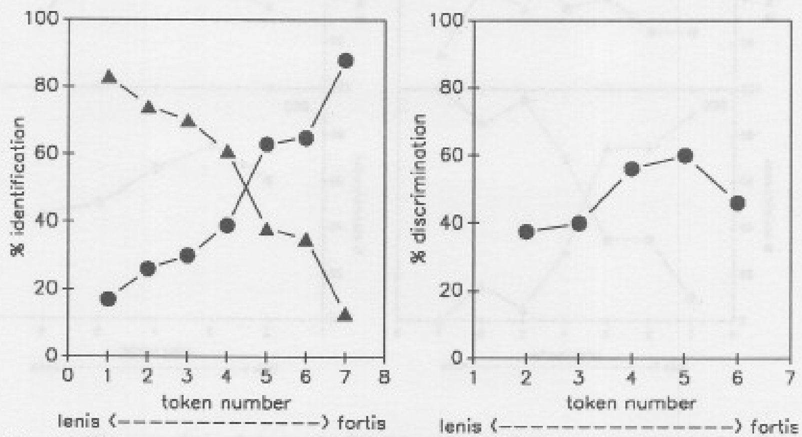


Fig. 6: Average response functions for the Korean native speakers on the identification and discrimination tasks.

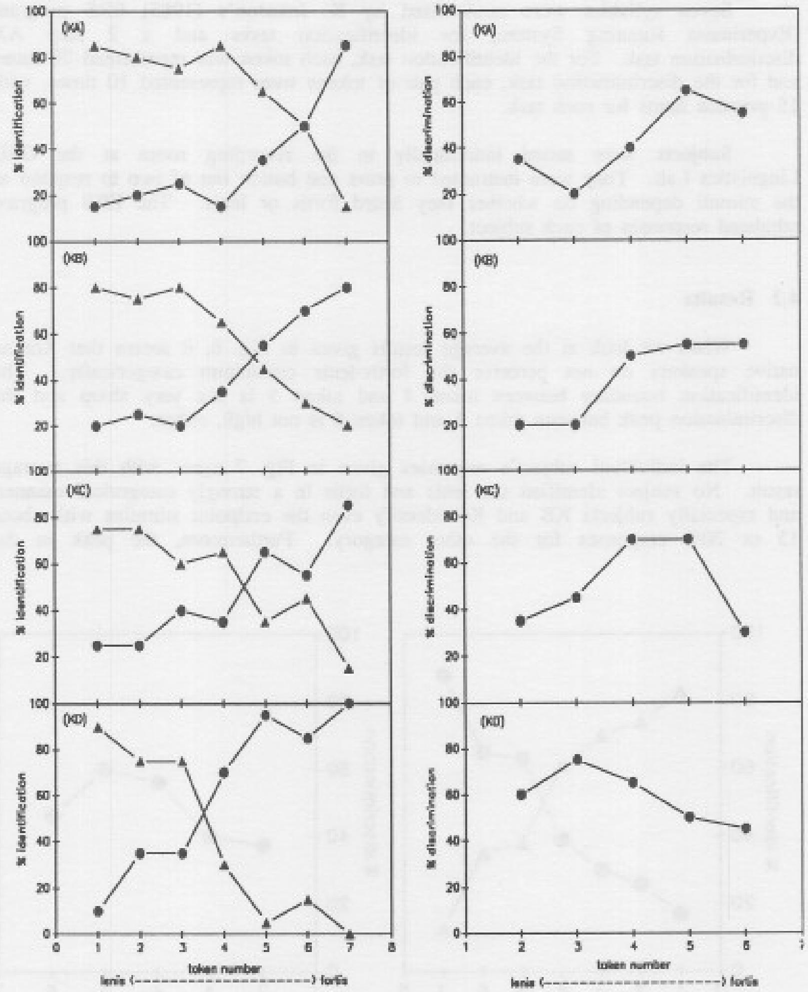


Fig. 7: Individual response functions on identification and discrimination tasks.

discrimination task is not so high, and especially subjects KA and KB have discrimination peaks just around 50 or 60. It is interesting to note that Korean subjects did not have identification boundaries at the same point along the continuum: subject KA has it around token 6; and subjects KB and KC, between token 4 and token 5; subject KD, between token 3 and token 4.

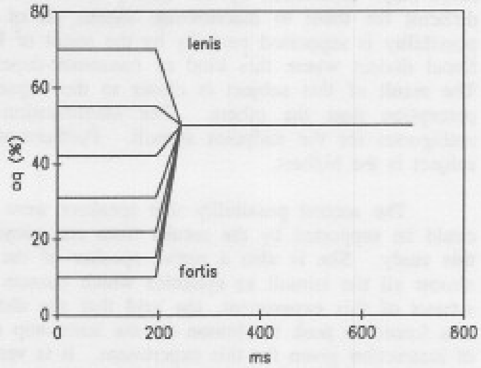


Fig. 8: Value of the parameter 'oq' in unequal steps.

It is also interesting that subjects KA, KB, and KC's discrimination peaks are expanded on the last two or three pairs, i.e., on pairs 3-5, 4-6, and 5-7. This observation could be interpreted as follows. The same articulatory or acoustic distance could be perceived as different distances depending on where in the continuum they are. For the lenis stop part along the continuum, the perception distance may be much closer than for the fortis stop part. In other words, in order for perception distance to be the same between each token, some parameters could be varied in unequal steps. For example, the parameter 'oq' should be varied in a larger step size around the lenis part and in a smaller step size around the fortis stop part along the continuum as in Fig. 8.

4.3 Discussion

We could think of some possible reasons for these unexpected results. One possible reason is that the parameters employed here in producing the lenis-fortis continuum do not work equally. Some parameters should be varied in smaller step sizes than other parameters or some in larger step size than others. To test this possibility we need to synthesize stimuli with each parameter changing separately and see what Korean speakers perceive.

Another possible reason is that most subjects were influenced by their native dialects. Three out of four subjects who participated in this experiment are native speakers of the Chonnam dialect where the pitch of the vowel is determined by the type of the preceding consonant in accentual phrase initial position: when the fortis or aspirated stop precedes a vowel, it is realized with a low or rising tone depending on its length (Jun 1990, this volume). Note that stimuli employed in this study were long (700 ms) and fundamental frequency was high (132 Hz) considering the fact that formant values were appropriate for a male speaker. So, the native speakers of this dialect, KA, KB, and KC, might identify the lenis and fortis stops sometimes by the quality of the following vowels. It might be very difficult for them to discriminate tokens all of which are long and high. This possibility is supported partially by the result of KD who is a native speaker of the Seoul dialect where this kind of consonant-dependent vowel pitch does not exist. The result of this subject is closer to the expected patterns of categorical speech perception than the others. The identification function is sharper and is less ambiguous for the endpoint stimuli. Furthermore, the discrimination peak of this subject is the highest.

The second possibility that speakers were influenced by their native dialects could be supported by the results from one subject which I have not mentioned in this study. She is also a native speaker of the Chonnam dialect. She perceived almost all the stimuli as syllables which contain a fortis stop. After serving as a subject of this experiment, she said that she did not hear any lenis stops but she was forced to push the button for the lenis stop sometimes because of the influence of instruction given for this experiment. It is very likely that she was concentrating on the vowel quality, not on the consonant quality, when she was asked to identify the type of consonants. To test this possibility, we need to get more subjects while controlling for dialect and put the syllables in non-initial position.

5. Conclusion

The first part of this paper describes an attempt to synthesize the Korean alveolar fortis and lenis stops and to figure out what kinds of articulatory and acoustic properties must be involved based on previous studies about those two types of Korean stops. To produce an acceptable fortis stop, the parameter 'av' (amplitude of voicing), 'oq' (percent of voicing of period with glottis open), and 'f0' (fundamental frequency) were employed which reflect respectively differences in prominence of formant structure, duration of opening phase of glottis and the fundamental frequency at the onset of the vowel following fortis and lenis stops as observed in earlier studies. To get an acceptable lenis stop, two more parameters, 'tl' (spectral tilt of voicing) and 'ah' (amplitude of aspiration) were used to reflect the observed weaker higher formants (f3 and/or f4) and some amount of intensity of the aspiration noise with voicing at the onset of the vowel following lenis stops.

