DOCUMENT RESUME

ED 282 278	CS 505 627

AUTHOR	O'Brien, Nancy, Ed.
TITLE	Status Report on Speech Research: A Report on the Status and Progress of Studies on the Nature of Speech, Instrumentation for Its Investigation, and Practical Applications, October-December 1986.
INSTITUTION	Haskins Labs., New Haven, Conn.
SPONS AGENCY	National Institutes of Health (DHHS), Bethesda, Md.; National Science Foundation, Washington, D.C.
REPORT NO	SR-88(1986)
PUB DATE	86
CONTRACT	NICHHD-NO1-HD-5-2910
GRANT	NICHHD-HD-01994; NIH-BRS-RR-05596; NINCDS-NS-13617; NINCDS-NS-13870; NINCDS-NS-18010; NSF-BNS-8520709
NOTE	269p.
AVAILABLE FROM	U.S. Department of Commerce, National Technical Information Service, 5285 Port Royal Rd., Springfield, VA 22151.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC11 Plus Postage.
DESCRIPTORS	*Communication Research; Language Handicaps; *Language Processing; Memory; Neurological Organization; Phonology; Second Languages; *Speech Communication

ABSTRACT

One of a series of semiannual reports, this paper presents articles exploring the status and progress of studies on the nature of speech, instrumentation for its investigation, and practical research applications. Titles of the papers and their authors are as follows: (1) "Lexical Organization and Welsh Consonant Mutations" (S. Boyce, C. P. Browman, and L. Goldstein); (2) "The Emergence of Cerebral Asymmetries in Early Human Development: A Literature Review and a Neuroembryological Model" (Catherine T. Best); (3) "The Role of Coarticulatory Effects in the Perception of Fricatives by Children and Adults" (Susan Nittrouer and Michael Studdert-Kennedy); (4) "Hemispheric Asymmetries in Phonological Processing" (G. Lukatela, Claudia Carello, M. Savic and M. T. Turvey); (5) "Processing Lexical Ambiguity and Visual Word Recognition in a Deep Orthography" (Shlomo Bentin and Ram Frost); (6) "Some Word Order Effects in Serbo-Croat" (Z. Urosevic, Claudia Carello, M. Savic, G. Lukatela, and M. T. Turvey); (7) "Short-Term Memory, Phonological Processing and Reading Ability" (Susan Brady); (8) "Phase-Locked Modes, Phase Transitions, and Component Oscillators in Biological Motion" (J. A. S. Kelso, G. Schoner, J. P. Scholz, and H. Haken); (9) "Beyond Anatomical Specificity" (M. T. Turvey); (10) "Perceptual Normalization of Vowels Produced by Sinusoidal Voices" (Robert E. Remez, Philip E. Rubin, Lynne C. Nygaard, and William A. Howell); (11) "The Stop-Glide Distinction: Acoustic Analysis and Perceptual Effect of Variation on Syllable Amplitude Envelope for Initial /b/ and /w/" (Susan Nittrouer and Michael Studdert-Kennedy); and (12) "The Development of Language Structure in Children with Down Syndrome" (Anne E. Fowler). (Citation information for each article is included.) (HTH)

HASKINS LABORATORIES

STATUS REPORT ON SPEECH RESEARCH

SR-88 *OCTOBER-DECEMBER 1986*

A Report on the Status and Progress of Studies on the Nature of Speech, Instrumentation for its Investigation, and Practical Applications

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED: The information in this document is available to the general public. Haskins Laboratories distributes it primarily for library use. Copies are available from the National Technical Information Service or the ERIC Document Reproduction Service. See the Appendix for order number of previous Status Reports.

Haskins Laboratories
270 Crown Street
New Haven, CT 06511

Michael Studdert-Kennedy, Editor-in-Chief
Nancy O'Brien, Editor
Yvonne Manning, TEXnician*

---

* This report was produced using TEX, the computer typesetting system developed by Donald E. Knuth at Stanford University. TEX is a registered trademark of the American Mathematical Society.

The research reported here was made possible
in part by support from the following sources:

NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT

Grant HD-01994

NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT

Contract NO1-HD-5-2910

NATIONAL INSTITUTES OF HEALTH

Biomedical Research Support Grant RR-05596

NATIONAL SCIENCE FOUNDATION

Grant BNS-8520709

NATIONAL INSTITUTE OF NEUROLOGICAL AND COMMUNICATIVE DISORDERS AND STROKE

Grant NS 13870
Grant NS 13617
Grant NS 18010

PERSONNEL IN SPEECH RESEARCH

Investigators

Arthur S. Abramson*
Peter J. Alfonso*
Thomas Baer
Fredericka Bell-Berti*
Catherine Best*
Geoffrey Bingham**
Gloria J. Borden*
Susan Brady*
Catherine P. Browman
Etienne Colomb***
Franklin S. Cooper*
Stephen Crain*
Robert Crowder*
Laurie B. Feldman*
Janet Fodor*
Anne Fowler**
Carol A. Fowler*
Ram Frost†
Dave Garrett‡
Louis Goldstein*
Vicki L. Hanson*
Katherine S. Harris*
Leonard Katz*
J. A. Scott Kelso*
Andrea G. Levitt*
Alvin M. Liberman*
Isabelle Y. Liberman*
Diane Lillo-Martin
Leigh Lisker*
Virginia Mann*
Ignatius G. Mattingly*
Nancy S. McGarr*
Richard S. McGowan*
Kevin G. Munhall
Hiroshi Muta[1]
Susan Nittrouer‡
Patrick W. Nye
Lawrence J. Raphael*
Bruno H. Repp
Philip E. Rubin
Elliot Saltzman
Donald Shankweiler*
Michael Studdert-Kennedy*
Betty Tuller*
Michael T. Turvey*
Douglas H. Whalen

Technical/Support

Philip Chagnon
Alice Dadourian
Michael D'Angelo
Betty J. Delise
Vincent Gulisano
Donald Hailey
Raymond C. Huey*
Sabina D. Koroluk
Yvonne Manning
Bruce Martin
Nancy O'Brien
William P. Scully
Richard S. Sharkany
Edward R. Wiley

Students*

Joy Armson
Dragana Barac
Eric Bateson
Suzanne Boyce
André Cooper
Margaret Dunn
Carole E. Gelfer
Joseph Kalinowski
Bruce Kay
Rena Krakow
Deborah Kuglitsch
Karen Kushner
Hwei-Bing Lin
Katrina Lukatela
Paul Macaruso
Harriet Magen
Judy Mahoney
Sharon Manuel
Diana Matson
Gerald McRoberts
Sheri L. Mize
Mark Pitt
Nian-qi-Ren
Lawrence D. Rosenblum
Arlyne Russo
Richard C. Schmidt
Jeffrey Shaw
Caroline Smith
Robin Story
Mark Tiede
David Williams

---

*Part-time
**NIH Research Fellow
***Fogarty International Fellow, Lausanne, Switzerland
†Postdoctoral Fellow, Hebrew University, Israel
‡NRSA Training Fellow
[1] Visiting from University of Tokyo, Japan

Status Report on Speech Research

Haskins Laboratories

CONTENTS

LEXICAL ORGANIZATION AND WELSH CONSONANT MUTATIONS

S. Boyce, C. P. Browman, and L. Goldstein

Abstract. *The role of phonological form in the lexical organization of morphologically related words is investigated using consonant mutations in Welsh. Consonant mutations are a regular form of non-affixal morphology, in which the initial consonant of a word changes depending on its syntactic context; for instance, the word* **pont** *'bridge' may appear as* **bont** *or* **font**. *In English, it has been shown that affixal variants, such as POUR-POURED, prime each other strongly, while non-affixal variants, such as HUNG-HANG, show weak or nonexistent priming (Kempley & Morton, 1982; Stanners, Neiser, Hernon, & Hall, 1979). Using the task of auditory repetition priming, we show that mutation is similar to affixing in English in that mutated variants prime each other. We further show that abstract morphological categories, rather than identity of phonological form, are required to organize the Welsh lexicon, thus suggesting that current phonologically based lexical models need to be revised. An alternative model utilizing an underspecified autosegmental representation is proposed.*

INTRODUCTION

Recent studies (e.g., Kempley & Morton, 1982) have shown that morphologically related words such as POUR and POURED share a common lexical representation. In this paper, we investigate the extent to which such sharing is constrained by the similarity in form between the related words. In particular, we ask whether such sharing is possible when the related words differ in a more complex way than the simple suffixation seen in English. To this end, we make use of an interesting feature of Welsh, the fact that initial consonants in Welsh words undergo systematic changes as a function of their syntactic and lexical context. These consonant changes, or *mutations* as they are known, are the focus of our study.

Words can be defined as those unique intersections (associations) of semantic and phonological material that can stand independently as minimal utterances. Thus, BOIL, GRAPHIC, and DOG are each a different combination of meaning and phonological form. But what of sets such

---

Aknowledgment. This work was supported by NIH grant HD-01994. We thank Dr. Frank Gooding, of the University of North Wales, Bangor, for making the facilities of the Linguistics Department available, and Carolyn Iorwerth, of the Department of Extra-Mural Welsh Instruction, for help in finding subjects. We also thank Patrice Beddor, David Fay, Len Katz, and an anonymous reviewer for helpful criticisms.

as POUR, POURING, POURED, etc? These also fulfill the definition as different words, yet there is something intuitively unsatisfying in saying that their relationship is analogous to that between, for example, BOIL, GRAPHIC, and DOG. Traditional linguistic analysis (and common sense) address this issue by pointing out that each of these words has a smaller pattern in common, POUR, which is itself a unique intersection of meaning and phonological shape. It is in this sense that POUR, POURED etc. can be alternately described as the same word in different forms, or, in the terms of the "minimal utterance" definition of a word, as different words with the common semantic/phonological pattern POUR. This common semantic/phonological pattern is referred to as the **root** morpheme, or base. In addition, linguistic theory posits **grammatical morphemes** that capture the grammatical function and phonological form that distinguish the morphological variants POUR, POURING, POURED from each other—for example, the forms -ING and -ED. In this paper we will refer to the entire set of morphologically related forms as a **morphological complex**, and words that belong to the same morphological complex as **morphological variants**. For inflectional processes (such as those involving -ING and -ED), we usually think of the morphological variants as forms of the same abstract "word" (or "lexeme," cf. Matthews, 1974), in the sense that we would expect to find only a single entry in a dictionary, corresponding to the whole morphological complex.

It has often been proposed that this special relationship between morphological variants is reflected in the organization of the lexicon; that is, words are generated or accessed according to morphological grouping (e.g., Lukatela, Gligorijević, Kostić, & Turvey, 1980; MacKay, 1979; Murrell & Morton, 1974; Stanners, Neiser, Hernon, & Hall, 1979; van der Molen & Morton, 1979). Assuming that such is the case, there are logically two ways by which these groups or access routes might be organized in the lexicon: according to the phonological form of the likeness between variants or according to their functional (semantic or syntactic) relationship. Putting this somewhat less categorically, we can ask whether there are any constraints on the kind and degree of phonological likeness, or the kind and degree of functional relationship, which in and of themselves exclude certain words from participating in such morphologically based groups or access routes. Evidence for such constraints on the possible groupings, or access routes, would obviously provide useful limitations on possible models of lexical organization.

It happens that in the most frequent and familiar English case (as in POUR, POURING, etc.) functional and formal analyses converge, thus making it difficult, if not impossible, to distinguish the two. In this "regular" English system, it is possible to decompose the phonological form of inflected words into two independent, physically continuous surface forms, the base and the suffix. These forms bear a one-to-one relation to the root meaning and grammatical tense marker. There are, however, more peripheral examples in English where formal and functional analyses diverge, in particular, in the case of "irregular" inflected nouns and verbs, whose formal relationship is often relatively arbitrary (SHAKE-SHOOK, CREEP-CREPT) but whose functional relationship is straightforwardly parallel to that of regular forms (PRESENT-PAST TENSE). In addition, there are cases of derived forms in English where formal relationships are transparent, but where functional relationships are not completely predictable (e.g., FORM-FORMAL; LYRIC-LYRICAL). These irregular examples have been the focus of a good deal of recent research (e.g., Fowler, Napps, & Feldman, 1985; Kempley & Morton, 1982; Stanners et al., 1979), and we will return to a discussion of this work, below.

Models of lexical access have tended to echo the pattern of the English regular paradigm, picturing the lexicon as based on physically continuous, separable, shared units. In one well-known model, the logogen theory of Morton (Kempley & Morton, 1982; Morton, 1969, 1970), the lexicon (or at least the **input** lexicon, cf. Morton, 1979) is a set of receptors or recognition units, each of which fires upon detecting the presence of a particular stimulus in the incoming signal. For a word such as POURED there would be two such receptors, corresponding to the base morpheme POUR and the affix -ED. The base morpheme POUR is considered to be a shared receptor, or logogen, which fires alike for each variant of the morphological complex containing POUR: POURING, POURED, etc. The logogen POUR is only excited by other variants and not by accidentally similar auditory or orthographic sequences. Moreover, because receptors can only fire for discrete, continuous units, morphological variants whose surface forms do not share this same unit cannot be grouped together, and would be represented by separate logogens (Kempley & Morton, 1982, p. 443). Presumably, this would be a problem both for pairs such as SHAKE-SHOOK, whose formal relation is irregular, and DECIDE-DECISION whose relationship is regular but requires the application of vowel and consonant-altering phonological rules. Similar notions of shared entry, and the separation of affix from base, have been proposed by Taft and Forster (1975) and Manelis and Tharp (1977), among others. In all these models, because the lexicon is organized by shared phonologically continuous units, morphological variants that do not share such a unit cannot be grouped together, or share recognition units.

Not all models of the lexicon, of course, are constrained in this way. Even if the lexicon is viewed as listing only base forms, it is possible to generate the variants by the application of stored rules (e.g., MacKay, 1979), where these rules can be either concatenative (e.g, "make a past by adding -ed") or form-changing (e.g., "make a past by changing the vowel"). Thus, this type of model allows a dictionary base entry to be shared among morphological variants that do not share a continuous surface sequence corresponding to the base. However, perceptual experiments have not, in general, supported this possibility. As discussed below, various experiments have suggested (e.g., Kempley & Morton, 1982; Stanners et al., 1979) that such sharing does not occur in the case of English irregular forms (e.g., SHAKE-SHOOK). Therefore, lexical models based on these experiments have assumed that morphological variants can share a lexical entry only if the base occurs in unmodified form in each variant.

These shared-entry approaches to lexical access are plausible in English and the more commonly studied European languages, where the primary device for encoding morphological variation is affixation. Because the exceptions to this format—primarily irregular nouns and verbs of the SHAKE-SHOOK variety—are few and idiosyncratic, it has been possible to assume that the lexicon treats them as associated, but not as sharing a lexical entry. There are many languages, however, in which there are regular morphological processes that are more complex than simple concatenation. The variants related by such processes often violate the constraint of shared, continuous base forms. Some languages, such as Tagalog, violate this requirement by inserting grammatical morphemes into the body of a base morpheme, as in /sulat/ 'to write', /sumulat/ 'wrote', /sinulat/ 'was written', from the combination of the base form /sulat/ with infixes /-um-/ and /-in-/. Many languages challenge both continuity and separability by modifying the base morpheme internally. In Semitic languages, for instance, roots are varied by vowel substitution and prosodic adjustment while maintaining a common consonantal core. In Hebrew /sapar/ 'barber' and /siper/ 'he cuts', the contrast between inserted vowels /a-a/ and /i-e/ differentiates the noun and the verb (cf. Bentin, Bargai, & Katz, 1984). For another example, compare Arabic /daras/ - /daaris/ - /darraas/ - /diraasat/, to their English equivalents 'study' (verb), 'studying', 'student', 'study' (noun), where variations in placement and doubling of the vowels /i/ and /a/ and the consonant /r/ differentiate large sets of related words (cf. McCarthy, 1981). As a third example, consider Indo-European, which had a regular process of vowel change from /e/ to /o/ in root forms that made (among other things) past from present verb forms. (German strong verb alternations, and our present-day English irregular forms SIT-SAT, are fossilized remnants of this once regular system.) In each of these cases, the non-affixal form of morphological variation is known to be, or have been, a regular and productive process in the language.

The existence of such processes (and the complex paradigms found in "inflecting" languages like Latin and Russian) has led some linguists to question the general utility of the base plus affix morpheme model (e.g., Anderson, 1982; Matthews, 1974); it also calls into question the universal applicability of the base plus affix model of lexical access. How, then, will the lexicon be organized in those languages whose regular morphological processes do **not** involve separable, phonologically continuous morphemes? If continuity and separability are essential to lexical organization, related morphemes in these languages will behave like separate words. Conversely, if the related morphemes in these languages behave like regular, affixed forms, then it would seem that lexical structure can be adapted to accommodate the various kinds of regular grammatical alternations found in the world's languages.

The issue of shared lexical representations, and in particular the problems posed by irregular forms, has been primarily investigated using variants of the repetition priming paradigm. This technique relies on the fact that if a word is presented twice to a subject it is recognized more easily, and faster, on the second repetition. This has been demonstrated using the lexical decision task (e.g., Forbach, Stanners, & Hochhaus, 1974), and recognition of words in noise (e.g., Murrell & Morton, 1974). In its application to morphological investigation, two different morphologically related words are presented instead of repeating exactly the same word twice; the extent of facilitative influence from the first presentation (prime) to the second (target) is measured. In the logogen model, facilitation is explained by positing an activation threshold for each logogen; activation then lowers its threshold so that the logogen is more easily activated for some following period of time. Priming between morphologically related forms is explained by the effect of previous activation on the logogen they have in common. As noted above, this shared logogen necessarily corresponds to the base morpheme, that is, the continuous string of phonological elements they hold in common. All shared-entry models echo this concept (e.g., Manelis & Tharp, 1977). In terms of priming, therefore, any word that contains a base morpheme such as POURED or POURING should be as effective a prime for the word POUR as POUR itself would be.

Inflectionally affixed forms such as POURED have been shown to successfully prime a base form POUR. The magnitude of such priming is consistently less, but statistically equivalent to, that for POUR-POUR. This has been shown for lexical decision tasks (Fowler et al., 1985; Stanners et al., 1979), and for auditory word recognition in noise (Kempley & Morton, 1982). That is, identical variants prime themselves best, and regular, inflectionally related variants prime each other nearly as well. This result for the regular variants has been referred to as "full" priming and has been interpreted as evidence for shared entries between morphologically related forms. However, the situation is different for irregular inflections, such as irregular strong verbs (WOVE-WEAVE) (Kempley & Morton, 1982; Stanners et al., 1979) and suppletive (WORSE-BAD) forms (Kempley & Morton, 1982). Kempley and Morton found no priming between irregularly related inflectional forms, while Stanners et al. found "partial" priming, that is, a significant effect, but one that was significantly less than that for regularly related forms. Both papers interpreted these results to mean that irregular variants do not have enough phonological material in common to share a single lexical representation and are therefore represented as separate entities.

As Kempley and Morton point out, the failure of irregular inflectional variants to prime each other also argues against an alternative interpretation of the repetition priming effects. This alternative attributes the priming to the semantic relations between the prime and the target. Since semantically related (but morphologically unrelated) words can prime each other (e.g., Meyer & Schvaneveldt, 1976, for lexical decision), priming by inflectionally related words might be completely attributed to the semantic overlap between them, rather than to their sharing a common lexical entry or recognition unit. However, the semantic relations are equally tight in the case of irregular inflections, and thus, the failure to obtain priming with them argues against the strictly semantic alternative. In addition, Henderson, Wallis, and Knight (1984) have shown that priming between inflectionally related forms is significantly greater than priming between words related only semantically, when the two types are controlled for semantic closeness. Thus, repetition priming between inflectionally related forms seems to be revealing an independent layer of organization in the lexicon that is morphological in nature.

Results of priming between derivationally related variants presents a somewhat more confused picture. Stanners et al. (1979) looked at priming using derived adjectives (e.g., SELECTIVE-SELECT) and derived nouns (e.g., DECISION-DECIDE). Results showed a significant decrease in priming from derived words to their bases compared to the priming of base to itself. On the basis of this pattern, Stanners et al. suggest a compromise model of the lexicon, in which inflectionally related variants share entries while the reduced priming observed between irregularly and derivationally related variants is modeled as resulting from associations between separate entries. This conclusion is further supported by the fact that the degree of change in (phonological or orthographic) form between the derivation and the base did not seem to influence the degree of priming. This would be predicted if the connection between such items were solely semantic. Thus, Stanners et al. conclude that shared, separable phonological units entail shared entries in the lexicon for regular base and affixed variants, but not for words whose formal or functional relationship is irregular (as derivational relations may be).

The results of Fowler et al. (1985) call into question the strong form of Stanners et al.'s conclusion. They replicated the above experiments, but also ran conditions with a longer lag between prime and target, and found that with this longer lag, full priming could be observed from derived forms to their bases. (They attribute this to reduction of "episodic" aspects of priming.) While this result casts some doubt on the ability to decide whether priming in any particular case is the result of shared entries or associative connections, it is still the case that the derived forms behave differently than regular inflected forms. That is, there exist conditions under which priming among derived forms is not full, while regular inflections show full priming under the same conditions. Stanners et al. attribute the difference between derivation and inflection to the fact that derivational processes in English are less *regular* than derivational ones. Note that regularity is a complex concept; even when no change in base form is involved, English derivations may be less regular in two ways: a particular derivational affix may attach only to a subclass of eligible items (e.g., -IVE cannot be added to all verbs, or all transitive verbs), or there may be unpredictable changes in meaning associated with the derivations (as in FORM-FORMAL). Inflectionally related variants, in contrast, must have regularly, that is predictably, related meaning (Bybee, 1985). Stanners et al. argue that shared entries are involved only if two conditions are met: (1) there is a continuous, shared phonological sequence between the forms and (2) the relationship between the forms is regular.

It should be clear by now, however, that in English, the factors regularity and shared form are (partially) confounded. There are few (if any) cases in which a completely regular morphological process involves a change in form. For this reason, we chose to examine a language in which a change in form is associated with a completely regular process.

WELSH

Welsh is an Indo-European language, one of the few members of the Celtic branch to have survived into modern times. Other members of this group are Irish, Scots, Gaelic, and Breton. Welsh is spoken by approximately half a million people in the western portion of the British Isles (1981 census figures). The following discussion and examples of Welsh were culled from several sources: primarily Jones (1977), a grammar of colloquial Welsh, Awbery (1973, 1986) and Willis (1986), the latter three references being linguistic analyses of the mutation system in modern Welsh using native speakers. Each of these primarily treats the South Wales dialects. Information on the Northern dialects comes from the above works and Fynes-Clinton (1913).