The second part of this paper concentrated on the categorical perception experiments: the identification task and discrimination task using seven synthesized tokens along the lenis-fortis continuum. The result could be summarized as follows: Korean native speakers did not perceive the stimuli categorically. Their identification boundary was not very sharp and the discrimination peak was not very high, either.

To interpret the unexpected results of this study, two possible reasons were proposed: one is that all the parameters employed here may not work equally in cueing Korean lenis and fortis stops. Another is that most subjects in this experiment were influenced by their native dialects where the pitch of the phrase-initial vowel is determined by the type of the preceding consonant. In later work, I will test these two possibilities.

Acknowledgements

I am deeply grateful to Mary Beckman, David Odden, and Wayne Cowart for their advice and comments throughout the investigation. I am also grateful to Gina Lee, Keith Johnson and Ken deJong for giving me much of their time in discussing these issues and encouraging me.

References

Abberton, E. 1972. "Some Laryngographic Data for Korean Stops". *J. Int. Phon. Assoc.* **2**: 67-78.

Abramson, A. S. and L. Lisker. 1964. "A Cross Language Study of Voicing in Initial Stops: Acoustical Measurement". *Word* **30**: 384-422.

_________. 1971. "Voice Timing in Korean Stops". *Status Report on Speech Research* (Haskins Laboratories, New Haven) **27**: 179-184.

Borden, G. L. and K. S. Harris. 1984. *Speech Science Primer*, 2nd ed. Williams & Wilkins, Baltimore.

Dart, S. N. 1987. "An Aerodynamic Study of Korean Stop Consonants: Measurements and Modeling". *J. Acoust. Soc. Am.* **81(1)**: 138-147.

Fant, G. 1983. "Preliminaries to Analysis of the Human Voice Source". *Speech Transmission Laboratory Quarterly Progress and Status Report* 4/1982: 1-27.

Han, M. S. and R. S. Weitzman. 1967. "Acoustic Features in the Manner-differentiation of Korean Stop Consonants." *Studies in the Phonology of Asian Languages* **V**, Univ. of Southern California, Acoustics Phonetics Research Laboratory, Los Angeles.

_________. 1970. "Acoustic Features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/". *Phonetica* **22**: 112-128.

Hardcastle, W. J. 1973. "Some Observations on the Tense-Lax Distinction in Initial Stops in Korean". *J. Phon.* **1**: 263-272.

Hirose, H., C. Y. Lee and T. Ushijima. 1974. "Laryngeal Control in Korean Stop Production". *J. Phon.* **2**: 145-152.

Hirose, H., H. S. Park and M. Sawashima. 1983. "Activity of the Thyroarytenoid Muscle in the Production of Korean Stops and Affricates". *Annual Bulletin* **17**, Research Institute of Logopedics and Phoniatrics, University of Tokyo.

Hombert, J. M. 1979. "Consonant Types, Vowel Quality, and Tone". *Tone: A Linguistic Survey*, V. Fromkin, ed. Academic Press.

Johnson, K. A. 1985. *A System of Running Speech Perception Experiments*. MA thesis, Ohio State University.

Jun, S. A. 1990. "The Accentual Pattern and Prosody of the Chonnam Dialect of Korean". This volume.

Kagaya, R. 1971. "Laryngeal Gestures in Korean Stop Consonants". *Annual Bulletin* **5**, Research Institute of Logopedics and Phoniatrics, University of Tokyo.

________. 1974. "Fiberscopic and Acoustic Study of the Korean Stops, Affricates and Fricatives". *J. Phon.* **2**: 161-180.

Kim, C. W. 1965. "On the Autonomy of the Tensity Feature in Stop Classification (With Special Reference to Korean Stops)". *Word* **21**: 339-359.

________. 1970. "Theory of Aspiration". *Phonetica* **21**: 107-116.

Klatt, D. H. 1980. "Software for a Cascade/Parallel Formant Synthesizer". *J. Acoust. Soc. Am.* **67(3)**: 971-985.

Ladefoged, P. 1982. "The Linguistic Uses of Different Phonation Types". *UCLA Working Papers in Phonetics* **54**: 28-39.

Lee, C. Y. and T. S. Smith. 1972. "Oral and Direct Subglottal Pressure in Korean Stops". *J. Acoust. Soc. Am.* suppl. **51 (1)**: K10.

Appendix: Constant and variable parameter time functions that control the synthesizer.

SYM	V/C	MIN	VAL	MAX
sr (sampling rate)	c	5000	10000	20000
du (duration)	c	30	700	5000
ui (update interval)	c	1	5	20
f0 (fundamental frequency)	v	0	1000	5000
f1 (first formant)	v	180	500	1300
f2 (second formant)	v	550	1500	3000
f3 (third formant)	v	1200	2500	4800
f4 (fourth formant)	v	2400	3250	4990
f5 (fifth formant)	v	3000	3700	4990
f6 (sixth formant)	v	3000	4990	4990
fz (frequency nasal zero)	v	180	280	800
fp (frequency nasal pole)	v	180	280	500
ah (amplitude of aspiration)	v	0	0	80
at (amplitude of turbulence)	v	0	0	80
af (amplitude of frication)	v	0	0	80
a1	v	0	0	80
a2	v	0	0	80
a3 (amplitudes parallel	v	0	0	80
a4 formants)	v	0	0	80
a5	v	0	0	80
a6	v	0	0	80
an (amplitude parallel nasal formant)	v	0	0	80
ap (amplitude voicing parallel)	v	0	0	80
g0 (overall gain control)	v	0	60	80
db	v	0	0	400
nf (number of formants in cascade vocal tract)	c	1	5	8
ss (source switch)	c	1	2	2
rs (random seed)	c	1	1	99
av (amplitude of voicing)	v	0	60	80
b1 (first bandwidth)	v	30	60	1000
b2 (second bandwidth)	v	40	90	1000
b3 (third bandwidth)	v	60	150	1000
b4 (fourth bandwidth)	v	100	200	1000
b5 (fifth bandwidth)	v	100	200	1500
b6 (sixth bandwidth)	v	100	500	4000
bz (bandwidth nasal zero)	v	40	90	1000
bp (bandwidth nasal pole)	v	40	90	1000
oq (percent of voicing period with glottis open)	v	10	50	80
tl (spectral tilt of voicing)	v	0	0	34

SYM	V/C	MIN	VAL	MAX
sk (skew to alternative periods)	v	0	0	100
p1	v	30	80	1000
p2	v	40	200	1000
p3 (bandwidths parallel	v	60	350	1000
p4 formants)	v	100	500	1000
p5	v	100	600	1500
p6	v	100	800	4000
ab (bypass path amplitude)	v	0	0	80
os (output waveform selector)	c	0	0	20
df	v	0	0	100

OHIO STATE UNIVERSITY WORKING PAPERS IN LINGUISTICS **38** (121-140)

The Accentual Pattern and Prosody of the Chonnam Dialect of Korean[1]

Sun-Ah Jun

Ohio State University

Abstract: This paper examines the pitch accent properties of the South Cholla (Chonnam) dialect. The Chonnam dialect has two kinds of accentual patterns, Low-High-Low and High-High-Low, which are determined by the properties of the first segment of a phrase. The domain of the accentual pattern is a Phonological phrase in the prosodic hierarchy postulated in Selkirk (1980, 1984) and Nespor and Vogel (1982, 1986), assuming a phonological word is the same as a morphological word, a lexical word plus any postnominal particles and inflectional endings. Basically each phonological word can form one accentual phrase. However, more than one phonological word can form one accentual phrase due to semantic factors or focus variation. It is also argued that the Chonnam dialect has two more units of prosodic structure above the phonological phrase within an utterance, namely the Intermediate Phrase and the Intonational Phrase (Beckman and Pierrehumbert 1986). An Intermediate phrase in Chonnam is characterized as the domain of downstep between phonological phrases. An Intonational phrase is characterized by a High or High-Low boundary tone. The data used in this paper are from the dialect spoken in Kwangju, the main city in the Chonnam province.