In common with other Celtic languages, Welsh shows a phenomenon known as "initial mutation," in which the initial consonant of a word changes as a function of its lexical and syntactic context. There are three classes of such changes, traditionally called the SOFT, ASPIRATE, and NASAL mutations; of these, we will be concerned with only two, the SOFT and the ASPIRATE mutations. The SOFT mutation changes initial /p,t,k/ into /b,d,g/ and initial /b,d/ into /v,ð/ (/ð/ is the **th** sound in English **breathe**), while initial /g/ is deleted. (The SOFT mutation also affects other initial consonants, specifically /m/, the voiceless aspirated trill /r̥/, and the voiceless lateral fricative /ɬ/; however, we will be concerned only with /p,t,k,b,d,g/ in this paper.) The ASPIRATE mutation changes the initial consonants /p,t,k/ into /f,θ,x/ (/θ/ is the **th** sound in English **thing** and /x/ is the **ch** sound in German **Bach**); it does not affect initial /b,d,g/. These changes are summarized in Table 1.

Table 1

Initial consonant changes for mutating words in Welsh

Citation	Aspirate Mutation	Soft Mutation
p	f	b
t	θ	d
k	x	g
b	b	v
d	d	ð
g	g	-

Thus, for example, a word whose dictionary form is /pont/ 'bridge' will appear in a SOFT mutation context as /bont/ and in an ASPIRATE mutation context as /font/. A word whose dictionary form is /brawd/ 'brother' appears as /vrawd/ in a SOFT mutation context but stays as unchanged /brawd/ in an ASPIRATE mutation context.

It is important to note that, in terms of the phonology of Welsh, all the changes we are concerned with are phonemic, that is, they involve change from one phoneme to another (Awbery, 1986). (The phonemic consonants of Welsh are listed in Table 2.) While in theory this could result in neutralization of the difference between minimal pairs of words, such as /pis/ 'peas' and /bis/ 'finger', in practice this happens very rarely, partly because many phonemes (e.g., /v/, /θ/, and /ð/) appear word-initially in base forms primarily as the result of borrowing (Awbery, 1973), in which case the different phonological structure of borrowed words makes neutralization unlikely. In addition, the fact that the mutated and citation, or base, forms appear in very different linguistic contexts makes confusion unlikely.

Table 2

Phonemic consonants of Welsh (from Awbery, 1986)

	bilabial	labiodental	dental	alveolar	palatoalveolar	palatal	velar	uvular	glottal
voiceless stops	p			t			k		
voiced stops	b			d			g		
voiceless fricatives		f	θ	s,ɬ	ʃ			x	h
voiced fricatives		v	ð						
nasals	m			n			ŋ		
liquids				l,r					
glides	w					j			

Note: Awbery lists the phoneme /x/ as being uvular rather than velar. This is probably a matter of dialectal variation, as Jones (1926) lists it as velar. In any case, velar and uvular place of articulation are not contrastive in Welsh phonology.

Certain consonants are not subject to mutation. These nonmutating consonants include the consonants /f,θ,v,s,ʃ,r,l,h/ and the glides /w,j/ as well as the nasal /n/ (/ŋ/ occurs initially only as a result of the nasal mutation). Words beginning with these consonants do not change form, regardless of surrounding context; thus, /frind/ 'friend' remains /frind/ in all circumstances.

The mutation classes are differentiated from each other not only by the particular phonological changes they engender, but also by the different sets of linguistic circumstances that trigger them. The contexts that trigger mutation are predominantly lexical (i.e., occurring with a particular word), but also can be syntactic (i.e., occurring in a certain type of grammatical phrase or syntactic structure). The contexts have no obvious phonological, structural, or semantic commonality in modern Welsh. As a historical note, however, membership in the class of mutation-triggering contexts, which seems so unmotivated in the modern language, stems from developments in the 5th and 6th centuries, when early phonological rules affected the pronunciation of word-initial consonants according to the final sound of the preceding word; later, the preceding vowels and consonants ceased to be pronounced, but the custom of mutating initial consonants was preserved and even extended (Willis, 1986, p. 42).

Examples of lexical contexts in which consonants are obliged to undergo the SOFT mutation include grammatical function words such as /i/ 'to', as in /i vaŋgor/ 'to (the city of) Bangor'; and /faint/ 'how many', as in /faint ðaith/ 'how many came?' (from /daith/ 'he/she/it came'). In these cases, the preceding word itself is the trigger, that is, synonyms and homonyms do not have the same effect. The vast majority of the SOFT mutation contexts are of this type, as are all of the ASPIRATE mutation contexts (Awbery, 1973): for example, consonants undergo the ASPIRATE mutation after the word /a/ 'and', as in /te a xofi/ 'tea and coffee'; and after the word /tri/ 'three', as in /tri xant/ 'three hundred' (from /kant/ 'hundred').

The SOFT mutation is also used in sytactic contexts. For example, it occurs in adjectives following a feminine singular noun, as in/ə vasged bert/ 'the pretty basket' (from /basged/ 'basket' and /pert/ 'pretty'); in words used vocatively, as in /bore da, blant/ 'good morning, children' (from /bore/ 'good', /da/ 'morning' and /plant/ 'children'); in a direct object following a verb (with or without intervening material), as in /gweloð Tom gi/ 'Tom saw a dog' (from /gweld/ 'see' and /ki/ 'dog'); and in the subject of a sentence when an adverb or other material is inserted between the verb and subject (Welsh has verb-subject-object sentence order). In some cases, the application of mutation by itself carries functional or semantic load; for instance, according to Jones (1977, p. 335) "the Soft Mutation is the sole interrogative marker in spoken Welsh," as in /welest ti ve/ 'did you see him?' (from /gweld/ 'see' and /ti/, /ve/ 'you', 'him').

In spite of the current arbitrariness of the triggering contexts, mutation in Welsh is a regular and productive process, in the sense that the mutations apply (with minor exceptions) to all phonologically eligible words in the language. Moreover, new words and borrowings are mutated just as entrenched Welsh words would be. It is also obligatory: use of the mutations cannot be varied for stylistic reasons and native speakers are identified by the consistency and ease with which they use mutations. Indeed, it is as pervasive a presence in Welsh sentences as, for instance, the use of the suffix /-s/ in plurals and the present tense is in English: a typical Welsh sentence will contain at least one instance of mutation and often more. All changes in initial consonants due to mutation are reflected in the writing system, that is, *pont* is written *bont* in a mutation context.

The Welsh mutations present some interesting opportunities for testing the hypotheses about lexical organization discussed above. First, the relationship among the forms /pont/, /bont/, and /font/ is not an instance of affixation, analogous to English POUR-POURED, but rather an instance of internal modification analogous to the English HANG-HUNG case. (Although the three share a continuous unit, ONT, this unit is not meaningful by itself.) Thus, if a shared, continuous base form is required in order for words to share a lexical entry or recognition unit, then /bont/ or /font/ would be expected to prime /pont/ no more than HUNG primes HANG (i.e., partial priming in lexical decision, Stanners et al., 1979; and no priming in auditory word recognition, Kempley & Morton, 1982). On the other hand, if the key factor in lexical organization is paradigmatic regularity, then the regular Welsh variants /pont/ and /bont/ would be expected to prime one another as much as the regular English variants POUR and POURED prime each other.

The particular pattern of phonological changes in Welsh mutations provides us with Welsh-internal controls on the role of form change in priming. In particular, words beginning with voiced stops show no change in ASPIRATE mutation, although words beginning with voiceless stops do show such changes. Thus, we can compare, for example, how well variants in the ASPIRATE mutation prime BASE forms, when this difference is or is not accompanied by an actual change in phonological form. This kind of control is important, since without it, the effect of form change could only be assessed by comparing results across experiments involving different languages.

The primary goal of Experiments 1 and 2 is to examine the role of form change in priming using the Welsh mutations. Robust, or full, priming would indicate that sharing a lexical entry or recognition unit does not depend on shared, continuous surface sequences, as long as the variation is regular. In addition, the experiments were designed to allow evaluation of alternative lexical organizations for morphologically related words. One alternative is the model that Stanners et al. (1979) propose for derivations and irregular inflections in English. In this model, these irregularly related forms do *not* share lexical entries, but rather constitute separate, closely associated entries. In addition, Stanners et al. hypothesize that processing of irregular inflections (and derivations) requires activation of the entry for base form in memory, although this activation is less strong than if the base and related form shared the same entry (as they do for regular inflections). This hypothesis was motivated by their results showing that derivations and irregular inflections require a longer lexical decision time (when unprimed) than their associated base forms. The Stanners et al. model predicts an asymmetry in priming: since processing irregular inflections and derivations requires the activation of the base forms, we would expect measurable priming of the bases by such variants. Conversely, activation of the variants is presumably not required in order to process the base (otherwise there would be no accounting for the differences in unprimed decision time), and thus we would expect bases to prime these variants much less strongly. Experiment 1 was additionally designed, therefore, to look for evidence of this asymmetry, by examining both priming of mutations by base forms and priming of base forms by mutations. Such asymmetries, particularly if accompanied by overall weak priming effects, could be taken as evidence for a "separate entries" organization of the Welsh mutations.

Experiment 2 partially replicates Experiment 1, and is also designed to look for evidence of the "satellite" model of lexical organization proposed by Lukatela et al. (1980). This model, developed for Serbo-Croatian inflections, assigns a separate lexical entry to each member of an inflectional paradigm, but ties each one to a single core entry (the nominative singular form). Since the satellites are not related to each other, but only to the nucleus, we would expect less priming between two satellite entries than between the nucleus and a satellite. To the extent that this kind of organization is appropriate for Welsh mutations, we would expect the base form to be the nucleus, since it is the most common and representative of the possible variants: it is the form any Welsh speaker will give in response to the question "What's the word for X?," and the only form that can occur in isolation. Experiment 2 will allow us to test the predictions made by this kind of lexical organization by comparing priming between two mutations with priming between mutation and base forms.

Experiment 3 is designed to compare the overall recognition rate for words in different mutations. Of particular interest here is the fact that Welsh mutations specifically effect changes in the word initial segment. There is substantial evidence that beginnings of words have some special status in the lexicon (Browman, 1978; Fay & Cutler, 1977) and in word recognition (Cole & Jakimik, 1980; Marslen-Wilson & Welsh, 1978). Theories of word recognition model the importance of the initial portion of the word by assuming that recognition involves creating a "short list" of potential recognition candidates on the basis of the word-initial segments. Such a short list is either seen as being created on the fly by independent word recognition units operating in parallel (cf. Marslen-Wilson, 1984), or as inherent in how the lexical entries are organized in a mental dictionary (cf. Forster, 1976; Taft, 1979). For Welsh, such an organization might be expected to cause difficulties, particularly if the evidence from Experiments 1 and 2 points to a single lexical entry or recognition unit for all mutation variants. If there is a single entry for the Welsh forms /pont/, /bont/, and /font/, in which "short list" would it be found–the /p/-initial, /b/-initial, or /f/-initial lists? If it were entered with the /p/-initial forms (appropriate to its base), then this would suggest some added complexity in the process of recognizing the mutated forms. The recognition units for mutating words would have to be sensitive to the interaction of phonological form and syntactic environment. Experiment 3 looks for evidence of such complexity by comparing recognition of mutating words with control words that do not mutate.

GENERAL METHOD

An unusual aspect of the study described here is the fact that all three experiments were run simultaneously on the same set of speakers. This was done to make maximal use of a limited amount of time in Wales and a potentially limited number of Welsh-speaking subjects. Those design features and procedures the three experiments have in common are described here, while the aspects of design and methodology unique to each experiment are described in the separate experimental sections.

The experimental paradigm used was a variation of the auditory priming methodology found in Kempley and Morton (1982). In our experiments, as in theirs, subjects were first required to attend to a list of priming words, presented under optimal listening conditions, and then to identify, using written responses, a list of target words presented in noise. Each target word was associated with a particular prime-target condition, depending on the specific experimental design. Typical prime-target conditions include priming of a target by itself (the "identical" form in prime and target lists), priming of the target by the same word in a different form ("related" or "different" priming) and no priming ("none"). Subjects were divided into eight groups, with each group receiving a given target word in a different prime-target condition. The use of eight subject groups was dictated by the structure of Experiment 2, which compared eight prime-target conditions. In Experiments 1 and 3, both of which logically required only six subject groups. certain prime-target conditions were duplicated in order to fill out the stimulus tapes for subject groups 7 and 8.

For each experiment, the experimental treatments were distributed among the eight subject groups and the selected lexical items using a balanced fractional factorial design based on Latin Squares. For example, given 8 prime-target conditions and 8 words, subject group I received word A in condition 1, word B in condition 2, word C in condition 3, and so on, while subject group II had word B in condition 1, word C in condition 2, word D in condition 3, etc. Thus, each word occurred in a different experimental treatment for each group of subjects. The number of words used in any experiment was always an even multiple of the number of prime-target conditions defined by the experiment.

Stimulus Construction: Words

There were 72 words altogether, of which 24 were used in Experiment 1, 32 in Experiment 2, and 16 in Experiment 3. All except eight nonmutating words (used in Experiment 3) are regularly subject to SOFT and ASPIRATE mutations in Standard Welsh. The eight nonmutating words are unambiguously nonmutating; that is, there are no corresponding mutating words that share the same form with the nonmutating words in some mutation. Criteria for word selection were that the word be a monosyllabic or bisyllabic masculine noun with initial stress and end in a closed syllable. The number of words with a particular initial consonant in the base form (for mutating words /p,t,k,b,d,g/; for nonmutating words /f,x/) was balanced in each experiment. Equal numbers of monosyllabic and bisyllabic words were chosen. We tried to limit the list to frequently used words for commonplace objects or concepts, and to avoid words with limited areal (dialectal) distribution, or unstable gender identification, by checking with a native speaker. The complete set of words is listed in Appendix I.

Mutation frames

Although base (dictionary) forms may occur in isolation, Welsh speakers find mutated forms removed from their triggering context extremely strange. Accordingly, in the experiment all forms were always presented embedded in an appropriate syntactic frame. The SOFT and ASPIRATE mutation contexts consisted of the preceding possessive pronouns /ei/ 'his' and /ei/ 'her', plus a disambiguating postposition commonly used in colloquial speech. For the SOFT mutation, this was /o/, meaning 'he' or 'him'. For the ASPIRATE mutation, this was /hi/, meaning 'she' or 'her'. The BASE frame consisted of the preceding definite article /ə/ plus a postposition /əma/ meaning 'there', which when combined have a colloquial meaning 'that'. Since the definite article is a base form context for masculine (but not feminine) nouns, our experiments used only singular number masculine nouns. For example, presentation of the word /pen/ 'head' for the SOFT mutation occurred in the phrase /ei ben o/ 'his head', for the ASPIRATE mutation in the phrase /ei fen hi/ 'her head', and for the BASE form in the phrase /ə pen əma/ 'that head'.

Recording

Each word was recorded three times in each of these three frames by a native speaker of North Wales, in a sound-treated booth. The speaker was instructed to produce the words with identical intonation, speaking rate, and loudness (as much as possible), and was constantly monitored for signs of fatigue or change in these variables. To avoid encouraging systematic intonational differences between BASE and MUTATION forms, phrases were recorded in a randomized list.

The most intelligible and normal-sounding example of each phrase, as determined by a second Welsh speaker, was then digitized at 20 Khz using the PCM system at Haskins Laboratories.

Noise levels

Each target stimulus was presented in noise. The noise employed was signal-correlated, that is, noise manufactured from the corresponding stimulus by reversing the polarity of digitized signal points in a random pattern. Unlike white noise, which is known to differentially degrade recognition of different segment types, noise made by this technique, which matches the amplitude envelope of the noise to that of the signal, may affect different segments and words more nearly equally.

Addition of a given level of noise affects the recognizability of different words to different degrees, particularly as a function of word frequency (Howes, 1957; Pollack, 1963; Rubenstein & Pollack, 1963). In addition, individual subject differences in hearing or word recognition strategy may also affect word recognition under noisy conditions. Thus in analyzing the outcome of recognition tests in noise, it is desirable to control to some extent the effects of individual word and subject variation. In this study, we drew on previous experience with Welsh pilot studies to determine an anchor noise level at which subjects averaged 50% recognition. This anchor level differed for the monosyllabic and bisyllabic stimulus words. In addition, we attempted to calibrate noise levels for individual subjects. To this end, we administered a noise-level sensitivity pretest to each subject before commencing the experiment proper. The pretest session was designed to choose a noise level for each subject that would yield an average 40% recognition rate. [Details of the process by which noise levels were chosen can be found in Appendix II.]

Stimulus orders

For each subject group, the target words (in the appropriate mutation frame) for all three experiments were combined into a single target list; similarly, the priming words (in their appropriate frames) were combined in a single priming list. Extra lead-in items (also in appropriate frames) were then added to the priming and target lists. Both the priming list and the target list were randomized separately from each other, and separately for each subject group, using a block randomization procedure that guaranteed that each quartile of each list contained the same number of words from each experiment. In line with Kempley and Morton's (1982) procedure, the priming list was repeated twice in succession on the test tape. The priming list was differently randomized and contained different lead-in items for each presentation.

Thus, a test session for any one subject included, in order, (1) the noise level pretest described above, (2) the priming items assigned to his or her subject group, presented twice, and (3) the target list for that group. Eight separate test tapes were made, one for each group of seven subjects. One channel of each tape held the noise-free recordings of pretest, priming, and target stimuli; the second channel carried synchronized correlated-noise images of the pretest and target stimuli, at an amplitude equal to the amplitude of the words. (For the priming session, this channel carried only silence.) Noise and signal channels were then mixed (using a circuit that allowed variable attenuation of the two channels) at a signal-to-noise ratio determined by the subject's response on the pretest, as detailed in Appendix II.

Procedure

Subjects were instructed, at the beginning of the test session, to expect words to occur in only one of the three phrases described above: /ə ___ əma/, /ei ___ o/, and /ei ___ hi/. For the pretest, subjects were instructed to write down the entire phrase as they heard it. For the two presentations of the priming list, subjects were told that they would hear a number of different mutated and base-form nouns embedded within one of the three frames, and were requested to provide a written estimate of the frequency with which the central noun in the phrase was used in everyday conversation, on a scale of 0 to 100. They were provided with an answer sheet with appropriate scales. This task ensured the subjects' attention to each word on the priming lists, and provided two estimates of token frequency for each lexical item in the experimental corpus. The estimates of frequency are reported in Appendix III, and further discussed in the relevant experiments; in general, while there were frequency differences among the lexical items, these were distributed evenly among the various conditions, with the primary exception of slightly higher frequency judgments for items in their base form than in their mutated forms (base mean = 59, mutation mean = 52). Four seconds of silence intervened between the presentation of each item in the priming lists.

Approximately 5 minutes after the second presentation of the priming list, the subjects were presented with the list of target items. In the interim, they were instructed that they were now to write down the entire phrase as they heard it, just as in the pretest case. After five lead-in items (different from the priming lead-ins but identical for each subject group), the list of target items was played, with 8 seconds between each phrase.

Each subject was tested individually in a quiet room. The total priming list for any subject group (from all three experiments plus lead-ins) consisted of 53 items, the total target list of 77 items. The entire test session for any one subject lasted approximately 35 minutes.

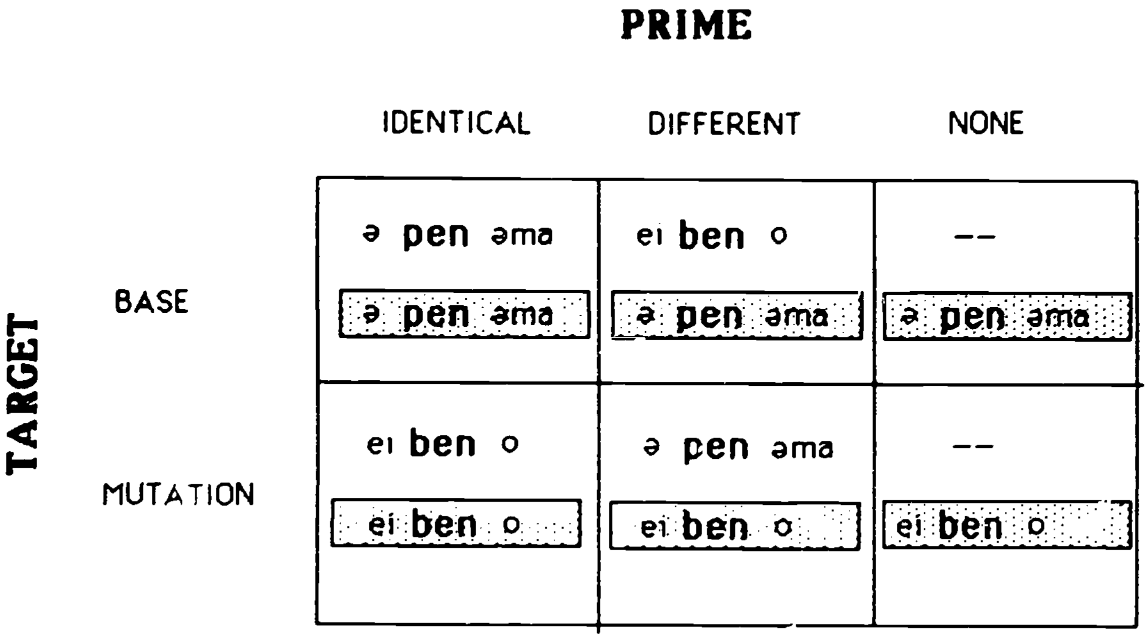
Subjects

Subjects were 56 native speakers of Welsh. Forty-seven subjects were students between the ages of 17 and 35 (45 at the University of North Wales in Bangor, 2 at Cambridge University in Cambridge, England), while 9 were older members of the Welsh speaking community in Cambridge, England. All had early schooling in Welsh, and many had been educated in Welsh up to university level.

All were fully bilingual in Welsh and English. Subjects were assigned to groups in a random fashion.

EXPERIMENT 1

As discussed in the introduction, the primary goal of Experiment 1 was to examine whether full priming is found between base and mutation forms of Welsh nouns. An affirmative answer would indicate that a shared, continuous phonological base form is not required for the variants to share a lexical entry. A second goal of the experiment was to test for possible asymmetries in the direction of priming. As argued above, the Stanners et al. (1979) model for irregular morphological relations predicts that mutations should prime base forms much more strongly than base forms prime mutations. Experiment 1 was therefore designed to show any differential effects of priming between BASE and MUTATED forms.



		PRIME		
		IDENTICAL	DIFFERENT	NONE
TARGET	BASE	ə pen əma ə pen əma	ei ben o ə pen əma	-- ə pen əma
	MUTATION	ei ben o ei ben o	ə pen əma ei ben o	-- ei ben o

Figure 1. Example of experimental design for Experiment 1. Within each box, the top item is the prime, the bottom item the target.

The design for Experiment 1 included three type of primes: IDENTICAL (prime and target were the same lexical item in the same mutation), DIFFERENT (prime and target were different forms of the same lexical item, one a base, the other a mutation), and NONE (no prime at all). In addition, for each prime type, targets were of two types: MUTATION and BASE. This yielded the six prime-target conditions shown in Figure 1, exemplified using the lexical item /pen/ ('head'). The six conditions were replicated for each of two mutation classes: the SOFT and ASPIRATE mutations.

A set of 12 words (6 monosyllabic, 6 bisyllabic; each group of 6 including one each of word-initial /p,t,k,b,d,g/) was assigned to each of these replications. The distribution of the 12 words across the six prime-target conditions varied over six subject groups according to the modified Latin Square design described earlier. In order to conform to the overall study design of eight subject groups, the prime-target patterns for subject groups 1 and 2 were repeated for subject groups 7 and 8.

Results

Subject responses were scored as correct if the entire stimulus, including mutation and surrounding frame, was correctly transcribed. Results are shown graphically in Figure 2. The top panel of the figure shows the percentage correct recognition for the six prime-target conditions for the SOFT mutation replication, the bottom panel shows the same for the ASPIRATE mutation.

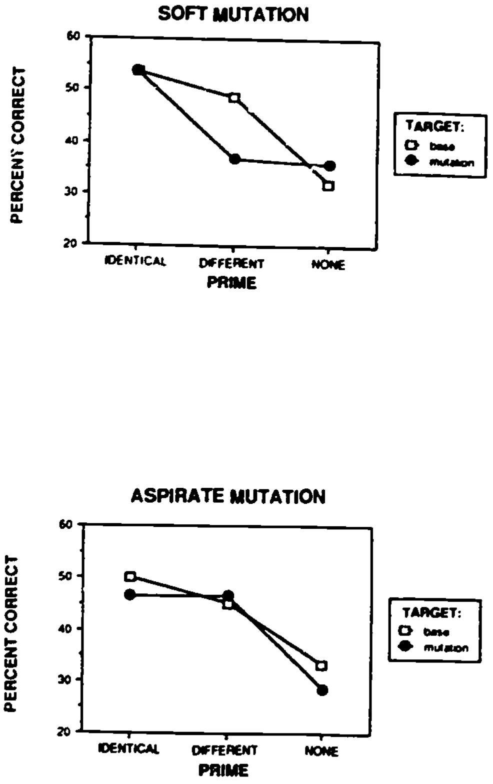


Figure 2. Percentage of correct recognition for each prime condition (Experiment 1).

The figures show that IDENTICAL priming yields the highest recognition rate, followed closely by DIFFERENT priming, which in turn shows better recognition than the NONE condition.

Two separate analyses of variance were performed (in BMDP4V), one using the variability of individual subjects and one using item means to form the error term. Factors were SYLLABLE (ONE vs. TWO), MUTATION (SOFT vs. ASPIRATE), TARGET (BASE vs. MUTATION form), and PRIME (IDENTICAL vs. DIFFERENT vs. NONE). Those factors involving different sets of words (SYLLABLE and MUTATION) were treated as within-group, or repeated measures, factors in the subject analyses and as between-group, or grouping, factors in the item analyses. There were no systematic differences between subject groups. Note that only data from subject groups 1 through 6 are reported here (the results did not differ when data from subject groups 7 and 8 replaced data from groups 1 and 2 in the analyses).

Only the main effect for PRIME was significant in both the subjects and items analyses: subjects, $F(2,82) = 16.78, p < .0001$; items, $F(2,40) = 12.53, p < .001$. Planned comparisons between the IDENTICAL and DIFFERENT conditions (the test for full vs. partial priming) showed the two to be marginally statistically different in the subjects analysis, $F(1,41) = 6.70, p = .013$, but not in the items analysis, $F(1,20) = 3.30, p = .084$. However, the DIFFERENT condition was reliably different from the NONE condition in both analyses: subjects, $F(1,41) = 11.49, p < .005$; items, $F(1,20) = 10.55, p < .005$. The effect of SYLLABLE was significant in the subjects analysis, but not the items: subjects, $F(1,41) = 23.65, p < .0001$; items, $F(1,20) = 1.71, p = .205$. The recognition rate was higher for monosyllables (monosyllabic mean = 49%, bisyllabic mean =

36%). This may indicate that the expected recognition advantage of bisyllabic words was overcompensated by the difference in noise levels; it might also reflect frequency differences between the monosyllables and bisyllables, since these differences are in the right direction, although not significant. No other main effects, nor any interactions, were significant. Note that the lack of a TARGET effect indicates that the small (but significant) difference in estimated frequency between base and mutated forms was not reflected in the recognition accuracy.

In Kempley and Morton (1982), a distinction was drawn between the results as scored by strict criteria and the results as scored by more lenient criteria in which a response containing the correct lexical item was considered correct, regardless of whether it was the correct morphological variant. In their experiment, this amounted to a distinction between "whole word" scoring (transcription of the stimulus as given) and "morphemic" scoring (responding with some variant of the given stimulus). In our experiment, there were several possibilities for "morphemic scoring," ranging through (1) responding with the right frame but wrong mutation; (2) responding with the wrong frame but right mutation; and (3) responding with the wrong frame and mutation. Informal analysis of the incidence of these response types revealed no systematic pattern; in addition, statistical reanalysis of the data obtained by scoring all three as correct responses showed the same results as the strict scoring procedure reported above.

Discussion

The most important result in this experiment is the significant effect of priming between BASE and MUTATED forms. Clearly, unlike Kempley and Morton's (1982) result for English, priming does occur between morphological variants that are not eligible to share a logogen. This, in turn, supports the view that the kind of intimate lexical sharing that is indexed by repetition priming does not require that the forms share a unique, continuous phonological sequence. Rather, the lexical organization among variant forms of a given lexical item in Welsh seems to be more akin to the organization of regular paradigms in English than to the organization of irregular English pairs.

This conclusion is also compatible with another major result of the experiment, namely, the failure to find any asymmetries in priming. The lack of a significant effect for TARGET and the lack of any interactions between PRIME and TARGET indicates that BASE and MUTATED forms are equally effective both as primes and as targets. Note that this result would not be expected under a theory of lexical structure that assumes that the variant forms have distinct lexical entries and that one of the forms is always accessed through the other, dominant form (a principal entries hypothesis). Rather, the current results argue for symmetrical lexical access.

It might be objected that this interpretation is not valid because it is based on partial rather than full priming, while most arguments for shared entries are based on full priming between related forms (cf. Stanners et al., 1979). However, we should point out that the difference between the IDENTICAL and DIFFERENT conditions only just approaches significance in our results, even in a powerful, a priori test. Further, it has been noted (Fowler et al., 1985) that full priming results in the literature have tended to show numerically, if not statistically, weaker priming for the DIFFERENT condition. That is, priming between related forms may be consistently less than priming between identical forms, but at a level that only sometimes reaches statistical significance.

It should be noted that, even if this difference between conditions is real, the differences are not necessarily due to the phonological differences between the prime and target words. The prime and target in the DIFFERENT condition also differ in their morphosyntactic (and semantic) mutation frames. The question of form vs. frame can be pursued, in a preliminary way, by breaking down the data for the ASPIRATE mutation further. Since words whose BASE forms begin with voiced /b,d,g/ do not undergo changes in the aspirate mutation, but words beginning with voiceless /p,t,k/ do undergo change to the corresponding fricatives, comparison of these two sets of stimuli could shed some light on the importance of phonological form change in contributing to the difference between IDENTICAL and DIFFERENT conditions. In fact, words with voiceless base forms show a decrement from 45.2% in the IDENTICAL condition to 39.2% in the DIFFERENT condition, while words with voiced base forms show 51.2% in the IDENTICAL condition and 52.3% in the DIFFERENT condition. While this difference suggests that phonological form plays a role, the design of Experiment 1 did not allow this difference to be properly evaluated statistically. In Experiment 2, therefore, the underlying voicing status of the target was explicitly set up as a factor in the design.

EXPERIMENT 2

Experiment 2 extends the findings of Experiment 1 to additional types of prime-target pairings, and attempts to explore more carefully the role of phonetic factors and form changes in priming. In Experiment 1, the issue of symmetrical vs. asymmetrical relations among forms of a given lexical item was tested using BASE-MUTATION and MUTATION-BASE prime-target pairs. Results suggested that the relations between the forms is symmetrical. It is possible, however, that whatever form of organization links the base form with the different mutation forms (and vice versa) does not operate between mutated forms. This would be the case, for example, in a "satellite" form of organization (e.g, Lukatela et al., 1980). In such a model, the base form would still be the central element of a cluster whose peripheral members remain unconnected with each other. In Experiment 2, therefore, priming between related forms was further explored by including conditions in which an item in one mutation form (SOFT or ASPIRATE) served as a prime for that item in the other mutation form.

The evidence for symmetrical relations in Experiment 1 lies in the absence of a significant TARGET effect, or of any significant interactions involving TARGET. However, as is clear in Figure 2, for the SOFT mutation there is a difference between BASE and MUTATION targets in the DIFFERENT priming condition (49% vs. 37%), while for the ASPIRATE mutation, such a difference is not found (45% vs. 46%). It is not clear whether this difference between the mutations is reliable (as noted above, interactions were not significant), or how to interpret it, if it should turn out to be reliable. It could reflect slight but genuine differences in lexical structure between the two mutations (i.e., relations are more symmetrical in the case of ASPIRATE mutation). Alternatively, it might be due to the fact that the SOFT and ASPIRATE mutation data come from separate sets of lexical items in Experiment 1. Or it may be the case that phonetic differences in the discriminability of consonants produced by the two mutations are, in fact, obscuring an underlying similarity between the two mutation classes. Experiment 2 was designed so that the same lexical items occurred in both mutation classes; and the experiment was set up to allow the phonetic effects of the target on recognition to be evaluated.

The experiment was designed as follows. All targets were presented in mutation form (either SOFT or ASPIRATE mutations), and they were primed under one of four conditions: IDENTICAL (same mutation form used for prime as for target), BASE (BASE form used as prime), OTHER (ASPIRATE form used as prime if target was SOFT and vice versa), and NONE.

The targets themselves were divided into four groups by two additional experimental factors. Half of the targets used in the experiment, which we will refer to as BITEMS, had base forms beginning with one of the voiced stop consonants (/b,d,g/), while the other half of the targets (PITEMS) had base forms beginning with one of the voiceless stops (/p,t,k/). Note (with the aid of Table 1) that the PITEMS undergo phonetic changes in both the SOFT and ASPIRATE mutations, while the BITEMS undergo changes only in the SOFT mutation.

The second factor on which targets varied was the mutation form in which the target was presented. However, rather than code the mutation form as SOFT vs. ASPIRATE, the target was coded according to the *phonetic* effect of the mutation, because of our interest in factoring out possible phonetic effects on recognition. Looking again at the consonant changes wrought by the two mutations (illustrated in Table 1), we see that for words beginning with stop consonants, there is one mutation context in which the phonetic feature of manner of articulation is changed from stop to fricative, and one mutation context in which the stop remains a stop. For BITEMS, this change from stop to fricative is induced by the SOFT mutation, while for PITEMS it is the ASPIRATE mutation that produces this change. Thus, the target is characterized according to the MANNER of its initial consonant—STOP vs. FRIC. Note that vowel-initial targets derived from /g/-initial base forms (by the SOFT mutation that derives /v/ from /b/ and /ð/ from /d/) are characterized as FRIC. This was done for the sake of symmetry in experimental design.

The four PRIME conditions were crossed with the four target conditions (2 BP X 2 MANNER), giving a total of 16 prime-target conditions. These are illustrated in Table 3, using the items /pen/ 'head' and /beic/ 'bicycle' as examples. Two sets of eight monosyllabic words were chosen, one PITEM set containing two words each beginning with /p/ and /k/, plus four words beginning with /t/, and one BITEM set containing two words each beginning with /b/ and /g/ plus four words beginning with /d/. Each set was distributed over the eight prime-target combinations appropriate to that item and over the eight subject groups as described above in Experiment 1. Each word appeared in each of the eight prime-target condition across different subject groups, and no subject heard the same words in more than one prime-target condition. Also as in Experiment 1, this structure was duplicated with a set of bisyllabic words.

Results

Subjects' responses were scored by strict and lenient ("morphemic") criteria, and repeated measure analyses of variance were carried out by subject and by item as described in Experiment 1. As before, results were the same by either method of scoring; only the results scored by strict criteria are reported below. Figure 3 shows the percent recognition for each of the 16 prime-target conditions. Lines connect the points for a given target type—PITEMS are connected by solid lines, BITEMS by textured lines. STOPS are represented by solid squares, and FRICS by open squares.

Table 3

Design of experiment 2

Within each box, the top item is prime, the bottom is target.

		IDENTICAL	BASE	OTHER	NONE
BITEMS	STOP	ei beic hi ei beic hi	ə beic əma ei beic hi	ei veic o ei beic hi	— ei beic hi
	FRIC	ei veic o ei veic o	ə beic əma ei veic o	ei beic hi ei veic o	— ei veic o
PITEMS	STOP	ei ben o ei ben o	ə pen əma ei ben o	ei fen hi ei ben o	— ei ben o
	FRIC	ei fen hi ei fen hi	ə pen əma ei fen hi	ei ben o ei fen hi	— ei fen hi

Factors for the analysis of variance included MANNER (STOP vs. FRIC), BP (BITEM vs. PITEM), PRIME (IDENTICAL, BASE, OTHER, and NONE), and SYLLABLE (ONE vs. TWO). As in Experiment 1, PRIME was the only factor that was significant in both subject and item analyses: subjects, $F(3,165) = 15.12, p < .0001$; items, $F(3,84) = 13.82, p < .0001$. As expected, recognition was highest in the IDENTICAL condition and lowest in the unprimed (NONE) condition, while priming between morphological variants (the BASE and OTHER-mutation priming conditions) showed intermediate, and approximately equal, levels. Contrasts showed no significant difference between these two intermediate conditions: subjects, $F(1,55) = .36, p = .55$; items, $F(1,28) = .21, p = .65$. To test for full vs. partial priming, we contrasted the IDENTICAL condition against the combined BASE and OTHER (non-IDENTICAL) conditions. This contrast was significant in both analyses: subjects, $F(1,55) = 10.16, p < .005$; items, $F(1,28) = 11.68, p < .005$. Finally, these latter two conditions contrasted significantly with the NONE condition: subjects, $F(1,55) = 19.36, p < .0001$; items, $F(1,28) = 25.87, p < .0001$.

Returning to the main analysis, overall differences in the recognizability of targets, as a function of the target factors (BP and MANNER), are visible in the separation between means corresponding to the four categories of targets (shown with connected lines in Figure 3). While MANNER itself (FRIC vs. STOP) was not a significant main effect, there was a significant effect for BP (BITEMS vs. PITEMS) in the subjects analysis, but not the items analysis, as well as a significant interaction of BP X MANNER (BP: subjects, $F(1,55) = 20.66, p < .0001$;

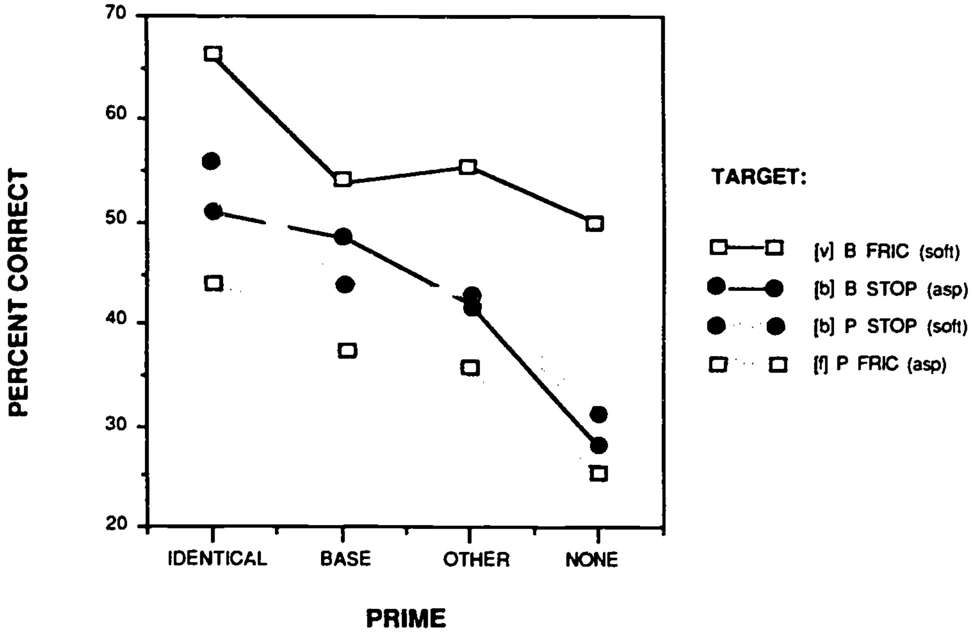


Figure 3. Percentage of correct recognition for each prime condition (Experiment 2).

items, $F(1,28) = 2.37, p = .13$; BP X MANNER: subjects, $F(1,55) = 19.46, p < .0001$; items, $F(1,28) = 3.8, p = .06$). The pattern of results visible in Figure 3, that is, the wide separation of lines for the voiced and voiceless FRIC targets and the clustering of lines for STOP targets midway between them, suggests that the significance of the BP factor could be entirely attributable to the MANNER X BP interaction. This would mean that the data can be explained phonetically by positing a recognition advantage for voiced fricative targets, relative to the stop baseline recognition, and a similar disadvantage for voiceless fricative targets. This interpretation is supported by the fact that an analysis of simple main effects shows that the BP factor is significant for FRIC targets, but not for STOP targets, (FRIC: $F(1,55) = 30.92, p < .0001$; STOP: $F(1,55) = .17, p = .69$; analysis done for subjects only, since the overall effects were only significant there). Alternatively, it is possible that, rather than being a function of the phonetic factors, the data reflect a basic difference in recognizability between the two mutations. (Note this is not reflected in the estimated frequencies of the items in the two mutations, which are not significantly different: 54 vs. 52.) Reanalysis of the data by the factor MUTATION (targets derived by the SOFT vs. targets derived by the ASPIRATE mutation) in fact revealed a significant difference between the two. All other results, including a significant BP effect, were the same in this reanalysis; there was, however, no BP x MUTATION interaction. From this reanalysis, it would appear that the two mutations have different baseline recognition rates, and that there is also some difference between recognition rates for BITEMS and PITEMS. At this point, there is insufficient evidence to decide between the alternative analyses. Experiment 3 will, however, be useful in this regard. In any case, none of these issues appear to affect priming; relationships between priming conditions is consistent for all categories of targets.

As in Experiment 1, the SYLLABLE effect was significant in the subjects but not items analysis: subjects, $F(1,55) = 7.97, p < .01$; items, $F(1,28) = .95, p = .34$. The monosyllabic recognition mean was 41%, the bisyllabic mean 47%. No other effects or interactions reached significance. There was a very small, nonsignificant difference, in the right direction, in the estimated frequencies of the monosyllables and bisyllables.

Experiment 2 was also set up to allow us to compare priming in BASE prime-target pairs that involve an actual change in initial consonant with those BASE prime-target pairs that do not actually show any consonant change. The relevant comparisons here involve BITEMS vs. PITEMS for the STOP manner, in which both BITEMS and PITEMS have targets that are mutations (BITEMS are ASP, PITEMS are SOFT) and that begin with /b,d,g/. In the IDENTICAL condition both sets of items are primed by targets beginning with /b,d,g/. However, in the BASE condition, primes for the BITEMS have the same initial consonants as the targets (e.g., /ə beic əma/ priming /ei beic hi/ in Table 3), while for PITEMS, the primes differ by having voiceless initial stops (e.g., /ə pen əma/ priming /ei ben o/ in Table 3). If form change *per se* between target and prime was a major determinant of strength of priming, then we would expect recognition to decrease from IDENTICAL to BASE conditions in the case of the PITEMS, but we would expect much less decrement in the case of the BITEMS. The points in Figure 3 are indeed in the right direction. However, a separate analysis of variance performed on these four cells alone revealed no interaction, and indeed, no significant main effect of PRIME, that is, IDENTICAL vs. BASE (PRIME: subjects, $F(1,55) = 1.91, p = .172$; items, $F(1,28) = 1.64, p = .211$; PRIME X BP: subjects, $F(1,55) = 1.04, p = .312$; items, $F(1,28) = .64, p = .43$).

Discussion

The apparent equivalence of the BASE and OTHER conditions, like the lack of asymmetries found in Experiment 1, argues against a kind of "satellite" lexical organization for Welsh mutations in which all the variant forms are related to the base, but not to each other. Once again, the results from Welsh suggest that the relationship among mutation variants is quite close, and there is no evidence to compel us to the view that each mutation variant must have its own independent lexical entry.

In Experiment 2, the difference between IDENTICAL and non-IDENTICAL (BASE plus OTHER) priming conditions reached significance (corroborating a trend found in Experiment 1). This suggests that priming between mutation variants, or between base and mutations, is only partial, rather than full. In this sense, the results for Welsh differ from those found for regular morphological processes in English. However, the priming difference does *not* seem to stem from the differences in phonological form between the lexical item in different mutations. As the separate smaller ANOVA on the BASE condition showed, indications of a difference in pattern for those lexical items that do show form changes (PITEMS) and those that do not show form changes (BITEMS) were not statistically significant. Thus, the fact that priming in the BASE and OTHER conditions is not full must be attributed to other differences between the conditions—for example, to syntactic or semantic differences between the mutation frames, or to abstract membership of an item in a particular mutation class. The fact that priming between items that show form change can be, statistically, as strong as between items that do not show such a change argues against defining logogens, or any notion of lexical units, strictly in terms of the criterion of unique shared physical sequences.

EXPERIMENT 3

In Experiment 3, we turn away from examining prime-target relationships, and instead look for differences in average recognition rates for different morphological variants. There was a suggestion of such differences in Experiment 2, where one interpretation of the results of the target manipulations was that there are differences in recognizability between the ASPIRATE and SOFT mutations. As discussed in the introduction, differences in recognizability among mutation variants are potentially quite interesting because of the important role initial segments play in lexical access (e.g., Cole & Jakimik, 1980; Taft, 1979). Since the evidence from Experiments 1 and 2 points to mutation and base variants sharing a single lexical entry, this single entry would have to be accessed by inputs varying in their initial consonants. If this kind of flexibility were problematic for the word recognition system, then we would expect to find some evidence of this difficulty in different recognition rates for the different morphological variants. In addition, we might expect mutating lexical items to be more difficult to recognize than words whose initial consonant is invariant across mutation environments (e.g., Welsh words whose base forms begin with fricatives).