0. Introduction

The Korean language can be divided into eight dialects, roughly corresponding to provincial areas: Hamkyung, Pyungan, Hwanghay, Central dialect (Kyungki, Kangwon), Chungchung, Kyungsang, Cholla, and Cheju. Among these, it has been claimed (e.g. Ramsey 1978) that the richest accentual systems are found in the dialects of Hamkyung and Kyungsang. It is also believed that the lexical pitch accent systems of these modern dialects of Korea are the remnants of the Middle Korean tone system. (Middle Korean was the language of 15th and 16th century Korea and apparently had lexical tone properties similar to those of Chinese.)

The pitch accent properties of the South Cholla (henceforth Chonnam) dialect are poorly studied. Yi (1983) only mentions that North Kyungsang and Chonnam

[1] To appear in Y. Yang, S. Kuno, I. H. Lee and J. Whitman, eds., *Harvard Studies in Korean Linguistics III* (1990).

dialects have pitch accents and still retain a vowel length contrast.[2] In addition to having a different pattern of intonation from that of the standard (Seoul) dialect, the Chonnam dialect differs from other dialects in that it has a property of phrasal pitch accent.

In this paper, I argue that the pitch accent of Chonnam dialect is not the property of lexical items, but of phrases, and that the phrasal domain of the pitch accent is the Phonological Phrase (P-Phrase) in the Prosodic Hierarchy postulated in Selkirk (1980,1984) and Nespor and Vogel (1982,1986), etc. Furthermore I argue that the Chonnam dialect has three more units of prosodic structure above the Phonological phrase - namely the Intermediate phrase, the Intonational phrase and the Utterance, based on such features as the domain of downstep. Finally I show that these prosodic structures are not isomorphic to syntactic constituents.

The data used in this paper are from the dialect spoken in Kwangju, the main city in the Chonnam province. The primary informant was the author, who lived in Kwangju until her mid-twenties. Corroborating judgements and utterances were obtained from other Chonnam native speakers, I. Park and Y. Yoon. Throughout this paper, *E* refers to a schwa sound, *N* refers to a velar sound, x^h an aspirated sound and *x'* a tense or glottalized sound. Also, a lenis stop is voiced between [+sonorant] segments within the domain of the pitch accent. But for easy recognition of a lexical item, the underlying form is used. Curly braces, { }, are used for representing an accentual phrase.

1. The Chonnam Accentual Pattern

In the Chonnam dialect there are two kinds of pitch accent patterns: one is Low-High-Low (LHL) and the other is High-High-Low (HHL). Each tone is associated with one mora from left to right within a domain, which I will call the accentual phrase, and which is the topic of the next section. Any leftover tone is associated with the final mora and any leftover mora is associated with the final tone. Since an accentual phrase is often exactly one phonological word, one might assume that a word has its own lexical tonal property. The following are representative examples illustrating accentual patterns of phrases which consist of a single lexical item only:

[2] Yi (1983) claims that a long vowel had a rising tone in Middle Korean and that even though the tone of a word disappeared around the end of the 16th century, two-dot words appear to have retained length even to the present time, especially in the first syllable. So he claims that vowel length contrasts in modern dialects of Korean indicate that Korean was a tone language.

(1)	LHL		HHL	
	pə̀̂m	'a night'	*sâl*	'flesh'
	pàâm	'a chestnut'	*s'áâl*	'rice'
	nàmûl	'herb'	*hánɨ̂l*	'a sky'
	nàámù	'a tree'	*t^{h}ók'î*	'a rabbit'
	àpÉci	'father'	*hálmÉnì*	'grandma'
	kòókùmà	'sweet potato'	*k^{h}ók'írì*	'an elephant'
	àrɨ́mtàùn	'beautiful'	*síïpk'è*	'easily'
	kìtárìtà	'to wait'	*c'árɨ́tà*	'to cut'

(Acute accent marks H tone and grave accent L tone. In monosyllabic words with short vowels, the HHL pattern is phonetically realized as HL by the twin sister convention.)

A Phonological word in Chonnam consists of a lexical item (Noun, Verb, Adjective, Adverb, or Determiner, etc.) and any postnominal particles or inflectional endings, since palatalization of dental stops, which applies only lexically, also applies between a lexical item and particles as in *tot* 'to rise' + *i* 'nominalizer' ==> *toci* 'rising' and *pat^h* 'a field' + *i* 'Subj.' ==> *pac^{h}i* 'a field - subject' (Cho 1987). Thus, when a noun is followed by more than one postnominal particle, they always form one accentual phrase as a single phonological word. For example, *sasɨm* 'a deer' + *tɨl* 'PL' + *k'aci* 'also' + *coc^{h}a* 'even' + *to* 'also' ==> {*sásɨ́mtɨlk'acicoc^{h}ato*}[3] 'even deer, too,...' or *pat* 'to receive' + *ass* 'PAST' + *ta* 'verbal ending' ==> {*patátt'a*} 'received'. However the fact that a phonological word often forms a domain of accentual phrase does not mean that the pattern of accentual phrase is a property of a lexical item. Four facts show that the tonal or accentual property is determined postlexically.

1.1 Arguments for the Post-lexical Status of the Accent Pattern

First, the accentual pattern of each lexical item is predictable and exceptionless. That is, the tonal property is not lexeme-specific. Whether a lexical item has the LHL pattern or the HHL pattern depends only on its first segment without exceptions. When the first segment has a laryngeal feature of either [+spread] or [+constricted] such as /*p^h, t^h, k^h, c^h, h, s, p', t', k', c', s'*/, the phrase shows the HHL accentual pattern and any phrase whose first segment does not have these laryngeal properties shows the LHL pattern. The following sets of related onomatopoeic words confirm the fact that only the phrase initial segment is relevant to the type of each accentual pattern.

(2) a. LHL: *tàlláNtallaN* 'tinkling jingling'
HHL: *t'álláNt'allaN* 'louder noise than usual tinkling jingling'
HHL: *t^{h}álláNt^{h}allaN* 'lighter noise than usual or sounds empty'

[3] Henceforth, any vowel lacking a tone mark is assumed to have a L.

b. LHL: *pànc'ákpanc'ak* 'twinkle twinkle'
HHL: *p'ánc'ákp'anc'ak* 'much more twinkling than usual'

c. LHL: *còólcol* 'trickling or bubbling noise or its appearance'
HHL: *c'óólc'ol* 'weakened noise of bubbling due to small amount of water'

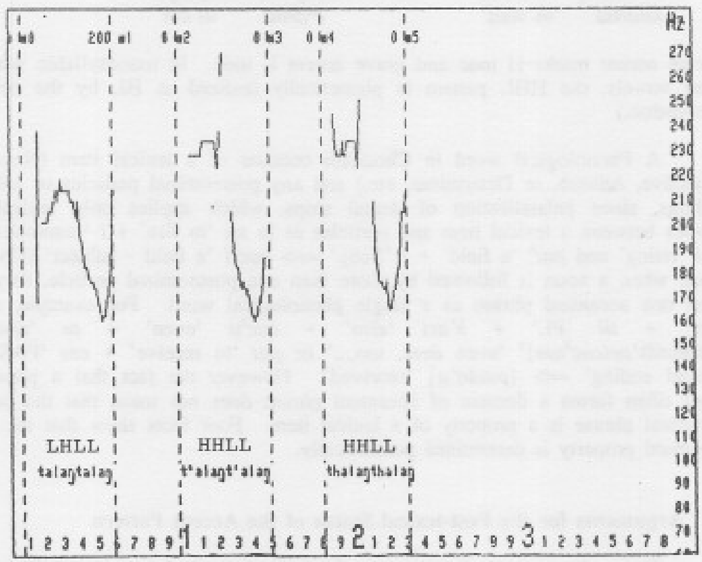


Figure 1: F0 tracks of utterances in (2a).