In this experiment, only the unprimed (NONE) and IDENTICAL prime-target conditions were employed. One comparison of interest to us was the contrast between the behavior of the BASE, ASPIRATE, and SOFT forms. A second comparison of interest was the difference between MUTATING and NONMUTATING words, both in terms of overall recognition rate and also in terms of similarity of behavior in the different mutation contexts. If the MUTATING and NONMUTATING words display the same pattern of behavior in the different mutation contexts, any difference in recognizability between the morphological variants can not be attributed to difficulties associated with a variable initial consonant. Rather, it must be attributed to some aspect of the different mutation contexts themselves.

Thus, given two prime conditions and three morphological variants, there were 6 prime-target conditions tested: BASE primed by BASE, SOFT MUTATION primed by SOFT MUTATION, ASPIRATE MUTATION primed by ASPIRATE MUTATION, and all three forms unprimed (NONE). Two sets of eight items were employed, MUTATING and NONMUTATING. The NONMUTATING words were chosen to begin with either /f/ or /x/, segments that are also the reflex of an initial /p/ or /k/ in the ASPIRATE mutating context. The MUTATING words began with /p,k,b,g/. The items were distributed across the six prime-target conditions so that for each subject, two items were presented in each of the BASE conditions (BASE-BASE and BASE-NONE), and one item in each of the other four conditions. This allowed us to take advantage of all eight subject groups. As in the other experiments, the distribution of items over conditions and subject groups was varied according to a modified Latin square.

Results

Figure 4 shows the percent correct recognition for each of the prime-target conditions, separately for the MUTATING and NONMUTATING words. Again, repeated measures analyses of variance were carried out by subject and by item as described in Experiment 1. The results reported below reflect strict scoring criteria; as in the two previous experiments, "morphemic" scoring results were parallel. Factors were WORDTYPE (MUTATING vs. NONMUTATING), TARGET (BASE, SOFT MUTATION or ASPIRATE MUTATION) and PRIME (IDENTICAL vs. NONE). The BASE level of the TARGET factor was represented by the mean (per subject

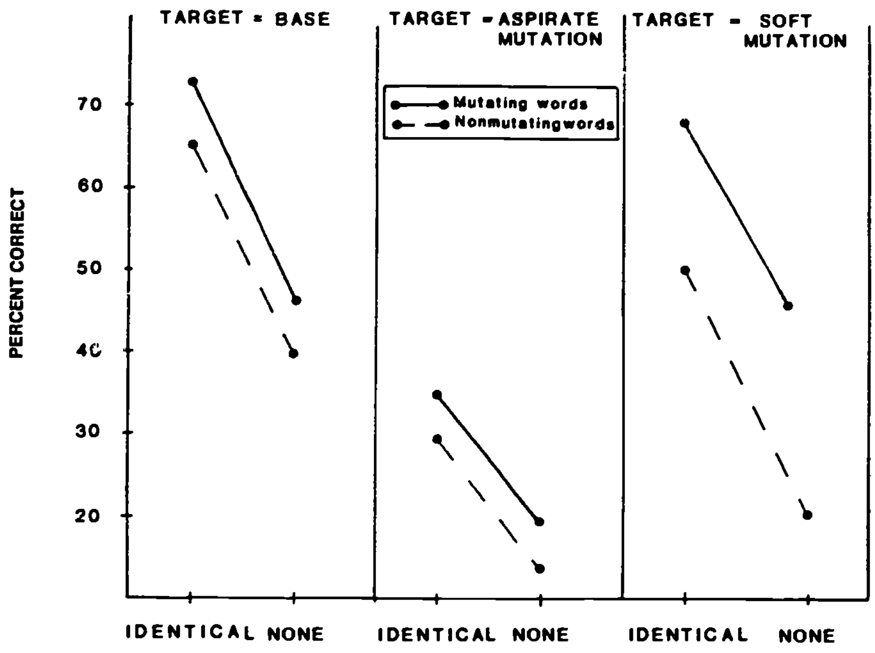


Figure 4. Percentage of correct recognition for each prime condition (Experiment 3).

or per item) of the two duplicated BASE-BASE and NONE-BASE prime conditions. (Separate analyses of the data using one or the other set of conditions showed parallel results.)

As expected, the IDENTICAL priming condition showed a significantly higher recognition rate than the NONE condition: subjects, $F(1,55) = 57.39, p < .0001$; items, $F(1,14) = 38.33, p < .0001$. The significant main effect for TARGET (subjects, $F(2,110) = 40.03, p < .0001$; items, $F(2,28) = 30.62, p < .0001$) was primarily due to the lower recognizability of the ASPIRATE mutation. The difference between SOFT and ASPIRATE mutations was significant in a planned comparison: subjects, $F(1,55) = 29.43, p < .0001$; items, $F(1,28) = 34.51, p < .0001$; however, the BASE and SOFT mutations also differed marginally in recognition: subjects, $F(1,55) = 7.81, p < .01$; items, $F(1,14) = 6.21, p = .0259$. The TARGET effect was similar and consistent for both MUTATING and NONMUTATING words, as shown in the non-significant TARGET x WORDTYPE interaction. WORDTYPE was significant in the subject analysis only: subjects, $F(1,55) = 21.54, p < .0001$; items, $F(1.14) = 1.24, p = .28$. No interactions were significant.

Discussion

The comparison of recognizability for NONMUTATING and MUTATING words fails to support the notion that variation in the initial consonant of a word produces difficulties in recognition. Specifically, it is not the case that NONMUTATING words were better recognized than MUTATING words, as might have been expected if variability in initial consonants decreases recognition. In fact, the NONMUTATING words were recognized (nonsignificantly) less well than the MUTATING words. This might be due to their lower estimated frequency; however, the frequency difference is the same as that noted between the monosyllables and bisyllables in Experiment 1, which had only a marginal effect on recognition accuracy.

The behavior of the mutations also fails to support the role of initial consonant variability in recognition difficulties. While overall recognizability is greater for the SOFT MUTATION and BASE than for the ASPIRATE MUTATION, this pattern holds true for both MUTATING and NONMUTATING words. Since the NONMUTATING words have the same initial consonants in BASE, SOFT MUTATION, and ASPIRATE MUTATION contexts, their adherence to this pattern indicates that any explanation must make reference to what the MUTATING and NONMUTATING words have in common: the frame in which the word occurs. It is not clear from this experiment, however, which properties of the frame are relevant. For example, we would like to know whether the effect is specific to the morphemes for 'his', /o/, and 'her', /hi/, or whether the same pattern would be obtained with other contexts for the ASPIRATE or SOFT mutations. In the former case, the effect would be specific to a particular morphological category; in the latter, it would be specific to an abstract mutation class (abstract in the sense that it doesn't always result in any phonological alterations).

Suggestions of the superiority of the SOFT MUTATION and BASE forms over the ASPIRATE MUTATION can be found in Experiments 1 and 2, as well. For Experiment 1, if we examine the ASPIRATE MUTATION results in Figure 2, we see that for both the IDENTICAL and NONE conditions, the BASE targets are superior to MUTATION targets. However, for the SOFT MUTATION results, this is not the case. In the IDENTICAL condition, BASE and MUTATION are equivalent, while for the NONE condition, the MUTATION form is actually better recognized. Thus, we see in Experiment 1 the same basic pattern that we have documented in Experiment 3.

As mentioned above, in Experiment 2 reanalysis of the data by mutation category showed a significantly higher recognition rate for the SOFT MUTATION targets than for the ASPIRATE MUTATION targets. In analyzing those results, we could not choose between an analysis based on inherent differences between the two mutations and an analysis based on acoustic phonetic properties. In Experiment 2 a large portion of the difference between targets derived by SOFT and ASPIRATE MUTATIONS could be ascribed to the higher rate of recognition associated with initial voiced fricative consonants, and the much lower rate associated with initial voiceless fricatives. However, the NONMUTATING words in Experiment 3 provide the crucial evidence—they show that there are systematic differences between SOFT and ASPIRATE MUTATION contexts even when there are no phonetic differences between the words in these contexts. This suggests that it would make more sense to interpret the results of Experiment 2 in the same way. This interpretation would be supported by other considerations as well. The alternative analysis rests on the assumption that voiced fricatives are more recognizable than stops, which are, in turn, more recognizable than voiceless fricatives. However, phonetic studies (at least for English) have tended to find a rather different order of intelligibility (e.g., Goldstein, 1977; Miller & Nicely, 1955)—stops, followed by voiceless fricatives, followed by voiced fricatives.

It is interesting to note that the apparent advantage of the SOFT over the ASPIRATE MUTATION is in accord with the linguistic facts. Awbery (1986) has shown that in many areas of Wales the phonetic changes associated with the SOFT MUTATION are replacing ASPIRATE MUTATION changes in ASPIRATE mutating environments. It is possible that what we see here is a reflection of the decreasing probability, over the language as a whole, that the ASPIRATE

MUTATION will be used. Before making this kind of conclusion, however, it would be important to know whether the effect generalizes to other mutation contexts, or is specific to 'his' and 'her'. It is unlikely to be a reflection of a specific weakness in the 'her' context, since this context is one of the strongest holdouts of the ASPIRATE MUTATION (Jones, 1977).

As mentioned above, unequal recognizability among morphological categories represents inherent differences in probability of access as a function of mutation class or of morphological environment. It is hard to know how to fit such a notion into current models of lexical access. The notion of word frequency, of course, can be used to account for differences among lexical items in probability of access: in the logogen model, for instance, high frequency words fire more readily because they have a higher state of resting activation. However, within the logogen model there is no mechanism available to register frequency of use for categories like mutation class (should that turn out to be the relevant domain of the effect) as abstract entities. A possible account would be that provided by the "satellite" model of Lukatela et al. (1980), which attempts to account for differences in access *time* between nominative and non-nominative forms in Serbo-Croatian. However, as we have seen, other aspects of the satellite model do not fit the Welsh situation very well: for example, the fact that in the model all inflected forms are separate entries connected only to the nominative (or basic) form, is not supported by the results of Experiment 2.

GENERAL DISCUSSION

The three experiments described all point to a single generalization about the organization of mutation variants in Welsh speakers' lexicons—namely, that the particular phonological relationships among the forms related by mutation processes seem to have little relevance for their relationships during lexical access. Experiments 1 and 2 showed that repetition priming is possible between mutation variants, and that this priming is equally robust between variants with and without change in phonological form. Further, Experiment 3 showed that there are differences among mutation variants in how readily they are recognized, but that such differences do not require corresponding differences in phonological form. Thus, it seems clear that the human cognitive capacity underlying lexical organization and access does not place severe limits on how phonological form can vary in lexical relations. From the Welsh results, in particular, we learn that morphological variants, in order to prime each other, (1) do not need to be concatenative (affixal); (2) do not need to share an unmodified base form; and (3) do not need to share an initial consonant. Thus, any theory of the lexicon that attempts to model human lexical retrieval, as opposed to language-specific behavior, must be equally appropriate for all the variations in phonological form that occur in natural language. As discussed in previous sections, neither the strong form of the logogen model nor the form of the shared entries hypothesis that requires an unmodified base form is adequate for this purpose.

A further problem for these models, and for the satellite model as well, stems from the fact that, in Experiments 1 and 2, priming was found to be symmetrical between variants. It appears that, for Welsh at least, all the variants are symmetrically represented in the lexicon. On the other hand, the fact that Welsh variants prime each other at all indicates that they must have some lexical structure in common, that is, they cannot be wholly independent entries. These findings, together with the differing recognizability of the mutations, mean that we need a model of lexical organization that does not entail asymmetries in priming among morphological variants, but that does accommodate asymmetries in ease of access to the variants.

With these points in mind, we feel that an adequate model of the lexicon should have two independent dimensions: a lexical dimension that accesses a single entry for each set of three morphophonological variants, and a morphological dimension that accesses the morphophonological classes of BASE, SOFT, or ASPIRATE mutations. It is important to emphasize that in this view of the lexicon, the dimension containing the mutation classes is autonomous. That is, the mutation class category can not be conflated with its phonological effects, because, as we have seen, the priming and recognition behaviors of the mutations are the same regardless of the presence or absence of phonological change in the first consonant. In particular, the nonmutating items, which do not show the phonological effects of the mutation classes, nevertheless show the recognizability differences for the different mutations. It should be noted that recent linguistic analyses have also emphasized the autonomy of the morphological level, both with respect to phonological structure (e.g., Anderson, 1982) and with respect to syntactic structure (Sadock, 1985).

The two proposed independent dimensions account for the dual aspects of priming symmetry and access asymmetry in the following way. First, priming a specific morphophonological variant entails recognizing two kinds of information: the specific lexical entry common to all the morphophonological variants, and the specific mutation class. The common lexical entry accounts for the fact that all the morphophonological variants prime each other equally, while the fact that both the lexical entry and the mutation class are primed accounts for the differences between priming with the identical form and priming with other variants. Further, the independence of the mutation class dimension permits differential access of the mutations in a manner analogous to the effect of frequency of occurrence on recognizability of lexical items. By modeling the mutation classes as separate entries on this independent dimension, it is possible to associate different degrees of recognizability with each mutation class, just as lexical items with different frequencies of occurrence are associated with differing amounts of recognizability. In fact, the differing ease of access (i.e., recognizability) of the different mutations may be a direct reflection of the frequency of the mutation, where the relevant index of frequency is the number of different lexical and syntactic environments in which the particular mutation is used (Boyce, 1983). Since the SOFT mutation is used in a much wider variety of environments than the ASPIRATE, it would have a higher frequency of occurrence, which, as we have seen, is associated with higher recognizability. Thus, the two dimensions allow recognizability to be affected both by the frequency of the individual lexical item and by the frequency of the particular mutation class.

This model of lexical organization could also be employed for languages with more conventional kinds of morphology–for example, the regular plural and past tense affixes in English. However, it should be recalled that there is one difference between the present results for Welsh and those for regular inflectional morphology in English. English does not show significant differences between priming with identical forms and priming with other variants. If we assume English has the same kind of organization that we are proposing for Welsh, then the difference must somehow reside in the susceptibility of the morphological classes themselves (mutation classes vs. past tense) to priming.

Thus far, the lexical organization we have proposed involves abstract lexical and mutation dimensions. To complete our proposed lexical representation, we need to relate these dimensions to the actual phonological form of words by specifying the internal structure of the lexical items. This internal structure must incorporate symmetrical relations among all the morphological variants, in order to accommodate the symmetry of priming we observed. That is, as noted above, a satellite model such as that proposed for Serbo-Croatian is not appropriate for Welsh. The minimal representation that meets the symmetry requirement is an "allomorphic" approach, in which all the morphological variants of a lexical item are fully specified as to their phonological form, and are linked together into a single lexical item either directly, or via connections to a single abstract node.

However, the use of fully specified allomorphs fails to capture, in the lexical structure, any phonological regularities that occur across the morphological variants. These regularities are most directly captured by "shared entry" models in which the phonological information common to all the variants is specified in a single shared phonological node, and the information unique to each variant is specified separately for each variant. For English and other inflectional languages, of course, this corresponds to the shared entry models of the logogen type, or (more generally) a concatenative, base plus affix approach. While a concatenative approach is not appropriate for Welsh, the basic insight of the shared entry models, that of analyzing phonological information into common and unique components, can be retained simply by relaxing the sequentiality and continuity constraints implicit in concatenative affixal morphology.

Recent work in *autosegmental phonology and morphology* has focused on precisely this issue: namely, the type of models appropriate for describing non-concatenative morphological processes (e.g., Lieber, 1984; Marantz, 1982; McCarthy, 1981, 1984). Rather than restricting phonological representation to strictly sequential units, autosegmental analyses decompose phonological form into parallel and autonomous simultaneously occurring units called *tiers*. Each tier provides a sequence of specifications for some subset of phonological features. For example, in the Semitic case discussed in the introduction, McCarthy (1981) has proposed that consonant and vowel "melodies" are represented on separate tiers. Lexical items are distinguished on the basis of consonantal sequences on the consonant tier, while different inflectional morphological categories are represented by different sequences of vowels on the vowel tier. The units on the two tiers are linked to positions on a third, common tier (the skeleton), and in this way the actual intercalation of consonants and vowels is represented.

In the Semitic case, each tier contains information about whole segments—consonants and vowels. Other morphological analyses have been suggested in which one of the tiers contains only a subset of the features required to completely specify a segment. Thus, Lieber (1984) has proposed a linguistic analysis of the morphological-phonological relations for a language (Fula) that shows patterns similar to those found in Welsh. Fula nouns may show as many as three different initial consonants as a function of morphological categories such as singular, plural, and diminutive. The actual forms for a given lexical item all share the same place of articulation, but may differ in that one is a continuant, one is a stop, and one is a prenasalized stop (e.g., the stem meaning 'free man' will have the variants /rim-/, /dim-/ and /ndim/). In Lieber's analysis, the morphological variants are decomposed into a stem whose initial consonant is unspecified for the features [continuant] and [nasal], and a separate tier that contains values for these features corresponding to the particular morphological categories. The combination of these two partially specified representations is seen as parallel to the concatenation of stem plus affix in the more commmon sequential case, in the sense that both separate the common and unique information.

It would be attractive to apply a similar analysis to Welsh. As in the Fula case, the morphological variants would be decomposed into a single underspecified stem whose initial consonant is specified only for the place of articulation, and a separate tier with the phonological information unique to each mutation. Although Lieber (1983) has attempted such an analysis, she has done so only for the SOFT mutation, and it is unclear how the analysis could be extended to account for the ASPIRATE mutation as well. The problem is that the mutations in Welsh do not have consistent phonological characteristics, as they do in Fula. That is, it is impossible to associate the mutations with any unique set of phonological features. While it is the case that the BASE, SOFT, and ASPIRATE mutations differ either in voicing or in degree of oral closure, these differences are not associated one-to-one with any single mutation. Thus, if the BASE form contains a voiceless stop, it will differ from the SOFT mutation in terms of voicing (e.g., /p/ vs. /b/), and from the ASPIRATE mutation in terms of degree of closure (e.g., /p/ vs. /f/). But if the BASE form contains a voiced stop, it will be the SOFT mutation, rather than the ASPIRATE, from which it differs in terms of degree of closure (e.g., /b/ vs. /v/); and the BASE and ASPIRATE mutations do not differ at all. The difficulties posed by these relationships have been clear to those who have tried to describe them by means of more traditional generative phonological rules. Awbery (1973), for example, notes that two completely unrelated rules are required to derive the SOFT mutation forms from the BASE forms (one for voiceless stops and liquids, and another for voiced stops), while the ASPIRATE mutation requires a rule that, in terms of its phonological effects (changing a stop to a fricative), ought to be collapsible with one of the SOFT mutation rules. But because the two stop-to-fricative rules apply in different morphological and syntactic environments, they cannot be collapsed. Thus, the phonological differences between these mutations are impossible to characterize in a simple and general way, at least using ordinary phonetic features.

However, it is possible to unify the descriptions of the phonological effects of the various mutations by considering them to be dynamic processes, as proposed by Griffen (1985). Griffen suggests that the mutations can be considered to be dynamic processes operating on a single scale of strength, where voiced fricatives are the weakest (designated by the value 1), voiced stops next (with a value of 2), voiceless stops next (3), and voiceless fricatives the strongest (4). (Acoustically, the strength scale corresponds to an increasing ratio of high-to-low frequency energy, according to Griffen, 1975.) For the labial place of articulation, this results in the following scale: (1) /v/ (2) /b/ (3) /p/ (4) /f/. In Griffen's analysis, the membership of each lexical item in the mutation system is defined in terms of the place of articulation of the beginning portion, and the strength of the BASE form. Starting from this classification of lexical items, the mutations are simply related in terms of their degrees of strength: the SOFT mutation is one degree weaker than the BASE form, and the ASPIRATE mutation one degree stronger. Thus, a lexical item whose BASE form begins with /p/ is defined as a labial with strength 3. Its corresponding SOFT mutation form would then have strength 2 (/b/) and the ASPIRATE mutation strength 4 (/f/). Similarly, a lexical item whose BASE form begins with /b/ is a labial with strength 2; its SOFT mutation form has strength 1 (/v/). The ASPIRATE mutation here has the same strength as the BASE, resulting in a gap in the application of the strength relations such that the ASPIRATE mutation only differs from the BASE if the BASE has strength 3.

While the notion of the strength scale has gaps in its application, it nevertheless provides a description of the phonological changes associated with the mutation classes that is simple enough to be used in an autosegmental account. This is true not only for the similarities and differences among the mutations and segments immediately relevant to this paper, but also, as Griffen (1985) shows, for the other mutations and segments in Welsh. And the strength scale is not simply an arbitrary scale. As noted above, it may reflect the acoustics of the segments. Moreover, increase and decrease in strength has been observed to be a frequently occurring historical process. This is true for Welsh, where the modern mutations are the fossilized remains of an earlier process that was quite phonologically regular (Jones, 1931). And changes in the strength of consonants have been observed for a number of languages, especially those in the Indo-European language family (Lass, 1984).

Thus, Griffen's strength scale appears to be well-motivated, as well as being well-suited to an autosegmental analysis of the phonological structure of the Welsh mutation classes. In such an analysis, the lexical item includes the underspecified phonological representation (with only place of articulation specified in the initial portion) and its link to the appropriate value on the separate strength tier (3 for voiceless BASES, 2 for voiced BASES). In addition, the mutation classes are associated with operations on the strength tier (adding values of 0, +1, and −1, respectively, for the BASE, ASPIRATE, and SOFT mutations). The mutation classes operate on the strength values associated with the lexical item to determine the correct strength for that morphophonological variant, and hence, the complete phonological representation. Certain restrictions are necessary to capture the gaps in the mutations, as well as to handle the nonmutating segments. Thus, the mutation classes will only change lexical strength values of 2 or 3, thereby exempting the nonmutating segments being considered in this paper, whose lexical item will be linked to the value of 4 on the strength tier. And the ASPIRATE mutation is further restricted to operate only on lexical items of strength 3.

Note that the phonological realization of the mutation classes with respect to the strength tier requires an operation, that of additivity. Current notions of the relationships among autosegmental tiers include linking between specific items, blocking of such links, removal of links, and re-assignment of links, but no arithmetic operations are posited. Thus, the use of operations such as addition for the combination of two (scalar) feature values represents an extension to autosegmental theory. This extension in fact brings the proposed autosegmental model of lexical organization quite close to a class of models currently being actively developed by psychologists, biologists, and computer scientists–parallel distributed processing or "connectionist" models (cf. Rumelhart & McClelland, 1986a; McClelland & Rumelhart, 1986). The image of a word suggested in autosegmental analyses–that of a set of links among nodes on independent tiers–is quite similar to the connectionist image of a network of interconnections among layers of simple, but dynamic, processing units. Each unit has some state of activation, which is influenced over time by the inputs from connected units.

In spite of this apparent surface convergence between the autosegmental and connectionist approaches, there is nevertheless a difference in underlying assumptions. The autosegmental approach assumes that the rules that combine information from the various tiers, for example, are not simply descriptive, but are also the mechanism by which the system operates. Connectionist models, in contrast, assume that rules serve to describe the behavior of a system but are not the mechanism whereby that behavior is achieved. Thus, as stated by Rumelhart and McClelland (1986b, p. 218), "[t]here is no denying that rules ... provide a fairly close characterization of ... performance ... We would only suggest that parallel distributed processing models may provide a mechanism sufficient to capture lawful behavior, without requiring the postulation of explicit but inaccessible rules. Put succinctly, our claim is that PDP [Parallel Distributed Processing] models provide an alternative to the explicit but inaccessible rules account of implicit knowledge of rules."

Rumelhart and McClelland (1986b) support the above claim by developing a connectionist model that learns both regular and irregular past tenses of English verbs, simply through repeated presentations of present and past tense forms. The system gradually develops a set of weights associating the nodes corresponding to phonological units of present tense forms with the nodes corresponding to past tense forms. In the course of learning, the system exhibits the various stages shown by children learning past tenses. That is, first a small set of verbs is learned with the correct past tenses, regardless of whether the past tenses are regular or irregular. In the second stage, the regular past tense is overgeneralized so that irregular verbs no longer have the correct past tense. In the final stage, a much larger set of verbs, with both regular and irregular past tenses, is learned correctly. What is interesting about this model in the current context is that even though the system "learns" regularities and can generalize, this knowledge resides only in the connections among the present and past tense pairs. The ability to generalize results from the similarity in the present-past associations for the multitude of regular verbs. That is, regularity consists of the same pattern being repeated in association after association—exactly the same situation as occurs in the Welsh mutation system. The exact nature of the pattern does not matter, only its regular occurrence.

From this point of view, the distinction between an allomorphic approach and an underspecified or autosegmental approach to characterizing lexical items may be more apparent than real. That is, the connectionist model associates fully specified allomorphs; as long as these allomorphs are specified using phonological features, it does not matter how the features associated with the morphological information are distributed throughout the lexical items. They can be completely isolatable in one portion of the lexical item, as an affix; or they can be only partially isolatable, as in the case of the mutations. The connectionist model should handle both these cases with equal ease. A similar conclusion is reached by Bybee (1985), who argues that these alternatives are endpoints on a continuum. For her, the lexicon consists of sets of dynamically linked allomorphs, each with variable lexical strengths (related to their frequency of occurrence: cf. also Stemberger & MacWhinney, 1986) and variable degree of connection to other morphologically related forms (at least partially as a function of phonological similarity).

This kind of connectionist model accommodates the basic similarities and differences between the English and Welsh morphological systems. Those patterns that recur in each language, regardless of their phonological nature, will result in generalizations based on their regularity. The specific nature of the generalization—affixal vs. mutation—will of course differ between the two languages. Nevertheless, certain details of the Rumelhart and McClelland (1986b) connectionist model for English are not appropriate for the facts of Welsh. In particular, the lack of a significant effect on priming due to phonological similarity (i.e., identity) among the mutations suggests that a simple connection via phonological feature nodes is not sufficient to characterize the connections among Welsh morphological variants. It might be possible to partially reconcile the Rumelhart and McClelland model with the Welsh data by positing a phonological identity effect for the first consonant that is so small as to be overwhelmed by the much larger effects due to the rest of the word and the additional differential recognizability of the individual morphological classes. Nevertheless, because the effect of the mutations remains strong regardless of their phonological realization, it is necessary to include an independent morphological level. This might be handled by additional input units or, possibly, by explicit "hidden units." The additional level in the Welsh lexicon might be capable of accounting for the fact, discussed above, that Welsh and English differ when comparing performance in the identical-priming and other-priming conditions. (However, note that a morphological level independent of phonological structure has also been suggested for English—cf. Fowler et al.'s [1985] suggested modification to Dell's [1984] network model.) In general, then, while it remains an appealing possibility that a connectionist model could be constructed that captures both the similarities and differences between English and Welsh, the possibility cannot be satisfactorily confirmed without the actual construction of an explicit model for testing.

To summarize, the two-dimensional model of lexical organization we have proposed, augmented with an autosegmental representation of the associated phonological structures, seems well suited to describing the kinds of lexical relations that develop in the case of regular nonconcatenative morphology like the Welsh mutations. Moreover, connectionist models offer the possibility of understanding the mechanism that gives rise to this kind of system behavior. By investigating languages with different morphophonological structures, it is possible to increase our understanding of the limits on the complexity of the relations among (semi-)autonomous linguistic structures (i.e., phonological, morphological, syntactic, semantic). The present paper has provided one attempt to extend our understanding of these limits.

References

Anderson, S. R. (1982). Where's morphology? *Linguistic Inquiry, 13*, 571-612.

Awbery, G. M. (1973). *Initial mutation in a generative grammar of Welsh* (Report No. 4, pp. 28-46.). Leeds: University of Leeds, Phonetics Department.

Awbery, G. M. (1986). Moves towards a simpler, binary mutation system in Welsh. In H. Andersen (Ed.), *Sandhi phenomena in the languages of Europe*. The Hague: Mouton.

Bentin, S., Bargai, N., & Katz, L. (1984). Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 353-368.

Boyce, S. E. (1983). A perceptual analog of change in progress in Welsh. *Haskins Laboratories Status Report on Speech Research, SR-76*, 159-165.

British Office of Population Censuses and Surveys. (1983). *Census 1981: Report for Wales*. London: Her Majesty's Stationery Office.

Browman, C. P. (1978). Tip of the tongue and slips of the ear: Implications for language processing. *UCLA Working Papers in Phonetics, 42*.

Bybee, J. L. (1985). *Morphology. A study of the relation between form and meaning*. Philadelphia: John Benjamins.

Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *The perception and production of fluent speech*. Hillsdale, NJ: LEA.

Dell, G. (1984). *A spreading activation theory of retrieval in sentence production* (Cognitive Science Technical Report). Rochester, NY: University of Rochester.

Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry, 8*, 505-520.

Forbach, G. B., Stanners, R. F., & Hochhaus, L. (1974). Repetition and practice effects in a lexical decision task. *Memory & Cognition, 2*, 337-339.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland Press.

Fowler, C. A., Napps, S. E., & Feldman, L. B. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory & Cognition, 13*, 241-255.

Fynes-Clinton, O. H. (1913). *The Welsh vocabulary of the Bangor district*. Oxford: Oxford University Press.

Goldstein, L. M. (1977). Perceptual salience of stressed syllables. *UCLA Working Papers in Phonetics, 39*, 37-60.

Goldstein, L. M., & van den Broecke, M. (1977). Response bias and subjective estimation of consonant frequency. *Journal of the Acoustical Society of America, 61(Suppl. 1)*, S65. (Abstract)

Griffen, T. D. (1975). *A new Welsh consonant shift, description and implications*. Doctoral dissertation, University of Florida, Gainesville. (University Microfilms International No. 76-4237.)

Griffen, T. D. (1985). *Aspects of dynamic phonology*. Amsterdam: John Benjamins Publishing Company.

Henderson, L., Wallis, J., & Knight, D. (1984). Morphemic structure and lexical access. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X. Control of language processes*. London: Erlbaum.

Howes, D. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *Journal of the Acoustical Society of America, 29*, 296-305.

Jones, J. M. (1931). *A Welsh grammar: Historical and comparative. Phonology and accidence.* Oxford: Clarendon Press.

Jones, S. (1926). *A Welsh phonetic reader*. London: University of London Press.

Jones, T. J. R. (1977). *Living Welsh*. Suffolk: Hodder and Stoughton.

Kempley, S. T., & Morton, J. (1982). The effects of irregularly related words in auditory word recognition. *British Journal of Psychology, 73*, 441-454.

Lass, R. (1984). *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

Lieber, R. (1983). New developments in autosegmental morphology: Consonant mutation. In M. Barlow, D. P. Flickinger, & M. T. Wescoat (Eds.), *Proceedings of the West Coast Conference on Formal Linguistics* (No. 2, pp. 165-175). Stanford, CA: Stanford Linguistics Association (Stanford University, Department of Linguistics).

Lieber, R. (1984). Consonant gradation in Fula: An autosegmental approach. In M. Aronoff & R. T. Oehrle with F. Kelley & B. W. Stephens (Eds.), *Language sound structure* (Chap. 17, pp. 328-345). Cambridge, MA: MIT Press.

Lukatela, G., Gligorijević, B., Kostić, A., & Turvey, M. T. (1980). Representation of inflected nouns in the internal lexicon. *Memory & Cognition, 8*, 415-423.

MacKay, D. G. (1979). Lexical insertion, inflection, and derivation: Creative processes in word production. *Journal of Psycholinguistic Research, 8*, 477-498.

Magen, H., & Manuel, S. Y. (1982, December). *Auditory recognition of morphologically related German words*. Paper presented at the meeting of the Linguistic Society of America.

Manelis, L., & Tharp, D. A. (1977). The processing of affixed words. *Memory & Cognition, 5*, 690-695.

Marantz, A. (1982). "Re reduplication." *Linguistic Inquiry, 13*, 435-482.

Marslen-Wilson, W. (1984). Function and process in spoken word recognition: A tutorial review. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X. Control of language processes* (pp. 125-150). London: Erlbaum.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29-63.

Matthews, P. (1974). *Morphology*. Cambridge: Cambridge University Press.

McCarthy, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry, 12*, 373-418.

McCarthy, J. (1984). Prosodic organization in morphology. In M. Aronoff & R. T. Oehrle with F. Kelley & B. W. Stephens (Eds.), *Language sound structure* (pp. 299-317). Cambridge, MA: The MIT Press.

McClelland, J. L., & Rumelhart, D. E. (1986). A distributed model of human learning and memory. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group, *Parallel distributed processing Explorations in the microstructure of cognition. Vol. 2. Psychological and biological models* (pp. 170-215). Cambridge, MA: The MIT Press.

Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental process. In C. N. Cofer (Ed.), *The structure of human memory*. San Francisco: W. H. Freeman.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*, 338-352.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76*, 165-178.

Morton, J. (1970). A functional model for memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press.

Morton, J. (1979). Word recognition. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics 2: Structure and processes* (pp. 107-156). Cambridge, MA: MIT Press.

Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology, 102*, 963-968.

Pollack, I. (1963). Reaction times to unknown word sets in noise. *Language and Speech, 6*, 189-195.

Rubenstein, H., & Pollack, I. (1963). Word predictability and intelligibility. *Journal of Verbal Learning and Verbal Behavior, 2*, 147-158.

Rumelhart, D. E., & McClelland, J. L. (1986a). PDP models and general issues in cognitive science. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1. Foundations* (pp. 110-146). Cambridge, MA: The MIT Press.

Rumelhart, D. E., & McClelland, J. L. (1986b). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2. Psychological and biological models* (pp. 216-271). Cambridge, MA: The MIT Press.

Sadock, J. M. (1985). Autolexical syntax: A proposal for the treatment of noun incorporation and similar phenomena. *Natural Language & Linguistic Theory*, 3, 379-439.

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior, 18*, 399-412.

Stemberger, J. P., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, 14, 17-26.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7, 263-272.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14*, 638-647.

Van der Molen, H., & Morton, J. (1979). Remembering plurals: Unit of coding and form of coding during serial recall. *Cognition, 7*, 35-47.

Willis, P. (1986). *The initial consonant mutations in Welsh and Breton*. Bloomington, IN: Indiana University Linguistics Club.

APPENDIX I: Stimuli

The stimuli are listed for each experiment in Welsh orthography, followed by their English gloss. Note that *w* and *y* represent vowels, pronounced roughly as in FOOL and FIN, *ch* represents a voiceless velar fricative, as in German *Bach*, *f* represents the English voiced labiodental fricative as in VENT, the double *ff* is the English voiceless labiodental fricative as in FOOL, the double *ll* is a voiceless lateral fricative, while the double *dd* is a voiced alveolar fricative as in English THAT.

Monosyllables

Exp. 1

beic	'bicyle'
brawd	'brother'
dial	'revenge'
drych	'mirror'
gwg	'frown'
gwallt	'hair'
pen	'head'
plyg	'fold'
trwyn	'nose'
tân	'fire'
cig	'meat'
caws	'cheese'

Exp. 2

bedd	'grave'
bocs	'box'
dwrn	'fist'
dyn	'man'
drwg	'evil'
darn	'piece'
gwin	'wine'
glaw	'rain'
pris	'price'
pwrs	'purse'
tolc	'dent'
taid	'grandfather'
trên	'train'
tir	'land'
cais	'attempt'
cwsg	'sleep'

Exp. 3

bwyd	'food'
budd	'profit'
gair	'word'
glyn	'glen'
pwys	'weight'
pôt	'pot'
cwm	'valley'
crys	'shirt

Bisyllables

Exp. 1

bwrlwm	'bubble'
bocsach	'boast'
dillad	'clothing'
deintydd	'dentist'
gorwel	'horizon'
priddfain	'brick'
talcen	'forehead'
tywydd	'weather'
colwyn	'puppy'
ceffyl	'horse'

Exp. 2

bachgen	'boy'
bywyd	'life'
diwrnod	'day'
darllun	'picture'
dyddiad	'date'
gelyn	'enemy'
plentyn	'child'
peiriant	'engine'
toriad	'cut'
tafod	'tongue'
tebot	'teapot'
tegan	'toy'
canol	'center'
capel	'chapel'

Nonmutating words

Exp. 3

ffŵl	'fool'
ffrind	'friend'
ffug	'fiction'
fflach	'flash'
chwant	'desire'
chwys	'sweat'
chwart	'quart'
chwyth	'breath'

APPENDIX II: Noise levels

The pretest consisted of thirty words (none of which was used in the experiment) set in the same BASE, SOFT MUTATION, and ASPIRATE MUTATION frames used in the experiment. The words were played to the subjects through a two-channel amplifier whose second channel could be attenuated in 2 db steps. The first channel (the noise) was mixed with the second channel (the signal words) and the output directed to the subject via stereo headphones. Beginning with the most intelligible of the signal-to-noise levels, the channel containing the signal words was attenuated by one 2 db turn of the dial every 6 items, a total of 5 different settings ranging in −2 db steps from +2 db to −6 db for monosyllables. The range for bisyllables was 3 db lower, from −1 to −9 db, to control for the higher recognition rate of bisyllables found in pilot studies (see Note 1). The subject's noise level for the target portion of the experiment was that noise level in which the subject had last guessed correctly one or more of the items within a 6-item group. The distribution of the different noise levels over subjects, and the effects of noise levels on recognition rates, are given in Note 2.

Note 1: The major problem for the technique of word recognition in noise is the possibility that a too-easy or too-difficult noise level for a particular combination of word and subject will swamp out any experimental effects due to priming for that combination. The ideal solution to this problem, of course, is extensive pretesting of individual subjects and individual words to establish performance under different noise conditions. In such a case, a different noise level can be chosen for each combination of word and subject that maximizes the likelihood that priming can be detected. Unfortunately, such testing is often not practical. Instead, investigators (Kempley & Morton, 1982; Magen & Manuel, 1982) have concentrated on choosing a single noise level for the experimental corpus that is free of ceiling and floor effects for a significant proportion of words. This can be done in several ways. In Kempley and Morton (1982) a pilot group of subjects was used to establish a 40% average recognition rate for the experimental corpus. In Magen and Manuel (an experiment on German) a pilot group of native and non-native speakers was tested on a sample corpus to locate a 40% average recognition rate. In our case, not being able to test the entire experimental corpus in New Haven due to the lack of native speakers, we drew on previous experience with small Welsh pilot studies on different subjects and different sets of words (some of which occur also in the experimental corpus) to estimate the average effects of various signal-to-noise ratios. The anchor level, that is, the 50% recognition level, was determined to be (approximately) at signal-to-noise equivalence, or a signal-to-noise ratio of 1 db for monosyllabic words and at −3 db for bisyllabic words. To control for this apparent general effect of redundancy (due to the increased word length), during the target phase of the experiment bisyllables were presented at a noise level 3 db greater than that for monosyllables.

Note 2: As mentioned previously, the noise level at which a subject heard the target tape was individually calibrated to each subject by means of a pretest. As it happened, subject performance on the pretest was an indifferent predictor of performance on the target list; in general, the subjects assigned the more difficult signal/noise levels had proportionately lower rates of recognition. Approximately half (30) of the subjects received a noise level of −4 db for monosyllables and −7 db for bisyllables, with the others normally distributed over remaining levels. The mean rate of recognition for all subjects, over all three experiments, was 46%. To check for the possibility that ceiling and floor effects influenced the pattern of the data reported here, the statistical analysis of each experiment was repeated using three subjects assigned noise level -4/-7 (monosyllables/bisyllables) from each subject group. Because, for each experiment, the results were the same as when the entire subject pool was used, but the reduced number of subjects in a subject group meant a loss of power, the results reported in the text are for the full complement of 56 subjects.

APPENDIX III: Estimated frequencies of stimuli

MUTATION

ITEM	MEAN	BASE	SOFT	ASP.	EXPERIMENT	TYPE
bedd	57	57	57	56	2	mono
bocs	64	72	67	54	2	mono
beic	66	65	-	68	1	mono
brawd	73	72	74	-	1	mono
bwyd	80	86	67	80	3	mono
budd	43	47	47	32	3	mono
dwrn	53	52	57	50	2	mono
dyn	69	83	58	68	2	mono
drwg	64	70	64	57	2	mono
darn	59	73	53	54	2	mono
dial	40	46	-	36	1	mono
drych	45	48	41	-	1	mono
gwin	58	68	53	52	2	mono
glaw	68	78	76	48	2	mono
gwg	25	28	-	21	1	mono
gwallt	70	68	71	-	1	mono
gair	65	68	67	59	3	mono
glyn	34	35	40	25	3	mono
pris	65	80	55	58	2	mono
pwrs	64	71	57	65	2	mono
pen	77	78	-	77	1	mono
plyg	27	29	23	-	1	mono
pwys	43	51	29	44	3	mono
pôt	57	67	49	42	3	mono
tolc	34	50	20	30	2	mono
taid	68	63	64	76	2	mono
tir	66	76	64	56	2	mono
trên	58	68	52	55	2	mono
trwyn	73	72	-	74	1	mono
tân	68	73	63	-	1	mono
cais	45	48	45	42	2	mono
cwsg	53	58	57	46	2	mono
cig	69	74	-	65	1	mono
caws	67	76	61	-	1	mono
cwm	56	63	43	54	3	mono
crys	75	72	72	82	3	mono
bachgen	72	73	74	70	2	bi
bywyd	70	68	66	76	2	bi
bwrlwm	29	32	-	28	1	bi
bocsach	18	19	17	-	1	bi
diwrnod	74	83	70	68	2	bi
darllen	61	69	60	56	2	bi

dyddiad	58	69	48	57	2	bi
dosbarth	64	70	68	52	2	bi
dillad	81	83	-	80	1	bi
deintydd	50	56	46	-	1	bi
golau	67	74	68	60	2	bi
gelyn	56	59	56	54	2	bi
gorwel	35	38	-	31	1	bi
gobaith	55	52	59	-	1	bi
plentyn	72	73	75	68	2	bi
peiriant	55	54	58	52	2	bi
pellter	48	60	-	40	1	bi
priddfain	15	15	15	-	1	bi
toriad	38	48	38	29	2	bi
tafod	64	62	75	53	2	bi
tebot	66	65	61	73	2	bi
tegan	51	50	55	48	2	bi
talcen	54	53	-	55	1	bi
capel	65	68	62	65	2	bi
colwyn	14	15	-	14	1	bi
ceffyl	61	64	58	-	1	bi
ffŵl	50	66	40	28	3	nonmutating
ffrind	77	73	78	83	3	nonmutating
ffug	30	28	26	40	3	nonmutating
fflach	36	38	38	30	3	nonmutating
chwant	48	52	49	37	3	nonmutating
chwys	53	54	58	47	3	nonmutating
chwart	34	31	45	28	3	nonmutating
chwyth	24	22	20	32	3	nonmutating

MEAN	55	59	54	52
S.D.	16	18	16	17
MIN	14	15	15	14
MAX	81	86	78	83

MEAN FREQUENCIES BY EXPERIMENT AND TYPE

exp1		exp2		exp3	
mono	bi	mono	bi	mono	nonmutating
58	44	59	62	57	44

Reliability of frequency estimates. The frequency estimates can only be considered to be approximations; the values in the table should be interpreted as an indication of, at best, the nearest decile. This is due both to measurement error and to the nature of the subject responses.

Measurement error for the frequency responses of individual subjects is at least ±5%, due to the experimenters' decision to round subjects' responses to the nearest 5. Additional possible sources of measurement error include lack of double checking during data entry into the computer (unlike the experiment proper, in which responses were double or triple checked, as necessary), and possible inconsistency in rounding decisions between the two individuals entering the responses.

In addition, the subject responses varied in accuracy. For some of the subjects, accuracy is ±20%; these subjects circled only the labeled frequency values on the answer sheets (frequency values of 0, 20, 40, 60, 80, and 100 were explicitly labeled). For other subjects, accuracy is ±10%; these subjects circled the labeled values and also the intermediate bisecting tick marks. A few subjects resorted to effectively binary choices (0 or 100): the rest of the subjects used the intermediate values, as expected.

As a rough measure of frequency, however, we believe these data may be fairly reliable, for several reasons. First, the presentation of the items in different mutation environments serves as a kind of replication; the relatively high correlations among the frequency estimates of the same item in different mutation environments (reported below) suggest that the estimates are at least fairly replicable. Second, Goldstein and Van den Broecke (1977) have shown that estimates of phoneme frequency correlate well with other measures of phoneme frequency, in English. Finally, anecdotal reports from subjects at the time of running the experiments were consistent with the frequency estimates for the low frequency items in the list.