Figure 1 shows the fundamental frequency (henceforth F0) tracks for the three words in (2a) above, illustrating the contrast between LHL and HHL. Since the two kinds of accentual patterns differ only in their first tone, L or H, and the tones are predictable from the phrase-initial segment, we can characterize the Chonnam accentual pattern as not lexeme-specific; it could be determined postlexically, as a phrasal accent.

The second argument for characterizing these patterns as a phrasal accent involves the interaction between the domain of phrasal accent and that of vowel shortening and that of resyllabification. In some dialects of Korean which still preserves the Middle Korean vowel-length contrast, a long vowel occurs only in word-initial position. The Chonnam dialect has a vowel shortening rule which applies both lexically and postlexically.

The rule is that a long vowel in a word-initial syllable becomes short when the word occurs non-initially in a compound such as *s<u>aa</u>ram* 'a man' of *nuuns'<u>a</u>ram* 'a snowman' or *s<u>uu</u>kEn* 'a towel' of *sons'<u>u</u>kEn* 'a handkerchief'.[4] This is the lexical application of the rule. The vowel also becomes short when the word occurs in certain non-phrase-initial positions such as *ǐ* 'this' + *s<u>áá</u>ràm* 'a man' ==> {*ì s<u>á</u>ràm*} 'this man' or *nòóràn* 'yellow' + *s<u>uu</u>kEn* 'a towel' ==> {*nòóràn s<u>ù</u>kEn*} 'a yellow towel'. This is the postlexical version of the rule. Note, however, that if these phrases form two accentual phrases with each word constituting one accentual phrase, then each second word is pronounced with the first syllable long, as in {*ǐ*} {*sááràm*} and {*nòóoràn*} {*súúkÈn*}. A formalization of the vowel shortening rule looks like the following.

(3) V: --> V / { C_0^2 VX __
(here, X is any string, and { is a domain boundary.)

In addition to the vowel shortening rule, the domain of postlexical resyllabification is also the same as that of an accentual pattern, since the vowel shortening rule is accompanied by a resyllabification of the coda of the preceding syllable if the shortened vowel is originally syllable initial. For example, *EEp* 'not exist' of {*sónÉpnɨn*} {*inhyÊN*} 'a doll which has no hand' (*sôn* 'a hand' + *ÈÊp* + *nɨn* 'relative marker' + *inhyÊN* 'a doll') is shortened since it is not a domain-initial syllable. It is also resyllabified by getting an onset from the coda of the preceding morpheme, here *n*, when the P-word *EEpnɨn* does not form its own accentual phrase. But when *EEpnin* forms an accentual phrase by itself the resyllabification does not apply between *son* and *EEpnin* and *EEpnɨn* is not shortened since it does not satisfy the structural description of rule (3) above. Thus, since the domain of the postlexical vowel shortening rule and resyllabification rules lines up with the domain of the accentual pattern, and an accentual phrase can contain more than one phonological word, our pitch accent must be a phrasal accent. The following examples clearly illustrate that the Chonnam accentual pattern is not a property of a phonological word.

---

[4] Only a few verbal adjectives preserve their vowel length even before a syllable beginning with a vowel of a different morpheme. For example,

i. *cooh* 'to be good' + *ta* 'verbal ending' ==> {*coót^{h}a*}
cooh + *as'* 'PAST' + *ta* ==> {*coóas't'a* (**coás't'a*)}
ii. {*cEÉkt'a*} 'to be small (amount)' - {*cEÉkEs't'a*} ' was small'
iii. {*maánt^{h}a*} 'to be a lot (amount)' - {*maáas't'a*} 'was a lot'
iv. {*caákt'a*} 'to be small (height)' - {*caáas't'a*} 'was small'

(4) *kʰɨn* 'big' + *kɨ̌ɨrìm* 'a picture'- *kwa* 'and' + *càákɨn* 'small' + *kɨ̌ɨrìm* 'a picture'
==> {*kʰɨ́nkɨ́rìmkwà*} {*càákɨ̀nkɨ̀rìm*} 'a big picture and a small picture'

In (4), since 'big' and 'small' are contrasted, each is focused by being the head of an accentual phrase containing the following noun. *kɨɨrim* has a long first syllable by itself and since its first segment has neither of the relevant laryngeal features mentioned above, it gets the LHL accentual pattern if it forms its own accentual phrase. But the long vowel becomes short when the word occurs in non-initial position, as in *kʰɨnkɨrimkwa* 'big picture and...'; regardless of the original LH tone, the shortened vowel gets a H tone since the shortened vowel is the second mora of the whole phrase. Here, the first segment of the whole phrase, *kʰ*, has the [+spread] feature, and thus the whole phrase gets HHL pattern. If the LH of *kɨɨ* were a lexical tone, then the shortened vowel would get a contour tone like *kʰɨ̂nkɨ̌rim*, or would get a contour tone like *kʰɨ̂nkɨrím* but it does not. This is true for all the cases of other long vowels shortened due to the position of a word within a single accentual phrase. Thus the non-initial lexical item within a phrase does not affect the accentual pattern of the whole phrase and the accentual pattern of the phrase is associated with each mora of the whole phrase postlexically.

The third argument for the accent pattern being post-lexical is seen in the comparison of the domain of the Korean Lenis Stop Voicing rule with that of the accentual pattern. A lenis stop is voiced between voiced sounds as in (5).

(5) Lenis Stop Voicing Rule

$$\begin{bmatrix}\text{-cont, -asp}\\ \text{-tense}\end{bmatrix} \text{--> [+voice]/[+voice]\_\_[+voice]}$$

Cho (1987) argues that this voicing rule applies both lexically and postlexically. So, in addition to the fact that *apEci* becomes *abEji*, *kɨ* 'that' + *cip* 'a house' becomes *kɨjip* and *kʰɨn* 'big' + *kɨɨrim* 'a picture' becomes *kʰɨngɨrim* by the Lenis Stop Voicing rule. Cho (1987) claims that in the Seoul dialect the domain of this rule is P-Phrase (with some stipulation over Selkirk's end-based theory). In the Chonnam dialect, the domain may correspond to the domain of accentual phrase because in Chonnam *cip* 'a house' or *kɨɨrim* 'a picture' is not voiced when it forms an accentual phrase of its own. Also the *p* of *pap* 'rice' in *kɨgamEknɨnbap* 'the rice he eats' (*kɨ* 'he' + *ka* 'Subj.' + *mEk* 'to eat' + *nɨn* 'Rel. marker' + *pap* 'rice') is voiced when the whole phrase forms one accentual phrase but is not voiced when it forms its own accentual phrase, as in {*kɨgámENnɨn*} {*pǎp*}. Therefore, since the domain of the postlexical Lenis Stop Voicing rule also lines up with that of the accentual pattern, the domain of the accentual pattern should be larger than a phonological word, and the accentual pattern should be determined post-lexically.

The fourth argument for characterizing these patterns as a phrasal accent involves words which have an alternate contracted form. When the contracted word forms one accentual phrase together with a following or preceding word, the tone of the deleted or contracted syllable ignores the original morpheme boundary and only refers to the number of the syllables in the whole phrase in which it occurs. Consider the following examples.

(6) a. *ùrí/ù̂l* 'our' + *Èmmâ* 'mother' ==> {*ùrìÈmma*}/{*ùrÉmmà*} 'our mother'
b. *cà̂n* 'small' + *òȉ /wè̂* 'cucumber' + *nɨn* 'Topic marker' ==> {*cànóinɨn*}/{*cànwénɨn*} 'a small cucumber is ...' or 'as for a small cucumber'

Here, *oi* has a contracted form as *we*, one syllable with a short vowel, and *uri* becomes *ul*, one syllable with short vowel.[5] However, when they form one accentual phrase together with the following or preceding noun (as indicated by the resyllabification of *ul* as *ur*), the tone of the deleted vowel does not care about its original morpheme boundary. Instead each mora gets a tone automatically according to the position within the newly formed accentual phrase. In other words, the tonal pattern, LHL or HHL, determined by the laryngeal feature of the first segment of the whole phrase, matches to morae from left to right within an accentual phrase.