For instance, all of the words with an overall score of 25 or less were spontaneously noticed by at least one subject as being odd or unnatural. *Priddfain* 'brick', for example, was remarked on as a word everyone knows from church but that is never used in conversation. (It may be worth noting that the North Wales dialect speaker we consulted in New Haven kept up her Welsh by attending a Welsh chapel.) Subjects also revealed that, for them, the primary denotation of *colwyn* 'puppy' was the nearby Colwyn Bay. *Bocsach* 'boast' and *gwg* 'frown' were apparently somewhat obscure, as several subjects claimed to have never encountered them.

Subject comments also revealed that five more words— *chwant* 'desire', *tegan* 'toy', *taid* 'grandfather', *bachgen* 'boy', and *cwm* 'valley'— have different frequencies in North and South Wales, *chwant* and *bachgen* being more common in South Wales, and *taid*, *tegan*, and *cwm* being used predominantly in North Wales. The estimated frequencies for these words therefore reflect the different experiences of North and South Wales dialect speakers among our subjects. Note, however, that we had approximately equal numbers of subjects from North and South Wales. Note also that, although the experiment took place in North Wales, the subjects were all familiar with the literary language, which is based on the dialect of South Wales.

The high frequency words in the list were also consistent with expectation. For instance, words with the highest frequency scores were common words for common objects or concepts: *bwyd* 'food', *crys* 'shirt', *dillad* 'clothing', and *ffrind* 'friend'. Words such as *brawd* 'brother', *dyn* 'man', and *diwrnod* 'day', which are expected to be frequent, have high scores also.

EXPERIMENT and TYPE effects. An analysis of variance was performed using the mean estimated frequency for each item, where the mean was the average across subjects and mutation classes (as reported in the leftmost numerical column of the above table). The grouping factors of EXPERIMENT (1, 2, or 3) and TYPE (MONO, BI, or NONMUTATING) were tested. Both EXPERIMENT ($p < .02$) and EXPERIMENT × TYPE ($p < .04$) were marginally significant; TYPE was not significant. Tests of simple main effects indicated that the EXPERIMENT effect was confined to a significant difference between the BIsyllables in EXPERIMENTs 1 and

2, with the BIsyllables in EXPERIMENT 1 being significantly lower in frequency than those in EXPERIMENT 2 ($p < .006$). No other simple main effects were significant.

Although TYPE was not significant in the overall analysis, nevertheless the NONMUTATING items in EXPERIMENT 3 were markedly lower in frequency than the MONOsyllabic items. Since this contrast was an important one, we felt justified in further tests of significance. TYPE remained non-significant when tested in EXPERIMENT 3 alone. When MONO and BI from all the experiments were grouped together and compared (without EXPERIMENT as a grouping factor) to the NONMUTATING items, thereby increasing the degrees of freedom, the NONMUTATING items were marginally significantly lower than the grouped other items ($p < .04$).

MUTATION effects. Here we used the estimated frequency averaged across subjects (as reported in the second through fourth numerical columns in the above table). The estimated frequencies for the three MUTATIONS, BASE, SOFT, and ASPIRATE, were fairly highly correlated: for BASE with SOFT, $r = .81$; BASE with ASPIRATE, $r = .81$; and SOFT with ASPIRATE, $r = .75$. The analyses of variance detailed below revealed that the items were judged to be significantly higher in frequency when they occurred in the BASE form than when they occurred in either of the other MUTATIONS; there was no significant difference in the estimated frequencies for items occurring in the SOFT and ASPIRATE MUTATIONS.

Using subsets of EXPERIMENT (1, 2, or 3) and TYPE (MONO, BI, or NONMUTATING) as grouping factors and MUTATION (BASE, SOFT, or ASPIRATE) as within-group factors, analyses were run to provide a rough indication of the significance of the differences of the frequency estimates for items in the different mutations. Since not all combinations of the grouping factors occurred in all the experiments, the analyses were run both on subsets of EXPERIMENTS and on subsets of MUTATIONS. No main effects, other than that for MUTATION, and no interactions were significant in any of the analyses reported below.

For EXPERIMENT levels 2 and 3, using all levels of MUTATION, MUTATION was significant ($p < .0001$). Further analyses revealed that the MUTATION effect was confined to the significantly higher estimated frequencies for items occurring in the BASE, compared to either of the other MUTATIONS. Thus, for all levels of EXPERIMENT (1, 2, and 3), items were estimated to be significantly higher in frequency when they occurred in the BASE than in the SOFT MUTATION ($p < .0058$). Again for all levels of EXPERIMENT, items were estimated to be significantly higher in frequency when they occurred in the BASE than in the ASPIRATE MUTATION ($p < .0003$).

For EXPERIMENTS 2 and 3, frequency estimates in the SOFT and ASPIRATE MUTATIONS were not significantly different. For EXPERIMENT 1, the SOFT and ASPIRATE MUTATIONS could not be directly compared since each lexical item occurred with only one of the two MUTATIONS. Thus, the SOFT and ASPIRATE MUTATION frequencies were averaged for each lexical item in EXPERIMENTS 2 and 3; for EXPERIMENT 1, the "average" consisted of the single frequency for whichever of the two MUTATIONS occurred with that lexical item. These average values were contrasted with the frequencies in the BASE for all levels of EXPERIMENT. The frequencies in the BASE were significantly higher than the averaged frequencies of the two MUTATIONS ($p < .0001$).

THE EMERGENCE OF CEREBRAL ASYMMETRIES IN EARLY HUMAN DEVELOPMENT: A LITERATURE REVIEW AND A NEUROEMBRYOLOGICAL MODEL *

Catherine T. Best†

Ever since Broca's century-old discovery of cerebral asymmetries in language functions, there has been speculation about the developmental emergence of human perceptual-cognitive asymmetries. The basic question has been: Do the asymmetries first appear only at some point after birth, starting from an initial state of bilateral equivalence or symmetry at birth, or are the hemispheres instead functionally asymmetrical from the start? A related question is whether functional asymmetries can be traced to some lateral bias in the structural development of the hemispheres, such that one hemisphere matures in advance of the other. Thus, the term "emergence" in the title of this chapter might refer either to functional asymmetries in infants or to the embryological development of the hemispheres. An underlying assumption of this chapter is that the ontogeny of functional asymmetries is influenced by an asymmetry in the formation and physical maturation of the cerebral hemispheres. Both issues will be addressed in the following discussion, beginning with a review of behavioral evidence for perceptual-cognitive asymmetries in early infancy, and ending with a proposed model for a lateralizing gradient in the neuroembryologic emergence of the cerebral hemispheres during prenatal development.

As for the development of functional asymmetries, Broca (1865) himself speculated that language becomes lateralized to the left hemisphere during language development (see Bever, 1978). Later, Samuel Orton (1937) expanded on this concept of "developmental lateralization" in his influential theory that dyslexic children suffer from a failure to establish cerebral dominance for language developmentally. More recently, Lenneberg (1967) further detailed the model of developmental change in lateralization of functions, proposing a critical period for language development, and hence for the progressive establishment of left-hemisphere language dominance, during the period between 2 years and puberty. More important for the present discussion, however, his model explicitly assumed that the child's cerebral hemispheres are equal in their capability to acquire language, a trait referred to as "equipotentiality," until at least 2 years of age. In fact, the focus on language in theoretical discussions of those times, to the exclusion of other cognitive functions, led to the reasoning that since infants have not yet acquired language, they should not

---

* To appear in S. Segalowitz & D. Molfese (Eds) *Developmental implications of brain lateralization.* New York: Guilford Press.

† Also Department of Psychology, Wesleyan University, Middletown, CT 06457

Aknowledgment. This work has profited from discussions with numerous colleagues and students. I am particularly grateful for discussions with, and careful critiques by, Marcel Kinsbourne, Donald Shankweiler, Ursula Kirk, and Marshall Gladstone. This work was supported in part by a grant from NICHD to Haskins Laboratories (NIH:HD-01994) and by a Biomedical Research Support Grant to the author through Wesleyan University.

show any hemisphere specialization. Implicit in the equipotentiality concept has been the notion that the infant's hemispheres show functional symmetry, or lack of behavioral differentiation.

Since the mid-1970's, however, evidence of functional cerebral asymmetries in young infants has indicated that the assumption of functional symmetry between the hemispheres in early development cannot be correct. Generally, this literature suggests a pattern of functional cerebral asymmetries by at least 2-3 months of age, and possibly even before full-term birth, that is analogous to the adult pattern of left hemisphere superiority for language-related functions and right hemisphere superiority for music and holistic perception of patterns and faces. The next part of the chapter will focus on perceptual-cognitive asymmetries in infants (for discussion of motoric asymmetries in infants, see the chapter by Turkewitz, this volume*), and particularly on behavioral evidence (electrophysiological data are presented in the chapter by Molfese, this volume*).

Functional Asymmetries in Infants

Before the specific findings are reviewed here, some preliminary qualifications are necessary. Up to this point, many questions about infant hemispheric specialization remain unanswered. It is not yet known, for example, whether infant asymmetries are fundamental responses to certain stimulus properties or classes, such as the physical characteristics of speech vs. nonspeech, or instead whether they reflect different processing styles, such as feature-analysis vs. holistic processing, as has been proposed for adults. In addition, the behavioral studies are actually quite few in number. Moreover, they have focused overwhelmingly on auditory asymmetries, particularly for human speech. This is due, in part, to a strong theoretical bias toward assessing language-related functions, but is also due to pragmatic constraints. The dichotic listening procedure is more obviously amenable to infant research than are the lateralized behavioral measures of asymmetries in other modalities (e.g., the requirements of the visual split-field tachistoscopic technique are obviously not suited to infants!).

Non-auditory Asymmetries.

Thus far, only two behavioral studies of infant cognitive-perceptual asymmetries in other, non-auditory modalities have been conducted, both of which assessed right hemisphere advantages for pattern recognition. One of these was inconclusive regarding functional asymmetries during infancy; the other is not yet published. In the first, Susan Rose (1984) tested 1-, 2-, and 3-year-olds for a left-hand advantage (right hemisphere superiority) in haptic perception of shapes. After blind, unimanual palpation of a 3-dimensional nonsense shape, children were tested for cross-modal shape recognition on a visual preference task in which they saw a picture of the palpated object presented alongside a picture of a differently-shaped object. Although all children showed preferences for the novel figure, and hence recognition memory for the palpated object, only the 2- and 3-year-olds showed a left-hand/right-hemisphere superiority. The 1-year-olds—the only infants in the study—failed to show a right-hemisphere advantage. However, as Rose argues, this cannot be taken as evidence for a lack of infant right-hemisphere specialization, because the visual test phase of the task involved bihemispheric, or non-lateralized, visual input. In fact, the left-hand effects even for the older children were rather small. Perhaps some other, more sensitive and completely lateralized test measure would detect tactile asymmetries in infants.

In the other non-auditory behavioral study, Witelson and Barrera (Witelson, personal communication) tested visual asymmetries in 3-month-olds. They presented the infants with side-by-side slides of two identical photographs, both of which were either of the infant's mother, or of a female stranger, or of a standard black-and-white checkerboard pattern. The infants showed a fixation-time preference for the left-side photo of the mother, as well as for the left-side checkerboard, suggesting greater activation of their right hemispheres. However, they did not show any side preference for the stranger. The authors' interpretation was that both mother and checkerboard constituted "gestalt" patterns to the infants, which they processed holistically, thus showing a right-hemisphere bias in activation. In contrast, the female stranger was not processed as a gestalt, and thus not handled preferentially by the right hemisphere. This argument, at least with respect to the infants' responses to mother vs. stranger, is consistent with developmental research on face recognition in children (Levine, 1985). Young children perceive unfamiliar faces in terms of salient features rather than holistically and show no hemispheric asymmetry for recognition of those faces. However, the same children do perceive familiar faces holistically, as well as showing a right hemisphere advantage for the familiar faces. Thus, the Witelson and Barrera results offer some suggestion of a right-hemisphere bias in holistic perception of faces (and patterns) by 3 months, which is compatible with the literature on visual asymmetries in adults. However, this suggestion must be viewed as still tentative, given that is is based only on a single, unpublished finding.

Auditory asymmetries

By comparison, the behavioral studies of auditory asymmetries have been more numerous, and have included assessments of both left- and right-hemisphere specialization in infants. The technique used is some modification of the dichotic listening procedure. In the first such study, reported at a conference held at Brock University in 1975, Anne Entus (1977) used a non-nutritive sucking measure with a dichotic habituation-dishabituation procedure. Two groups of infants, who averaged 2½ months in age, were tested for ear differences in discrimination of either musical notes played by different instruments, or of consonant differences in speech syllables. In the first phase of each test, the infants heard a rapidly repeated presentation of a dichotic pair of stimuli until they reached a criterion of habituation. At that point, the element in either the right or left ear was changed to, respectively, a new music note or syllable, while the other ear continued to receive its original habituation stimulus. The infants in the speech condition showed a greater recovery of the sucking response when the syllable changed in the right ear (REA) than when it changed in the left, indicating left-hemisphere superiority. The music group showed the opposite pattern, a left-ear (LEA) or right-hemisphere advantage. However, Vargha-Khadem and Corballis (1979) subsequently failed to replicate with 2-month-olds the speech REA that Entus found, a point to which we will return later in the chapter. In this later study, the infants discriminated the speech syllable change equally well with both ears.

In a similar dichotic habituation study with 3-month-olds, Glanville, Best, and Levenson (1977) used the heart rate measure of a deceleratory orienting response, reflecting interested attention to a stimulus, in order to introduce a memory component to the task. Friedes (1977) has presented evidence that, in adults, memory retrieval is more strongly associated with dichotic ear asymmetries than is a simple input processing dominance. Therefore, the intervals between presentations of the dichotic pairs in the Glanville et al. test were long enough ($M = 25 sec$) that the infants had to rely on short term memory in order to learn the habituation pair and to recognize the stimulus change on the test trial. In each test block, the habituation pair was

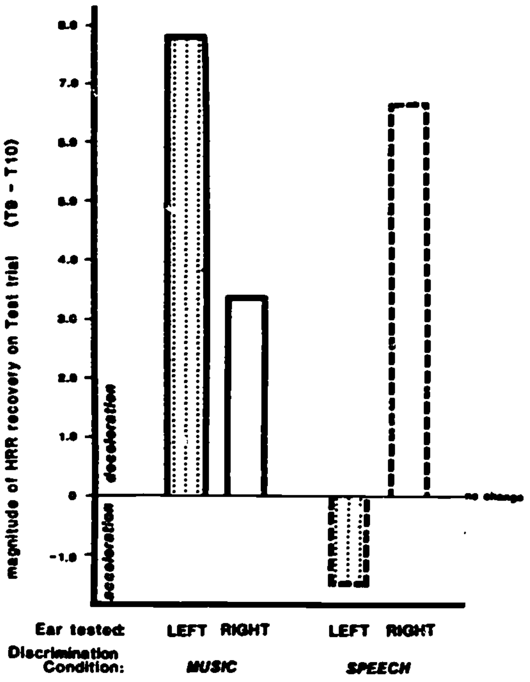


Figure 1. Ear differences in 3-month-olds' discrimination of speech syllables differing in initial consonant, and of music notes differing in instrument timbre. The task was a dichotic habituation-dishabituation task using heart rate deceleration as the response measure; represented here is the magnitude of dishabituation (cardiac deceleration) on the test trial (stimulus change in either right or left ear), relative to the cardiac response on the last habituation trial (trial 9). Redrawn from data reported in Glanville, Best, and Levenson (1977).

presented nine times, and the stimulus change was then presented on the tenth and final trial. All infants received separate left and right-ear discrimination test blocks each for speech syllables and for music notes. The results provided converging evidence with Entus' findings for an adultlike pattern among 3-month-olds of REA in response to speech syllable changes, and LEA in response to music changes (see Figure 1).

These first dichotic studies still left several important questions, two of which were addressed by subsequent research with infants. First, there have been two attempts to obtain a better specification of the speech properties to which the infant's left hemisphere is preferentially responsive. Second, age changes in behavioral evidence of auditory cerebral asymmetries during infancy have been assessed.

The basis of left hemisphere speech specialization. For a more detailed understanding of the infant's left-hemisphere response to speech, Best (1978) used the dichotic heart rate habituation procedure to determine whether $3\frac{1}{2}$-month-olds show different patterns of ear asymmetries for vowel vs. consonant discriminations. Several studies with adults had suggested that the REA for speech perception is greatest for consonant perception, while there is often a weaker or absent ear advantage for vowel perception (e.g., Darwin, 1971; Studdert-Kennedy & Shankweiler, 1970; Weiss & House, 1973). This pattern may be related to the fact that consonants involve rapidly

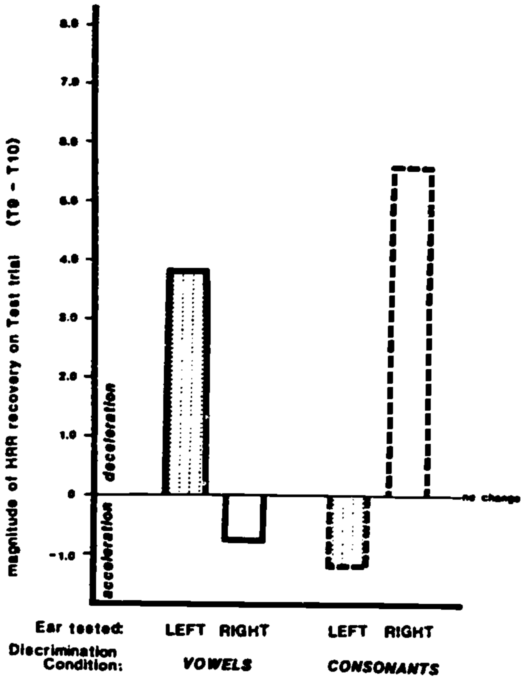


Figure 2. Ear differences in $3\frac{1}{2}$-month-olds' discrimination of vowels and of consonants in computer-synthesized syllables. The method and response measure are the same as described for Figure 1. Based on data presented in Best (1978).

changing acoustic properties, whereas vowels are associated with much more slowly-changing, or even steady-state, acoustic properties (see Cutting, 1974; Schwartz & Tallal, 1980). Therefore, a set of computer-synthesized syllables were developed, which exaggerated the rapidly changing acoustic properties vs. steady-state characteristics associated with consonants vs. vowels. The results revealed a REA for discrimination among the exaggerated consonants, consistent both with the earlier infant studies and with adult findings. However, the infants showed an LEA for steady-state vowel discrimination (see Figure 2), unlike adults.

These findings suggested that the infant's left hemisphere may be particularly responsive to rapidly-changing acoustic information, while the right hemisphere is more responsive to steady-state spectral information. The vowel LEA is compatible with John Sidtis' (1980) findings of a right-hemisphere advantage in adults' perception of steady-state harmonic information. The lack of an adult ear advantage for vowels suggests that this steady-state information may be easily transferred across the corpus callosum; the left-ear advantage in infants may be due to the immaturity of their corpus callosa (see also Molfese, Freeman, & Palermo, 1975; Molfese & Molfese, 1985; Studdert-Kennedy & Shankweiler, 1980).

MacKain, Studdert-Kennedy, Spieker, and Stern (1983) further explored the nature of the infant's left hemisphere specialization for speech perception, in a bimodal-matching study with

5-6-month-olds. The infants viewed two side-by-side synchronous video films of a woman repeating two different 2-syllable nonsense words, while they simultaneously heard a synchronous audio recording (over a centrally-located loudspeaker) that corresponded to one of the two video displays. Infants detected the cross-modal equivalence, as indicated by a looking preference for the film that matched the audio presentation, but only when the correct video was in the right-side video monitor. This finding implies selective left-hemisphere activation, and suggests a left-hemisphere specialization for perception of the common underlying articulatory pattern that produced the disparate information in the two sensory modalities.

Together, these two studies on cerebral asymmetries for the properties of speech suggest that the infant's left hemisphere may be specialized for recognizing articulatory patterns in speech, and particularly the rapid acoustic changes resulting from the dynamic articulatory gestures that produce consonant sounds. However, this still leaves open the question "Why the left hemisphere?" One possibility is a lateralized gradient in the maturation of the two hemispheres.

Lateral differences in hemisphere maturation? If there is a lateralized developmental gradient, uncertainty still remains as to whether the asymmetry in speech perception would result from earlier or later development of the left hemisphere relative to the right. Broca (1865) proposed a left-to-right gradient to explain language lateralization (see Bever, 1978); recently, Corballis and Morgan (1978) seconded the notion of a left-right gradient. However, Taylor (1969), Crowell, Jones, Kapuniai, and Nakagawa (1973), and Brown and Jaffe (1975) have argued for a right-to-left gradient, which is also suggested by recent embryologic evidence that cortical fissures appear consistently earlier in the right than the left fetal hemisphere (Dooling, Chi, & Gilles, 1983).

To test the possibility of early age changes in asymmetrical function, Best, Hoffman, and Glanville (1982) tested for ear asymmetries in memory-based discriminations of speech syllables vs. music notes by 2-, 3-, and 4-month-old infants. The 3- and 4-month-olds replicated the earlier findings of a REA/left-hemisphere advantage for speech and LEA/right-hemisphere advantage for music. However, the 2-month-olds showed only the LEA for music; they did not detect the speech syllable change in either ear (see Figure 3). These results suggest an increase in functional maturity of the left hemisphere sometime between 2 and 3 months of age, at least for auditory discriminations that depend on short-term memory capacities. Such a change in cortical maturity around 2-3 months of age is consistent with reports of widespread biobehavioral changes and maturation of cortical influences over behavior around that time (Emde & Robinson, 1979).

This finding may also help explain the negative report by Vargha-Khadem and Corballis (1979), which had failed to replicate findings by Entus (1977) of a speech REA in infants. Although their infant subjects discriminated the syllable change in both ears, discrimination under the rapid stimulus presentation conditions used in the sucking habituation procedure clearly does not depend solely on cortical involvement, since Frances Graham and her colleagues (Graham, Leavitt, Strock, & Brown, 1978) have found similar speech discrimination in a 6-week-old anencephalic infant. The conclusion of Best et al. (1982) was that the speech-specialized function of the left hemisphere may be insufficiently mature at 2 months to control behavioral responses in a memory-dependent discrimination task. In contrast, the analogous right hemisphere function appears sufficiently mature at that age to effect a LEA for memory-based music timbre discrimination, suggesting a right-to-left gradient in the maturation of asymmetrical perceptual memory functions. This does not necessarily imply that cerebral lateralization itself develops

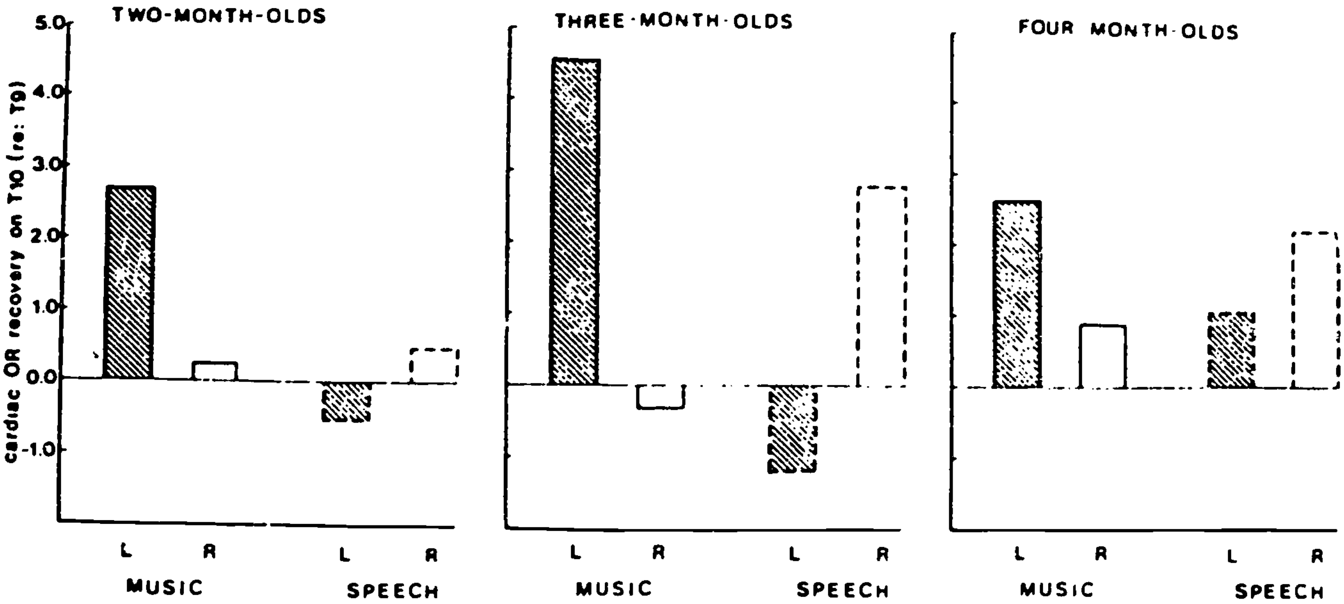


Figure 3. Age changes in ear differences for infants' discrimination of speech syllables and of musical timbre. Method and response measure are identical to the description of Figure 1. Reprinted with publisher's permission, from Best, Hoffman, and Glanville (1982), *Perception & Psychophysics, 31*, 75-85. Copyright ©1982 by the Psychonomic Society.

out of an unlateralized substrate. Alternatively, cognitive and perceptual functions may mature developmentally at different rates, but within the context of a neural substrate that is already laterally-specialized from the start (see Witelson, 1977, 1985; see also Kinsbourne, 1975).

If the latter view is correct, the question becomes: "What is the source of this lateralized gradient in functional maturation?" According to developmental biologists, morphologists, and particularly neuroembryologists, the patterns of embryologic development are ultimately responsible for the structure and form of the adult organism, including the brain, both at the gross morphological level and at the histological level. Given the basic neuropsychological assumption that variations in neuronal organization and development affect behavior and its development, then, fetal brain development should provide evidence of a lateralized developmental gradient. In the remainder of this chapter it will be argued that a right-to-left gradient in postnatal functional maturation parallels a similar gradient in the prenatal, embryologic development of the cerebral hemispheres.

Proposal for a Lateralized Gradient in Neuroembryologic Development

Gross Morphological Asymmetries

The notion that morphological properties of the brain find their expression during embryogenesis can be traced back at least to the early neuroscientist Pernkopf (see Keller, 1942), and is consistent with general principles of contemporary theory in embryology and evolutionary biology. Thus, the patterns of morphological asymmetries found in adult brains should be attributable to embryological growth patterns. From a complementary perspective, we should be able to "read" adult asymmetries in brain structure as a record of the forces in embryologic growth. Presented in the following pages is a working model of just this sort of interpretive view of the various structural asymmetries in the cerebral hemispheres that have been reported in the past 10-15 years.

The most well-known asymmetry, found in the majority of right-handed adults' brains, was first reported by Geschwind and Levitsky (1968), and replicated or extended in numerous subsequent reports (e.g., Galaburda, 1984; Galaburda, LeMay, Kemper, & Geschwind, 1978; Galaburda, Sanides, & Geschwind, 1978). The left hemisphere shows a larger surface area of the planum temporale (see Figure 4), which incorporates the auditory association area that is central to language comprehension, known as Wernicke's area. This asymmetry is found in the majority of infant (Witelson & Pallie, 1973) and fetal brains as well (Chi, Dooling, & Gilles, 1977; Wada, Clarke, & Hamm, 1975).

Another language-specialized area in the left hemisphere is Broca's area in the frontal lobe, encroaching on the Sylvian fissure. The size of this crucial speech production area in the majority of both adult and fetal brains is paradoxically smaller in the left hemisphere than the right, if measured as the visible surface area, according to Wada et al. (1975). However, several reports suggested that this region is more deeply fissurated in the left hemisphere. Following up on this suggestion, Falzi and colleagues (Falzi, Perrone, & Vignolo, 1982) measured the cortical surface area in the regions corresponding to Broca's area, pars triangularis and pars opercularis, in both hemispheres, such that their measurements included the cortex buried inside the sulci. They found this anterior speech region to be larger on the left than the right in three-quarters of their cases, when "hidden cortex" inside the sulci was taken into account in this manner (see Figure 5), thus implicating greater fissuration on the left, a point to which we will later return.

Let us focus now on another set of morphological asymmetries reported by Marjorie LeMay (1976, 1977, 1984; LeMay & Geschwind, 1978; LeMay & Kido, 1978), which exist in the majority of adults, children, and fetuses, as well as in corresponding measurements of prehistoric human skulls (LeMay, 1976, 1984). The pattern is a wider and more protruding right frontal lobe, a characteristic referred to as right frontopetalia; there is a converse left-hemisphere bias in the posterior portion of the brain, where a left occipitopetalia (greater backward protrusion) is found along with a wider left occipital region. These patterns are illustrated in Figures 6, 7, 8, and 9. Notice that the left occipitopetalia is generally more striking than the right frontopetalia. These characteristics are reflected in gross volumetric measures of frontal and occipital regions (Weinberger, Luchins, Morihisa, & Wyatt, 1982—see Figure 10) and by corresponding asymmetries in the skull itself. As mentioned earlier, they are also evident in fetal brains (Figure 11) and skulls (Figure 12—note the positions of the bone plates before the fontanelles have closed).

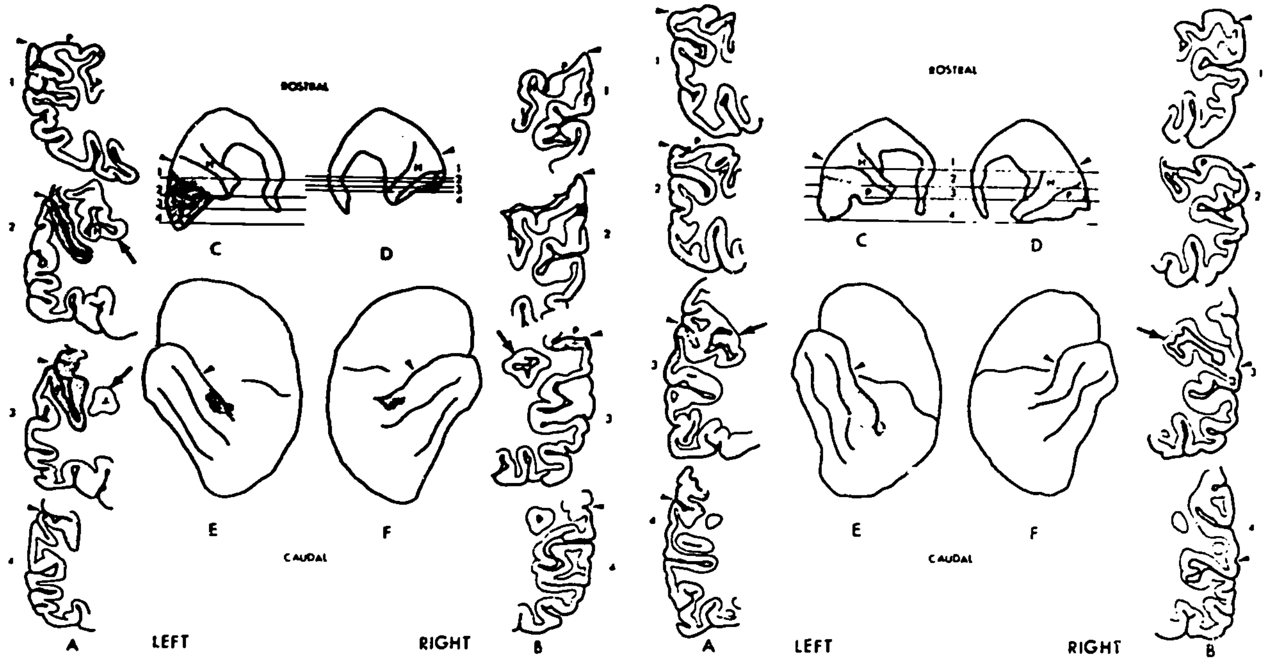


Figure 4. Anatomical analyses of two brains are shown here, illustrating the left hemisphere bias in size of planum temporale. Coronal sections are shown in columns A and B; superior temporal planes in C and D showing planum temporal (P) and Heschl's gyrus (H); and lateral surfaces of brain showing area Tpt (shaded) in E and F. Arrowheads point to Sylvian fissures. Note large asymmetry in Tpt and P. Also note buried temporal cortex on coronal sections (arrows). Reprinted with permission of senior author and publisher, from Galaburda, Sanides, and Geschwind (1978), *Archives of Neurology, 35*, 812-817. Copyright ©1978 by the American Medical Association.

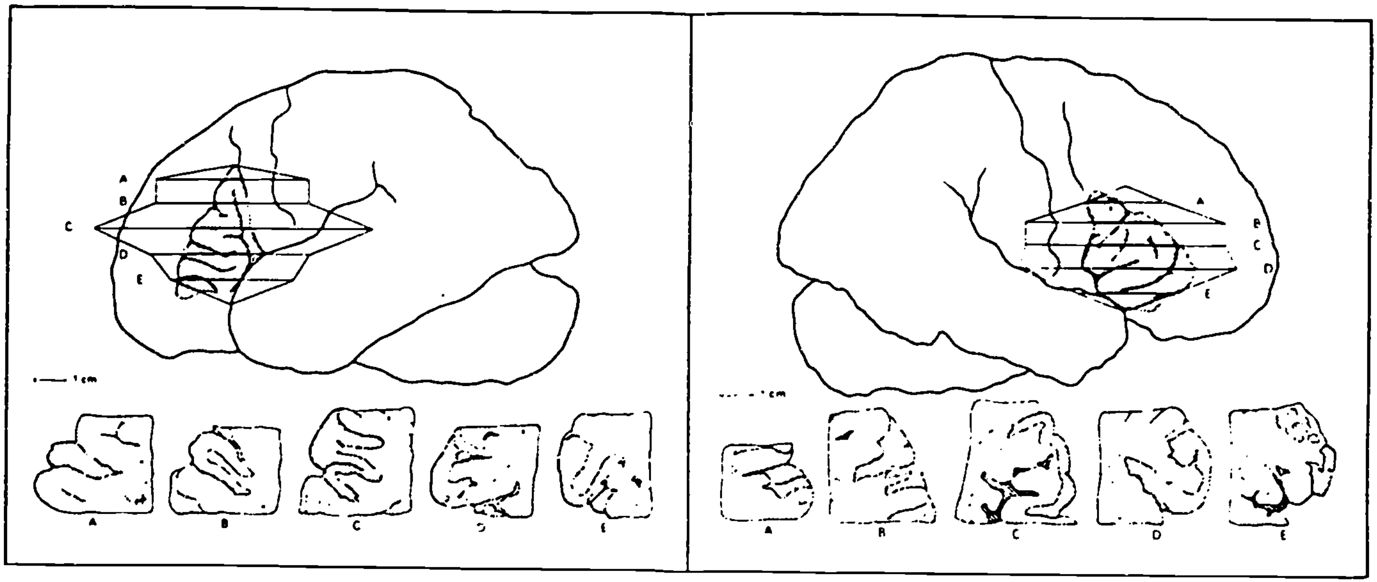


Figure 5. "Entire" anterior speech region (defined as pars opercularis and pars triangularis of the third frontal convolution), and its right hemisphere homologue, shown here superimposed within the polygons on the surface of both hemispheres. The figure illustrates one of the twelve right-handed cases reported in Falzi, Perrone, and Vignolo (1982). The authors measured both the extrasulcal and intrasulcal cortex of these regions (the polygons superimposed on the hemispheres show the sections made in order to measure intrasulcal cortex). There was more *intra*sulcal cortex found in the left hemisphere in 3/4 of their cases, indicating greater fissuration on the left than the right hemisphere. Reprinted with permission of the authors and publisher, from Falzi, Perrone & Vignolo (1982), *Archives of Neurology, 39*, 239-240. Copyright ©1982 by the American Medical Association.

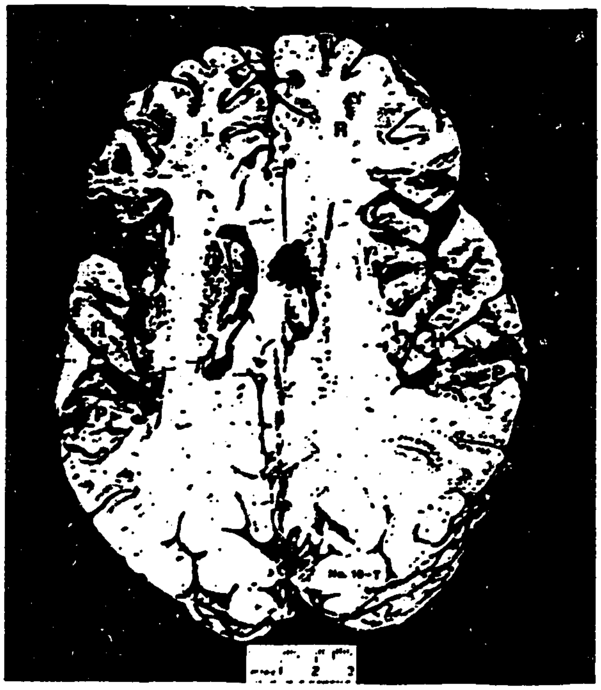


Figure 6. Horizontal section of an adult brain, exposing the plana temporale (P) and Heschl's gyri (H). Note both the larger planum on the left, and the right frontopetalia and left occipitopetalia. Reprinted with publisher's permission, from Witelson (1977), *Annals of the New York Academy of Sciences, 299*, 328-354. Copyright ©1977 by the New York Academy of Sciences.

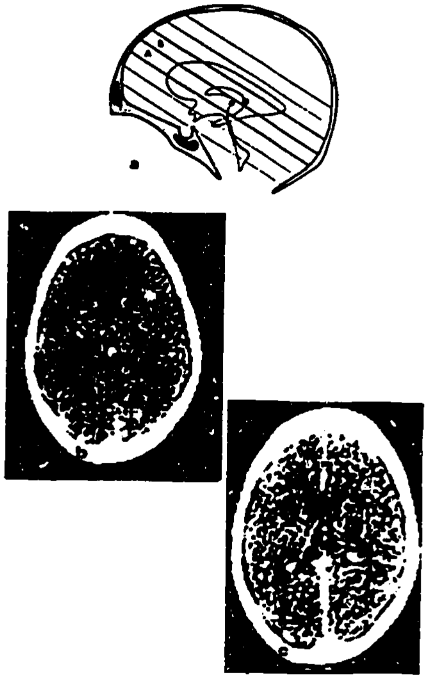


Figure 7. a. Diagram of sections taken through the brain during routine examination by x-ray computerized axial tomography (CT); b. CT scan though section A; c. CT scan though section B. Note the left occipitopetalia and wider right frontal region. Reprinted with publisher's permission from LeMay (1976), *Annals of the New York Academy of Sciences, 280*, 349-366. Copyright © 1976 by the New York Academy of Sciences.

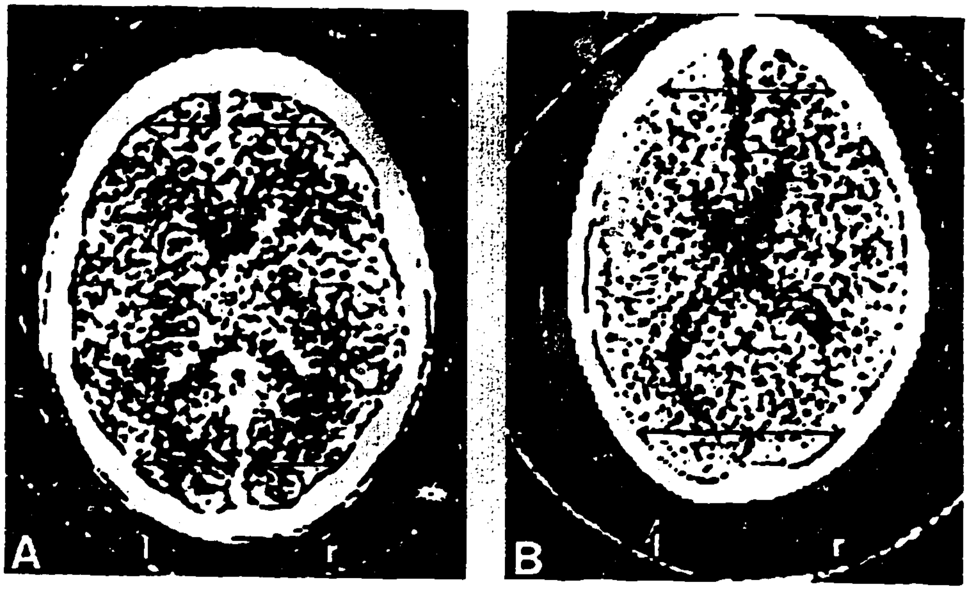


Figure 8. CT scans. The central arrows mark the inter-hemispheric region. Note the wider right frontal and left posterior parietal-occipital areas. a. Note the slight left occipitopetalia; b. Note the left occipitopetalia and slight right frontopetalia (patient B has slightly enlarged ventricles). Reprinted with publisher's permission, from LeMay and Geschwind (1978). Asymmetries of the human cerebral hemispheres. In Caramazza and Zurif (Eds.) *Language acquisition and language breakdown*. Baltimore, MD: Johns Hopkins University Press. Copyright ©1978 by the Johns Hopkins University Press.

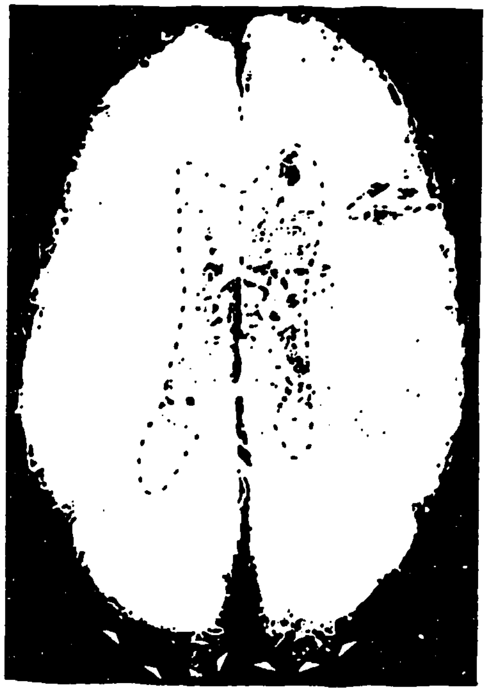


Figure 9. X-ray of a brain in which the blood vessels were injected with an opaque substance post-mortem. The tips of the occipital lobes are shown by white arrowheads. The ventricular outlines are shown by interrupted dark lines. The frontal and central portions of the right hemisphere are wider than the left. Note also the right frontopetalia and left occipitopetalia. Reprinted with permission of publisher, from LeMay (1976) *Annals of the New York Academy of Sciences, 280*, 349-366. Copyright ©1976 by the New York Academy of Sciences.

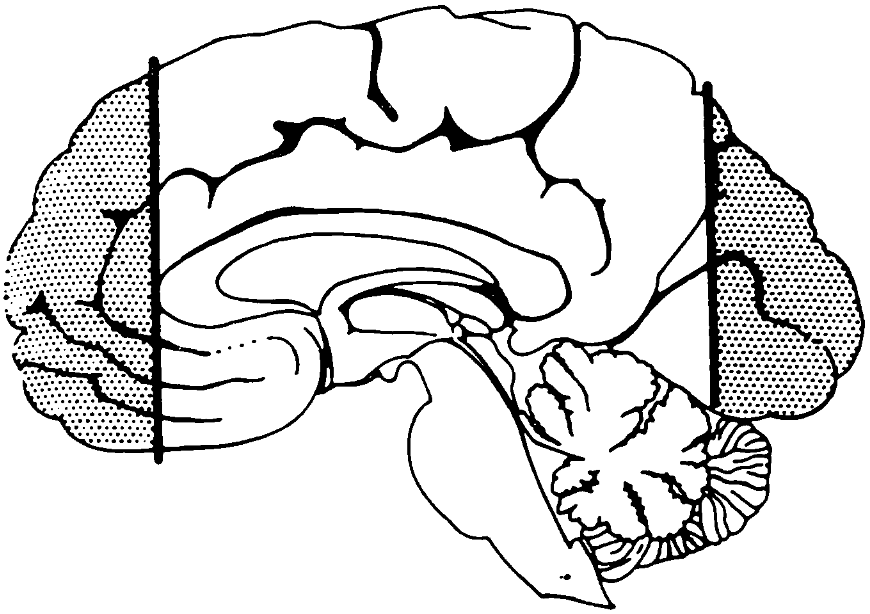


Figure 10. Sagittal section of the adult brain. Stippled areas were measured volumetically. The anterior region was larger on the right, and the posterior region was larger on the left, in the majority of brains studied (taken from the Yakovlev collection). Reprinted with permission of the publisher, from Weinberger, Luchins, Morihisa, and Wyatt (1982) *Annals of Neurology, 11*, 97-100. Copyright ©1982 by Little, Brown & Co.