So far we have seen that the Chonnam pitch accent is a phrasal accent. The pitch accent assignment rule, therefore, should be postlexical and the domain of this rule should be larger than the phonological word. It can be identified with the P-Phrase in the prosodic hierarchy suggested by Nespor and Vogel (1982, 1986) and Selkirk (1980, 1986), and others.

1.2. The Underlying Form and Derivational Rules

Since there are two accentual patterns in Chonnam, we might assume both of the accentual patterns as basic and assign each accentual pattern directly to phrases depending on their first segments. On the other hand, we may also assume one of the accentual pattern as basic and derive the other pattern by phonological rules. Under the second assumption, I argue that the LHL pattern is underlying and HHL is derived from that, since it is known cross-linguistically that aspirated or glottalized segments cause higher F0 at the onset of the following vowel (Hombert 1978). Moreover, Hardcastle (1973), Kim (1965, 1970), Kagaya (1974) and others all show that the F0 at the onset of a vowel following a Korean aspirated or fortis (or glottalized) initial stop is higher than that following a lenis (or voiceless unaspirated) initial stop, although there is some overlapping in the F0 values for some subjects.

---

[5] *uri* usually becomes *ul* before certain nouns referring to a close family relation and beginning with a vowel such as *Emma* 'mom', *EmmEni* 'mother', *ap'a* 'dad', *apEci* 'father', *Enni* 'elder sister', *op'a* 'elder brother', and *eeki* 'baby', but not before *imo* 'aunt' or *acEs'i* 'uncle' or 'a man of one's parents' age'.

These phonetic explanations, however, do not mean that the HHL pattern is a low-level product of the physiological effect. Rather, the physiological effect only provides a good motivation for assuming the LHL pattern as the underlying one. This is clear if we compare the F0 contours of the CV series in a non-accentual dialect of Korean like Seoul dialect with those of Chonnam. The Chonnam HHL accentual pattern is different from the pattern of phonetically raised F0 mentioned above in that the beginning of the Chonnam HHL pattern shows a steady state of high frequency until it reaches the second mora as shown in Figure 2, while the portion of the vowel with naturally raised F0 after aspirated or glottalized stops in the Seoul dialect is much shorter according to the reports of it in the literature (e.g. Han and Weitzman 1970).

Figure 2 shows the F0 tracks of nonsense words of stop and vowel uttered by a Chonnam native speaker.

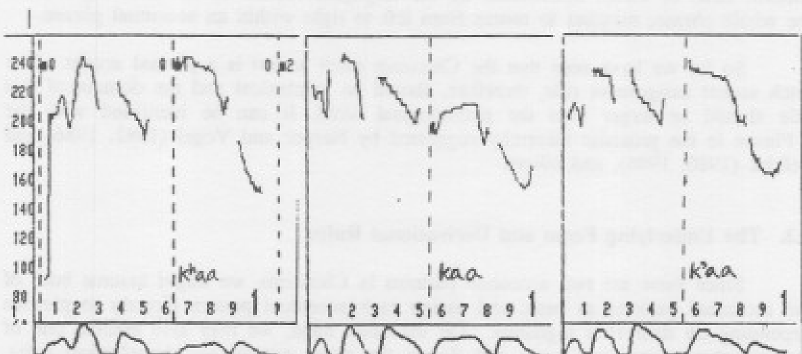
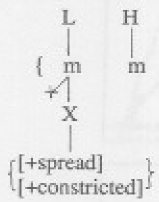


Figure 2: The nonsense words *kʰaa-kaa-k'aa* uttered by a Chonnam speaker in the carrier sentence '*IkEsin* ___ *ta*' 'this is ___'.

Accordingly, assuming the LHL accentual pattern is underlying, three phonological rules would be needed to derive the HHL pattern. First, within the domain of accentual phrase one-to-one left-to-right mapping would apply between tones and morae (Tone-Mora Mapping). Next, if the first segment has the relevant laryngeal features, delete the first L and finally, spread H to its left mora (L-delinking and H-spreading) shown in (7).

Since L-delinking and H-spreading only occur at the beginning of a P-Phrase we need a left end of P-Phrase mark for them. Conversely, since the L tone spreads only to the end of P-Phrase we need a right end of P-Phrase mark for the L-spreading rule.

(7) a. L-delinking & H-spreading



(where { is a mark for the beginning of the domain and 'm' for mora X for CV-skeleton and +| is Pulleyblank (1986)'s notation for something with no association to the left.)

Among segments containing the relevant laryngeal features, only the status of *s* is a problem, because Korean *s* has usually been assumed to belong with the lenis stop phonemes, /*p*, *t*, *k*/, for various reasons, one of which is that it is subject to the Korean Tensing rule after a nasal or obstruent coda as are the lenis stops (Obstruent Tensing rule).

(8)
a. *nun* 'an eye' + *tooNca* 'a ball' ==> *nunt'oNca* 'an eyeball'
b. *nun* 'an eye' + *pyEEN* 'disease' ==> *nunp'yEN* 'an eye disease'
c. *nun* 'an eye' + *cimcak* 'a guess' ==> *nunc'imcak* 'eye measurement'
d. *nuun* 'snow' + *kil* 'a road' ==> *nuunk'il* 'a snow road'
e. *nuun* 'snow' + *saaram* 'a man' ==> *nuuns'aram* 'a snowman'
f. *cip* 'a house' + *caNsa* 'a dealing' ==> *cipc'aNsa* 'a house dealing business'
g. *cip* 'a house' + *saaram* 'a man' ==> *cips'aram* 'a housewife'

However, Kagaya (1974) claims that Korean *s* belongs with the aspirated consonants in that it causes a higher F0 on the following vowel just as p^h, t^h, k^h do and also in that its glottal width over time in a CV sequence is similar to that of the aspirated stops. Figure 3 shows the glottal width over time for representative utterance samples of each type of consonant in a CV sequence (Kagaya 1974: 164). Since any phrase beginning with *s* starts with a High tone in Chonnam dialect, this supports the argument for Korean *s* belonging with the aspirated consonant phonemes rather than with the lenis phonemes.

2. Domains of the Accentual Pattern and the Prosodic Structure of Chonnam

So far we have seen that the Chonnam accentual pattern is a property of a phonological phrase at some level. Our next question is what determines the

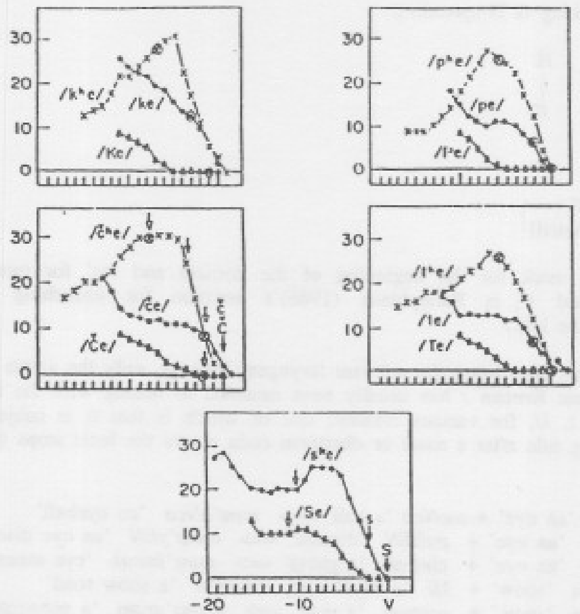


Figure 3: The ordinate gives the apparent glottal width in an arbitrary scale. A small circle marks articulatory explosion. An empty head arrow marks frication onset; a filled head arrow marks frication offset.

phonological phrasing at this level. Although the domain of the accentual pattern roughly matches syntactic constituents, in a conversation the phrasing is influenced by factors other than syntactic constituents, such as semantic factors, length of domain, emphasis on some specific morphemes, or morpheme inherent properties. However, based on the degree of generality across possible accentual domains, we can find several criteria by which we can predict a possible accentual phrasing.

2.1. Domains of the Accentual Pattern

Basically, each phonological word (henceforth P-Word) can potentially form one accentual domain (henceforth P-Phrase). But a P-word does not always form its own P-Phrase. Rather it can form one P-Phrase together with the preceding or following P-Word. When the P-word is a noun and it is 'old information' expectable from the context, or it is a function word such as *te* 'a place', *pa* 'a fact', *kEt* 'a thing', etc., (which are usually bound-morpheme-like pronouns), or it functions as a kind of semantic proto-type, the P-Word, a noun, does not form its own P-Phrase but is included in a larger P-Phrase together with the preceding P-Word. So, for example, when the word *saaram* is used specifically to mean 'a man' in opposition to 'an animal', *saaram* forms its own P-Phrase, but when it is used meaning 'someone' as a pro-word, it does not form its own P-Phrase but belongs to the P-Phrase of the preceding modifier. Some examples are as follows:

(9) a. *k^hɨn* 'big' + *caaNkap* 'a glove'==> {*k^hɨn*} {*caáNkap*} 'a big glove'
b. *k^hɨn* 'big' + *kEt* 'a thing' ==> {*k^hɨ́nkÊt*} 'a big thing'
c. *cE* 'that' + *ke* 'a thing' + *saaram* 'a man' + *i* 'to be' + *nya* 'question particle' ==> {*cEkê*} {*sááraminya*}? 'Is that a man?'
d. *pyEENwEn* 'a hospital' + *e* 'in' + *iss* 'to exist or to stay' + *nɨn* 'relative particle' + *ap^{h}ɨn* 'sick' + *saaram* 'a man' ==> {*pyEÉNwEne innɨn*} {*ap^{h}ɨ́nsaram*} 'someone who is sick in the hospital'

(9a) above is a normal case of two accentual phrases for {Adj} {Noun}, but in (9b) the adjective and noun form one P-Phrase {Adj+Noun} since the noun is a bound-morpheme-like pronoun. In (9d), *saaram* does not have any specific meaning but is a semantically proto-type noun, while *saaram* of (9c) has a literal meaning of 'something which is a man'. So the *saaram* in (9c) forms a separate P-Phrase of its own while that in (9d) does not.

Next, when a P-Word is intentionally or contextually focused or emphasized, it always forms one P-Phrase absorbing, regardless of any syntactic boundary, all the following P-Words within the Intonational Phrase, which otherwise would have their own P-Phrases. For example, consider (10) and (11).

(10) *acu* 'very' + *nooran* 'yellow' + *suukEn* 'a towel' ==>
a. {*acû*} {*noóran*} {*súúkEn*} 'a very yellow towel'
b. {*acú noran sukEn*} 'a VERY yellow towel'
c. {*acû*} {*noóran sukEn*} 'a very YELLOW towel (not a green towel)'
d. {*acû*} {*noóran*} {*súúkEn*} 'a very yellow TOWEL (not a yellow book)'

(11) *Emma* 'mom' + *ka* 'subject marker' + *mantɨl* 'to make' + *n* 'relative marker' + *mantukuk* 'a wonton soup' + *mEk* 'to eat' + *Ess* 'PAST' + *E* 'verbal ending' ==>

a. {*Emmáka*} {*mantɨ́n*} {*mantúkuk mEkEs'E*}[6] '(I) ate a wonton soup which mom made.'
b. {*Emmáka mantɨn mantukuk mEkEs'E*} '(I) ate a wonton soup which MOM made'
c. {*Emmáka*} {*mantɨ́n mantukuk mEkEs'E*} '(I) ate a wonton soup which mom MADE'
d. {*Emmáka*} {*mantɨ̂n*} {*mantúkuk mEkEs'E*} '(I) ate A WONTON SOUP which mom made'

Both examples are cases where the whole utterance consists of one Intonational Phrase. (10) is an example of a syntactic structure $_{NP}[_{AdjP}[[Adv][Adj]]_{AdjP}[Noun]]_{NP}$ and (11) has the syntactic structure $_{S}[_{NP}[_{S}[[NP][VP]]_{S}[head\ Noun]]_{NP}[VP]]_{S}$. The P-Phrases in (10a) and (11a) above are for the neutral utterance, 'neutral' in the sense that it is uttered without focusing any special P-Word. (10b) and (11b) examples are in the case of when the first P-Word is emphasized, with the whole utterance forming one P-Phrase. 'c' examples are of when the second P-Word is emphasized, resulting in two P-Phrases.

Sentences (10d) and (11d), where the final P-Phrase is emphasized, show the same structure of P-Phrases as (10a) and (11a). However, the final P-Phrase of the 'd' utterances is phonetically different from those of the 'a' utterances; the F0 peak of the final P-Phrase is higher in the 'd' version than it is in the case of the 'neutral' utterance, the duration of the word is longer than normal, and the intensity of the word is stronger than normal. The comparison of emphasized with unemphasized accentual patterns in (11) is displayed in Figure 4.

As shown above, if any word in a sentence is focused it initiates a new P-Phrase, which absorbs all following P-Words regardless of syntactic boundaries. In the case of (11c), a verb in a Relative clause, a head noun, and a main verb all form one P-Phrase. In other words, the principle 'one P-Word equals one P-Phrase' is overridden by semantic or pragmatic factors and, therefore, the arrangement of P-Phrases is not always isomorphic to syntactic structure. In summary, a Chonnam P-Phrase usually consists of one P-Word, thus, the domain of P-Phrase often matches to a syntactic constituent. However, this P-Phrase formation rule is overridden by semantic factors, or by focus, or by the length of a constituent.

---

[6] Usually, if a verb is closely related to a single object noun semantically and contextually, the verb P-Word does not form its own P-Phrase. So, for example, if a verb is *pEryEtt'a* 'threw' in this example, it is only when the noun is in narrow-focus that the Noun + Verb forms one P-Phrase. On the other hand, if 'ate' is emphasized the sentence would form four P-Phrases as in {*Emmáka*} {*mantɨ́n*} {*mantúkuk*} {*mEkÉs'E*}.

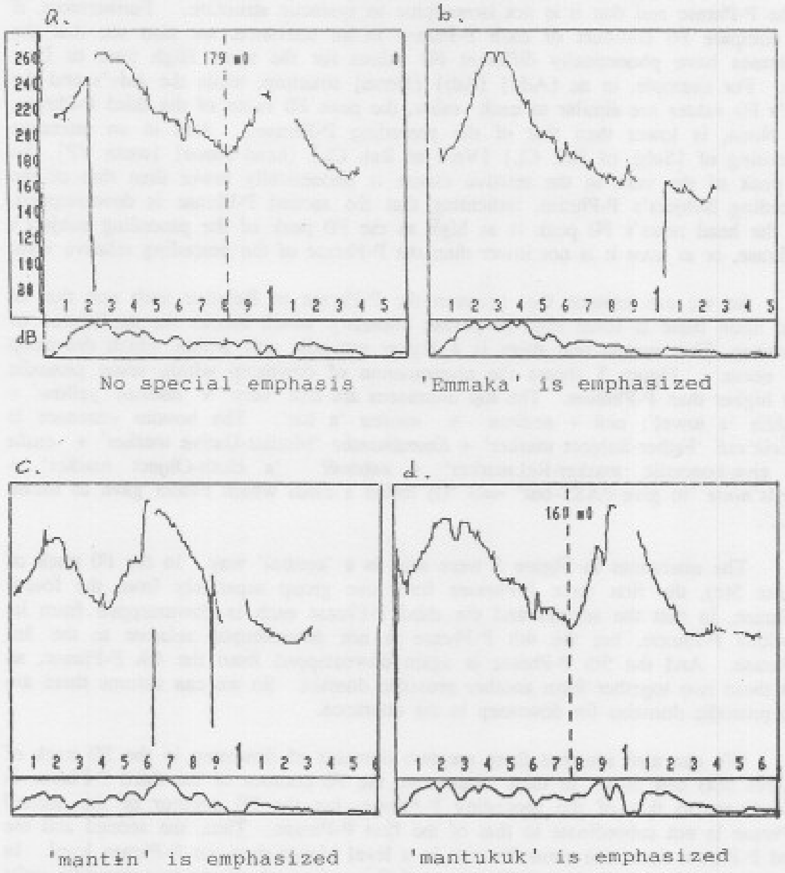


Figure 4: F0 tracks of utterances in (11) above.

2.2. Prosodic Structures of Chonnam

So far it has been shown that the domain of the Chonnam accentual pattern is the P-Phrase and that it is not isomorphic to syntactic structure. Furthermore, if we compare F0 contours of each P-Phrase in an utterance, we also see that the P-Phrases have phonetically different F0 values for the same High tone or Low tone. For example, in an {Adv} {Adj} {Noun} structure, while the Adv's and the Adj's F0 values are similar to each other, the peak F0 value of the third P-Phrase, the Noun, is lower than that of the preceding P-Phrases. And in an utterance consisting of {Subj. of Rel. Cl.} {Verb of Rel. Cl.} {head Noun} {main VP}, the F0 peak of the verb in the relative clause is phonetically lower than that of the preceding Subject's P-Phrase, indicating that the second P-Phrase is downstepped, but the head noun's F0 peak is as high as the F0 peak of the preceding subject's P-Phrase, or at least it is not lower than the P-Phrase of the preceding relative verb.

So we can assume that between the P-Phrase of Relative verb and that of head noun there is some kind of phrase boundary which blocks the application of downstep. This means that there is a higher prosodic unit within which downstep can occur. Figure 5 shows the phenomenon of downstep within some prosodic unit higher than P-Phrases. The top utterances are *acû* 'very' + *noóran* 'yellow' + *súúkEn* 'a towel'; *acû* + *noóran* + *moóca* 'a hat'. The bottom utterance is *apÉcik'esE* 'Father-Subject marker' + *Emmáhanthe* 'Mother-Dative marker' + *cusín* 'to give-honorific marker-Rel.marker' + *otkámɨl* 'a cloth-Object marker' + *marás'nɨnte* 'to give-PAST-but' ==> '(I) rolled a cloth which Father gave to Mom, but...'.

The utterances in Figure 5 were said in a 'neutral' way. In the F0 track of Figure 5(c), the first three P-Phrases form one group separately from the fourth P-Phrase, in that the second and the third P-Phrase each is downstepped from its previous P-Phrase, but the 4th P-Phrase is not downstepped relative to the 3rd P-Phrase. And the 5th P-Phrase is again downstepped from the 4th P-Phrase, so that these two together form another prosodic domain. So we can assume there are two prosodic domains for downstep in the utterance.

We can also see that there are two domains of downstep in the F0 track of Figures 5(a) and 5(b). In these utterances, the F0 contour of the third P-Phrase is subordinate to that of the preceding P-Phrase, but the F0 contour of the second P-Phrase is not subordinate to that of the first P-Phrase. Thus, the second and the third P-Phrase form one prosodic unit in a level higher than the P-Phrase level. In other words, there are three P-Phrases and above them there are two prosodic units which are the domain of downstep. At the same time, this shows that the higher prosodic phrase is also not isomorphic to syntactic structure since the adverb which is the first P-Phrase and the adjective that it modifies (the 2nd phrase) belong to different downstep domains even though they form an Adjective Phrase syntactically.

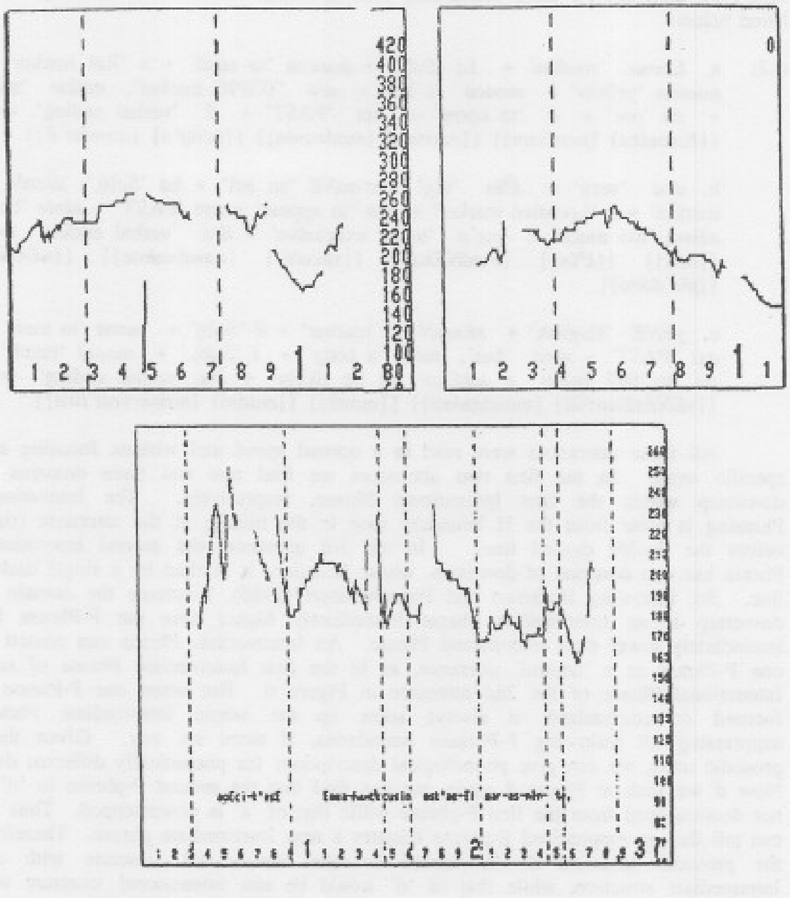


Figure 5: Downstep of P-Phrases.

Moreover, if we consider an utterance containing more than one Intonational Phrase we can also find that the Intonational Phrase is not the prosodic unit immediately higher than the P-Phrase, because we can find that within one Intonational Phrase, there is more than one phrasal domain where downstep occurs. The domain immediately higher than the P-Phrase has no special boundary tone while an Intonational Phrase necessarily has H or HL boundary tone. Figure 6 shows F0 contours for three utterances, each of which is two intonational phrases, listed below:

(12) a. *Emma* 'mother' + *ka* 'Subj' + *ponecu* 'to send' + *n* 'Rel marker' + *nooran* 'yellow' + *mooca* 'a hat' + *nɨn* 'TOPIC marker', *acikto* 'still' + *an* 'not' + *o* 'to come' + *ass* 'PAST' + *E* 'verbal ending'. ==> {{*Emmáka*} {*ponécun*}} {{*noóran*} {*moócanɨn*}} {{*acíkt'o*} {*anwáts'E*}}.

b. *acu* 'very' + *k^hɨn* 'big' + *c'aaNE* 'an eel' + *ka* 'Subj.', *siicaN* 'a market' + *e* 'Locative marker' + *nao* 'to appear' + *ass* 'PAST' + *nɨnte* 'but', *nEmu* 'too much' + *pis'a* 'to be expensive' + *tɨra* 'verbal ending'. ==> {{*acû*}} {{*k^hɨn*}} {*c'aáNEka*}} {{*siícaNe*} {*nawánnɨnte*}} {{*nEmû*}} {{*pis'átɨra*}}.

c. *yENE* 'English' + *sEnseN* 'a teacher' + *ɨl* 'Subj' + *mana* 'to meet' + *ass* 'PAST' + *nɨnti* 'but', *mom* 'a body' + i 'Subj.' + *maani* 'much' + *na* 'to feel better' + *ɨnkEkat* 'to be likely' + *ɨra* 'verbal ending'. ==> {{*yENÉsEnseNɨl*} {*manánnɨnti*}} {{*momí*}} {{*maáni*} {*naɨ́nkekatt'Era*}}.

All these utterances were read in a normal speed and without focusing any specific word. In the first two utterances we find two and three domains of downstep within the first Intonational Phrase, respectively. The Intonational Phrasing is clear from the H boundary tone in the middle of the utterance (right before the double dashed line). In the 3rd utterance, the second Intonational Phrase has two domains of downstep, whose boundary is marked by a single dashed line. So, following Beckman and Pierrehumbert (1986), I assume the domain of downstep is an Intermediate Phrase immediately higher than the P-Phrase but immediately lower than Intonational Phrase. An Intermediate Phrase can consist of one P-Phrase in a 'neutral' utterance, as in the first Intermediate Phrase of each Intonational Phrase of the 2nd utterance in Figure 6. But when one P-Phrase is focused or emphasized, it always takes up the whole Intermediate Phrase, suppressing all following P-Phrases boundaries, if there are any. Given these prosodic units, we can give phonological descriptions for phonetically different data. Now if we look at Figure 4 again, we can find that the second P-phrase in 'd' is not downstepped from the first P-phrase while that of 'a' is downstepped. Thus we can tell that an emphasized P-phrase initiates a new Intermediate phrase. Therefore, the prosodic structure of 'a' would be one Intonational structure with one Intermediate structure, while that of 'd' would be one Intonational structure with two Intermediate structures.

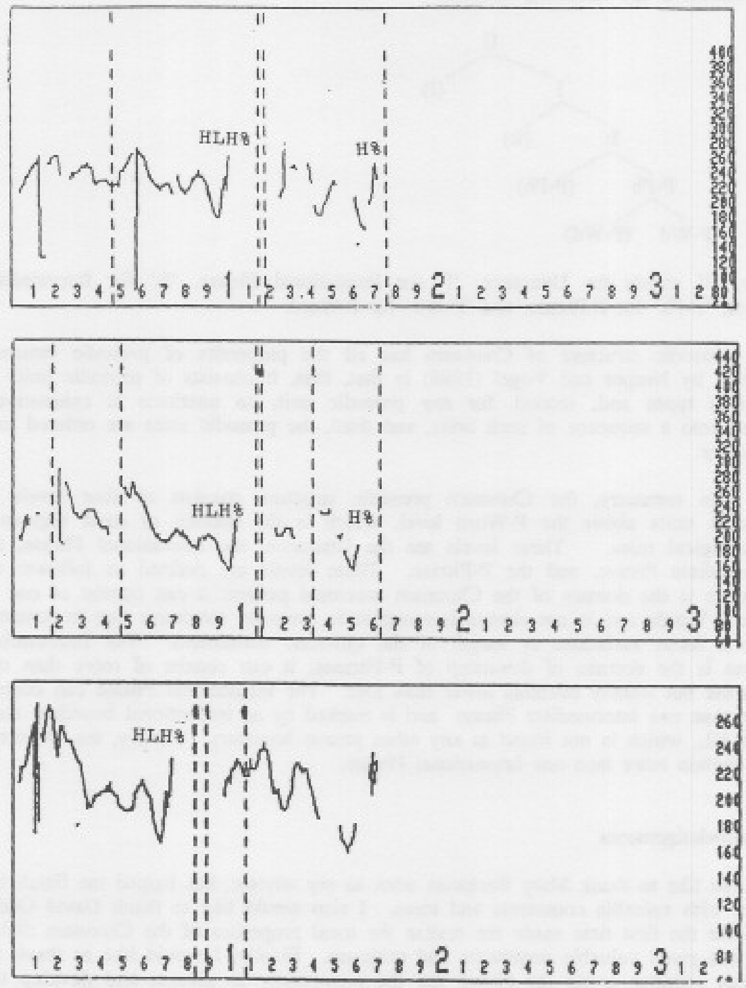
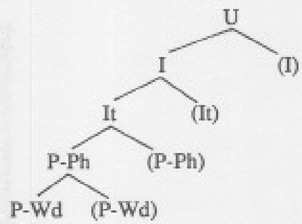


Figure 6: Utterances of two Intonational Phrases in (12). Single dashed line marks a boundary between Intermediate phrases and a double dashed line marks a boundary between Intonational Phrases.

Accordingly, the prosodic or phonological structure of Chonnam can be represented in the following fashion:

(13)



Here 'U' stands for Utterance, 'I' for Intonational Phrase, 'It' for Intermediate Phrase, 'P-Ph' for P-Phrase, and 'P-Wd' for P-Word.

This prosodic structure of Chonnam has all the properties of prosodic structure claimed by Nespor and Vogel (1986) in that, first, it consists of prosodic units of different types and, second, for any prosodic unit, an utterance is exhaustively parsed into a sequence of such units, and third, the prosodic units are ordered in a hierarchy.

In summary, the Chonnam prosodic structure consists of four levels of prosodic units above the P-Word level, which is the domain of most segmental phonological rules. These levels are the Utterance, the Intonational Phrase, the Intermediate Phrase, and the P-Phrase. These levels are defined as follows: the P-Phrase is the domain of the Chonnam accentual pattern; it can consist of one or more P-Words and is not always isomorphic to syntactic structures due to semantic factors, focus variations or length of the syntactic constituent. The Intermediate Phrase is the domain of downstep of P-Phrases; it can consist of more than one P-Phrase but usually contains fewer than four. The Intonational Phrase can contain more than one Intermediate Phrase and is marked by an intonational boundary tone, H or HL, which is not found at any other phrase boundary. Finally, the Utterance can contain more than one Intonational Phrase.

Acknowledgements

I would like to thank Mary Beckman who, as my advisor, has helped me finish this paper with valuable comments and ideas. I also would like to thank David Odden who for the first time made me realize the tonal properties of the Chonnam dialect and has given valuable comments and criticism. Finally, I would like to thank the Korean seminar group for giving me the opportunity to present and develop this paper.

References

Beckman, M. and J. Pierrehumbert. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook* 3: 255-309.

Cho, Young-mee Y. 1987. The Domain of Korean Sandhi Rules. Paper presented at the LSA meeting.

____________. 1988. Syntax and Phrasing in Korean. To appear in *Phonology - Syntax Connection.*

Clements, G. N. 1985. The Geometry of Phonological Features. *Phonology Yearbook* **2:** 225-252.

Han, M. S. and R. S. Weitzman. 1970. Acoustic Features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica* **22:** 112-128.

Hardcastle, W. J. 1973. Some Observations on the Tense-Lax Distinction in Initial Stops in Korean. *Journal of Phonetics* **1:** 263-272.

Hombert, J. M. 1978. Consonant Types, Vowel Quality, and Tone. Victoria Fromkin ed. *Tone: A Linguistic Survey.* Academic Press.

Iverson, G. K. 1983. Korean *s.* Dept. of Linguistics, Univ. of Iowa.

Kaisse, E. M. and P. A. Shaw. 1985. On the Theory of Lexical Phonology. *Phonology Yearbook* **2:** 1-30.

Kagaya, R. 1971. Laryngeal Gestures in Korean Stop Consonants. *Annual Bulletin*, Research Institute of Logopedics and Phoniatrics, University of Tokyo **5:** 15-24.

_________. 1974. A Fiberscopic and Acoustic Study of the Korean Stops, Affricates and Fricatives. *Journal of Phonetics* **2:** 161-128.

Kim, C. W. 1965. On the Autonomy of the Tensity Feature in Stop Classification (with special reference to Korean stops). *Word* **21:** 339-359.

_________. 1970. A Theory of Aspiration. *Phonetica* **21:** 107-116.

McCawley, J. D. 1978. What is a Tone Language? Victoria Fromkin ed. *Tone: A Linguistic Survey*. Academic Press.

Mohanan, K. P. 1986. *The Theory of Lexical Phonology.* D. Reidel Publishing Co. Dordrecht, Holland.

Nespor, M. and Irene Vogel. 1986. *Prosody Phonology.* Foris, Dordrecht.

Odden, D. 1987. Kimatuumbi Phrasal Phonology. *Phonology Yearbook* 3:13-36.

Pulleyblank, D. 1986. *Tone in Lexical Phonology*. D. Reidel Publishing Company.

Ramsey, S. 1978. *Accent and Morphology in Korean Dialect: A Descriptive and Historical Study*. Top Publishing Co. Korea.

Selkirk, E. 1980. Prosodic Domain in Phonology: Sanskrit Revisited. M. Aronoff and M. L. Kean eds. *Juncture*. Saratoga: Anma Libri.

_________. 1986. On Derived Domains in Sentence Phonology. *Phonology Yearbook* 3.

Yi, S.O. 1983. Remarks on Tone in Middle Korean. The Korean National Commission for UNESCO ed. *The Korean Language*, Pace International Research, Inc., Oregon, U.S.A.