Figure 11. Photograph of superior surface of a 32-week-old fetal brain showing a slight right frontopetalia and a more striking left occipitopetalia. Reprinted with permission of publisher, from LeMay and Geschwind (1978). Asymmetries of the human cerebral hemispheres. In Caramazza and Zurif (Eds.) *Language acquisition and language breakdown*. Baltimore, MD: Johns Hopkins University Press. Copyright ©1978 by Johns Hopkins University Press.

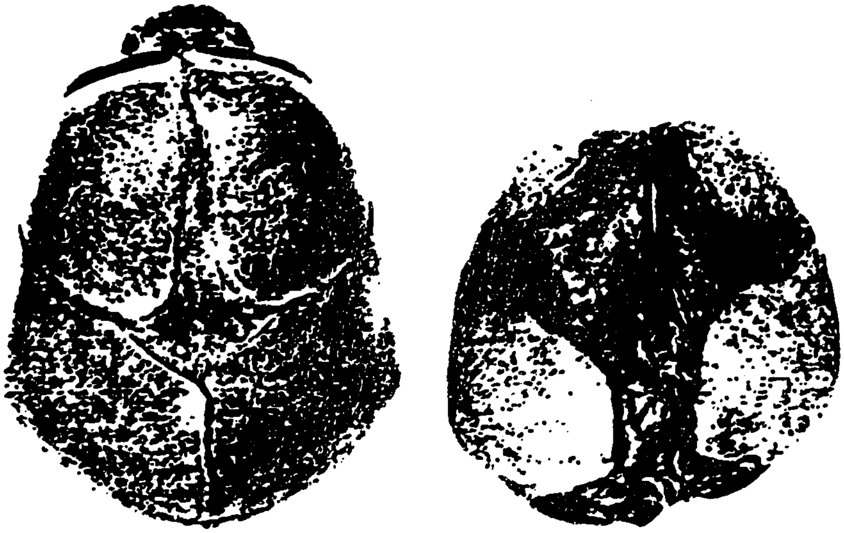


Figure 12. a. Fetal skull. The bone over the right frontal region, and the coronal suture, forehead, and lower rim of the orbit are farther forward than on the left side. The vault extends slightly more posteriorly on the left. b. Upper surface of the skull of a young fetus. The fetus probably had hydrocephalus, but again note the forward position of the right frontal region, the posterior extension of the left hemisphere beyond the right, and the positions of the bony islands of the developing vault. Reprinted with permission of publisher, from LeMay (1984). Radiological, developmental, and fossil asymmetries. In Geschwind and Galaburda (Eds.), *Cerebral dominance: The biological foundations*. Cambridge, MA: Harvard University Press.

Morphologic asymmetries as evidence of a lateralized neuroembryologic gradient. The argument put forth in this chapter is that these morphological asymmetries can be read as a record of a right-to-left gradient in the embryologic emergence of the cerebral hemispheres. This proposal depends on several considerations: First, the brain develops in a general anterior-to-posterior direction. Second, this general gradient is complicated by interaction with other developmental gradients, in the ventro-dorsal and primary → secondary → tertiary dimensions.[1] Third, it assumes (based on reasoning and some indirect evidence to be presented a bit later) that earlier onset in the formation and growth of a given region of telencephalon will strongly tend to result in a larger volume of that region (e.g., greater hemispheric width, but not necessarily a larger

---

[1] There is also a general mediolateral gradient, but Jacobson (1978) states that it is less consistent than the other gradients, i.e., in numerous regions it reverses to a lateromedial gradient or else there is no obvious gradient in either direction along this dimension. Furthermore, it is difficult to conceptualize a vector of greater than 3 dimensions cutting through time without some visualization aid such as computer animation modeling. Therefore, the mediolateral dimension will not be considered further in this discussion of the proposed model for hemisphere growth.

measurement of cortical surface area or gray matter), relative to the homologous contralateral region. This hypothesis of a right-to-left growth gradient is consistent with recent evidence that in fetal development, the major (primary region) fissures appear 1-2 weeks earlier on the right hemisphere than on the left (Dooling, et al., 1983).

The current proposal differs in one crucial respect from earlier proposals of a right-to-left or left-to-right gradient in brain growth. The earlier models assumed or implied that the earlier-maturing hemisphere would show advanced development over the other hemisphere *throughout its extent*. If we combined that assumption with the assumption stated in the previous paragraph that there should be greater volume for earlier-emerging regions, the resulting prediction would be that the earlier-emerging hemisphere should end up larger overall, which is simply not the case. In fact, the simple-minded lateral gradients of earlier models would have to posit some sort of post-hoc explanation (like Ptolemy's planetary epicycles) for the fact of larger right-hemisphere volume in the frontal regions but larger left-hemisphere volume in the posterior regions.

The present proposal refers to a dynamic, developmental gradient in the lateral right-to-left axis of the embryo; that is, it refers to a shift over time from right to left. The proposal also takes into account the fact that there are growth gradients along the other main axes of embryologic development, i.e., that emergence of the hemispheres takes place *over time in 3-dimensional space*. The right-to-left gradient is only one of several axial growth gradients, and its influence on morphological asymmetries can only be understood in the context of the other gradients. In other words, we need to conceptualize a *growth vector* cutting through at least three dimensions over time; this vector represents a *wave of leading growth activity*. There are actually three other developmental gradients that should be accounted for in this hypothesized growth vector (see Figure 13 for general reference on human fetal brain growth). The anteroposterior gradient refers to the general direction of growth from the frontal region toward the occipital region (e.g., Gilles, Leviton, & Dooling, 1983). This gradient, however, is complicated by a growth gradient that moves in the following direction: from primary motor and sensory zones to secondary association areas, and finally to tertiary association zones. This is important to keep in mind, because although the motor and premotor zones of the frontal lobe are early-emerging, the forward extension of the *pre*frontal area is one of the last developments of the hemispheres, and is a tertiary association area (e.g., Rabinowicz, 1979; Yakovlev & Lecours, 1967). The third developmental gradient to consider is the ventrodorsal gradient (Jacobson, 1978), from basal regions toward upper or superior regions. The ventrodorsal gradient is distorted, however, in hemispheric development by the fact that the hemispheres develop radially around the core of the basal ganglia and the insula (considered to be the basal or floor region of the hemisphere), moving in an inverted C-shaped direction, folding down and under around the back of the head and then turning forward to form, respectively, the occipital and temporal poles.

The resulting prediction of a 3-D growth vector starts with an earlier emergence of right primary motor (and premotor) and sensory regions that lie more frontal and ventral (initially, at least). With concurrent developmental shifts along all gradients, the advancing wave of the growth vector would then proceed toward *earlier* emergence of the *left* side for tertiary association regions (including Wernicke's area) that lie more dorsal and posterior (e.g., superior parietal), again at least initially (but recall the late and presumed leftward bias in protrusion of prefrontal regions). At some point midway between these extremes, growth should reach equilibrium between the two sides (possibly in secondary association areas).

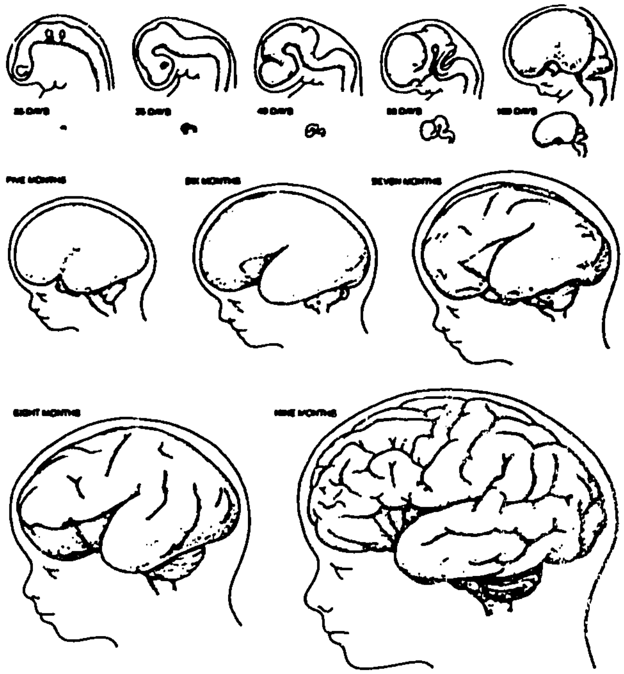


Figure 13. Developing human brain, seen from the left side in a succession of embryonic and fetal stages. The illustrations for the bottom two rows are approximately 4/5 life size, and drawn to scale. Those in the top row are enlarged to show structural details; the insets show $\frac{4}{5}$ life size scale. Reprinted with permission of publisher, from Cowan (1979), *Scientific American, 241*, 112-133. Copyright ©1979 by Scientific American, Inc. All rights reserved.

The effect of this growth vector is a counterclockwise torque evident in the shape of the developing as well as the adult brain. This can only be illustrated here in two dimensions at a time. Figure 14 shows a brain viewed from above (from LeMay, 1976), to illustrate the combined influences of the anteroposterior and right-left gradients, producing a counterclockwise torque, as though some force had molded the brain with a fore-to-aft twist on the left, concurrent with an opposing twist on the right. The counterclockwise torque resulting from the combined effects of the ventrodorsal and right-left gradients is seen in coronal views of fetal brains (from Dooling, Chi, & Gilles, 1983), as seen in Figure 15 (easiest to view in the 34-week brain at top).

The overall effect on the hemispheres is as though some force had twisted the left hemisphere rearward and dorsal, while twisting the right hemisphere forward and ventral. LeMay's (1984) observations of asymmetries in the positions and angles of the central (Rolandic) fissure and the Sylvian fissure are consistent with this image—the Rolandic fissure appears farther forward (and tilted more vertically) on the right hemisphere, even in fetal brains, whereas the Sylvian fissure slants more horizontally (i.e., lower) on the left, with a lower and more posterior endpoint (Sylvian point). The right Sylvian fissure angles more sharply upward, and has a more anterior endpoint.

This model of embryologic hemisphere development leads to several predictions. First, the gross morphologic effect of earlier-emerging right frontal-motor regions may become attenuated

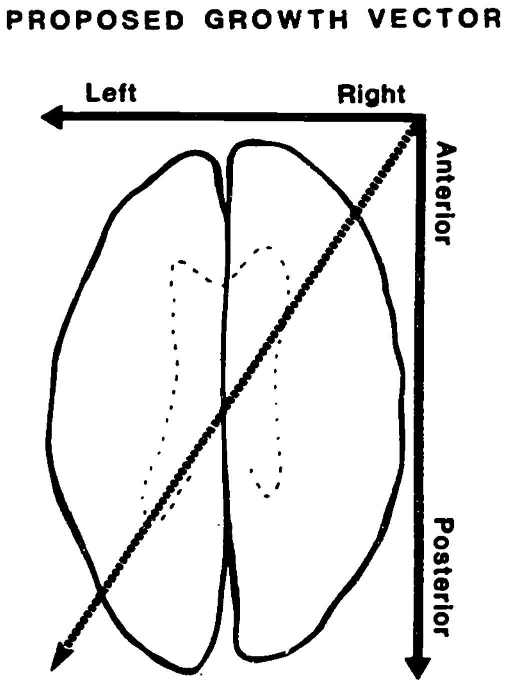


Figure 14. Schematic diagram of the proposed growth vector, proceeding from right anterior to left posterior regions of the hemispheres. This diagram is a simplification of the vector, since it shows only the left-right and anterior-posterior dimensions (it omits the primary → secondary → tertiary dimension and the ventral-dorsal dimension).

by the later, *left*-biased growth of the tertiary association cortex in the prefrontal region. This would contrast with the convergence of the left-side bias for posterior regions and left-side bias in the growth of the posterior tertiary association areas. The result should be a more striking left occipitopetalia than right frontopetalia, at least in adult brains. LeMay's (1976) data are in agreement with this pattern. Moreover, since the left-biased tertiary association areas are late to emerge in development, we should expect to see *greater* evidence of right frontopetalia in fetuses than in adults, but *lesser* left occipitopetalia in fetuses than adults. Again, LeMay's data (1977) are in accord with these predictions. Also in accord is the Wada et al. (1975) finding of greater left-side bias in planum temporale among adult brains than among fetuses.

Another prediction is for a left-side bias in earlier emergence of tertiary sulci and gyri, in the tertiary association regions of prefrontal and posterior cortex. There are no data available on this possibility, since only primary sulci have been carefully mapped out on left vs. right hemispheres in this manner (see Gilles et al., 1983). In fact, the earlier right hemisphere appearance of fissures on the superior temporal surface actually refers to the formation of the transverse (Heschl's) gyrus, or primary auditory cortex, and not to the formation of Wernicke's association area (a tertiary area) itself (Dooling et al., 1983).

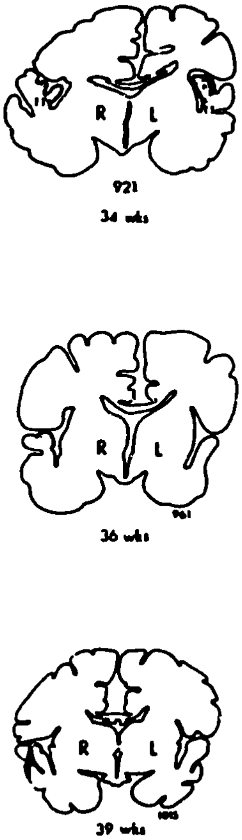


Figure 15. Tracings of photographs of sections in frontal plane, of representative fetal brains of different gestational ages. Reprinted with permission of publisher, from Gilles, Leviton, and Dooling (1983), *The developing human brain.* Boston: John Wright PSG Inc.

Effect of the Growth Vector on Neuronal Organization

The model also carries implications for asymmetries in neuronal organization of cortical areas, and hence for functional development and plasticity of the various regions on the two sides. The impact of the growth vector upon neuronal organization must be understood in the context of the sequential development of the six layers of neocortex and their differing contributions to the development of cortical fissuration and gyration, that is, the development of cortical folding. The five cortical layers that contain actual cell bodies of neurons develop in an inside-out sequence (Figure 16), with the cells of layer 6, the deepest inner layer, reaching their target positions and developing dendritic and axonal connections earliest (e.g., Rakic, 1980). The earliest-developing layers, 5 and 6, contain primarily efferent cells projecting to regions outside of cortex per se, and so give rise to long, myelinated axons, that is, subcortical white matter. Layer 4 is also directly associated with subcortical white matter, because in most regions of cortex it is the layer that receives initial, primary input from afferents to cortex. That is, its cells synapse with incoming long, myelinated axons that arise largely from subcortical areas or from other, relatively distant cortical regions. Thus, the lower layers 4-6 of cortex contribute disproportionately to the "white matter" side of regional measurements of ratios of gray matter (aggregated neural cell bodies and short unmyelinated axons within the cortical mantle itself) to white matter (subcortical myelinated axon bundles) (e.g., Gur et al., 1980; Meyer et al., 1978; McHenry et al., 1978). For those areas with relatively higher proportions of subcortical white matter (low gray/white ratio), there should be a tendency toward greater width and/or volume than in areas with a higher

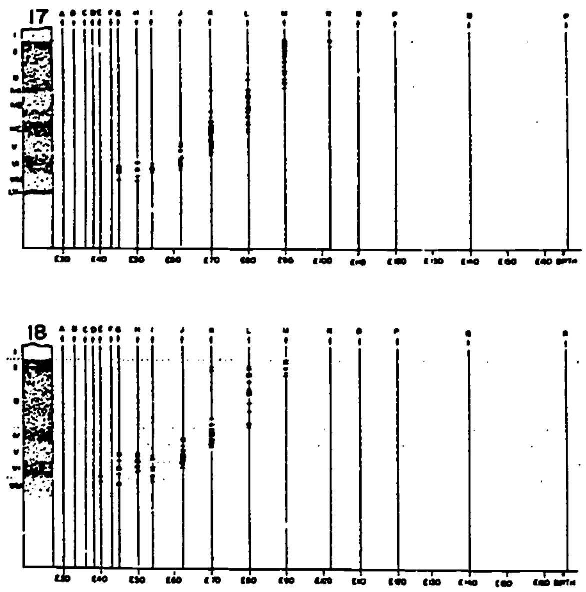


Figure 16. Diagrammatic representation of heavily labeled neurons in the visual cortex of juvenile monkeys that had been injected with [^{3}H]-TdR at selected embryonic days. The top picture shows Brodman's area 17, the bottom shows area 18. On the left of each diagram is a drawing of cortex sections with cresyl-violet staining, with Brodman's layers indicated in Roman numerals. WM: white matter, LV: lateral ventricles. Embryonic days (E) are shown on the X axis, starting with the end of the first fetal month (E27) and ending at term (E165). The vertical lines represent the embryonic days on which subsets of the animals received a pulse of [^{3}H]-TdR. On each vertical line short horizontal markers indicate the positions of all heavily labeled neurons. Since [^{3}H]-TdR labels the cells undergoing mitosis at the time of injection, these diagrams indicate that cortex is built in inside-out order, with layer VI neurons generated earliest and layer II neurons latest. Reprinted with permission from Rakic (1981). Developmental events leading to laminar and areal organization of the neocortex. In Schmitt, Worden, Adelman and Dennis (eds.) *The organization of the cerebral cortex*. Cambridge, MA: MIT Press. Copyright ©1981 by the Massachusetts Institute of Technology.

gray/white ratio, given that white matter makes up more of the bulk of the width and volume of the hemispheres than does cortical gray matter.

Conversely, the cells of the most superficial cellular layer, layer 2, form latest. This latest-developing layer contains neurons that make predominantly local connections with other cells lying within the nearby cortical layers. That is, it does not contribute substantially to subcortical white matter, and thereby contributes relatively more to cortical gray matter. Layer 3, which is next-to-last in development, also contributes more to gray matter than to subcortical white matter, in that its cells make mostly intracortical connections, including connections with the contralateral hemisphere via the corpus callosum. Thus, layers 2 and 3 contribute disproportionately to the "gray matter" side of the gray/white ratio. Furthermore, the late-developing, superficial layer 2 is primarily responsible for the process of cortical fissuration and gyration, during the period of fetal ontogeny when that layer greatly expands in thickness relative to the lower layers, as a result of its developing dendritic processes and proliferation of glial support cells (Jacobson, 1978). Therefore, later development and higher gray/white ratios should be associated with deeper, denser fissuration but also with *lesser* width and volume in that area of the hemisphere, given the argument made in the previous paragraph.

The growth vector would be expected to have the following influence on the development of cortical gray and white matter: for earlier-developing cortical regions relatively greater growth emphasis would be seen in the earlier-emerging deeper cell layers, which should result in smaller ratios of gray matter to white matter and greater regional width/volume. Figure 17, taken from Gur and colleagues (1980), illustrates that indeed the gray-white ratio is smaller in the right hemisphere for at least the primary-motor and -sensory cortical regions. In contrast, later-maturing regions should reflect greater growth emphasis of the later-emerging, more superficial cortical layers, 2 and perhaps 3. This should yield smaller width/volume but a more deeply fissurated region with higher gray/white ratios, which may imply a somewhat higher degree of local intracortical organization of neuronal connections. In accordance, there is a higher gray/white ratio in the left hemisphere for motor, premotor, and primary sensory areas (Gur et al., 1980), consistent with Semmes' (1968) claim that the left hemisphere is focally organized, whereas the right is diffusely organized (based on her studies of somatosensory deficits in unilateral brain-damaged patients). Correspondingly, it is those same areas in which LeMay (1976) found smaller hemispheric width on the left, and Weinberger et al. (1982) measured a smaller volume on the left. Also, at least for the anterior speech area in the premotor region (Broca's area), there is deeper fissuration on the left (Falzi et al., 1982).

Interestingly, and consistent with the predictions of the growth vector model, there is a higher gray/white ratio in the right hemisphere for tertiary association areas in posterior cortex (temporal and parietal association areas, numbered 5 and 6 in Figure 17). Recall the model's prediction that tertiary association areas in posterior cortex would emerge later in the right than in the left hemisphere. Also, in posterior regions, LeMay (1976) found lesser hemispheric width, and Weinberger et al. (1982) found smaller volume, on the right side.

As for development of dendritic processes and of neuronal connections, the influence of the growth vector should be a bias toward later-emerging characteristics in the later-maturing regions of the cortex. Arnold Scheibel (1984) has thus far provided the only data relevant to this issue, and it appears to corroborate one part of the hypothesis. In layer 3 of the anterior speech region in the left hemisphere (LOP in Figure 18), the dendritic trees of the neurons show relatively greater elaboration of the later-emerging dendritic features, such as proportionally more higher-order branching points, than is true of the homologous region in the right hemisphere (ROP in Figure 18).

Implications for Functional Development

This morphological growth vector should also have implications for the development of perceptual and cognitive functions, and for asymmetrical patterns in developmental plasticity. Gross morphological asymmetries of the adult brain, both in vivo (Ratcliff, Dila, Taylor, & Milner, 1980) and postmortem (Witelson, 1983), are associated, in fact, with at least some measures of functional asymmetries, notably speech lateralization and handedness. Development of regional functional maturity should proceed according to the same growth vector as already outlined for

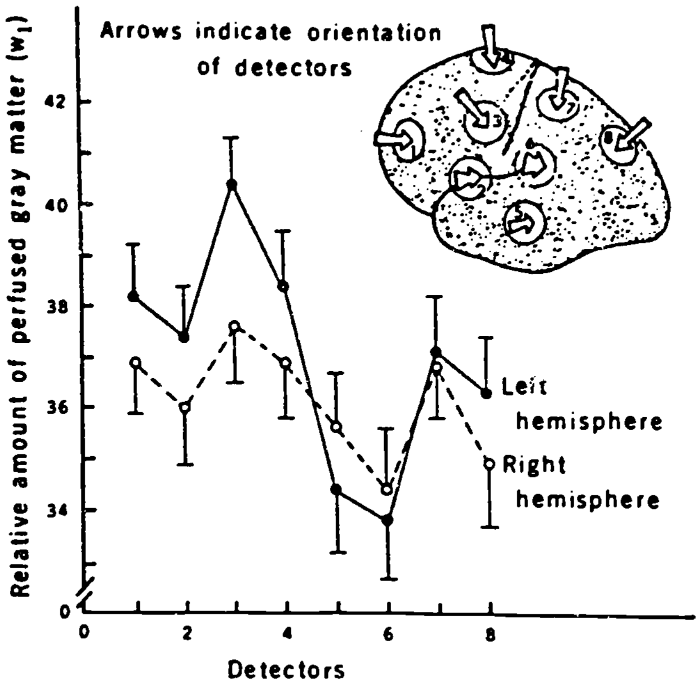


Figure 17. Relative amount (± standard errors) of perfused gray matter (w_1) plotted separately for eight homologous regions in the left and right hemispheres. Reprinted with permission of the authors and publisher, from Gur, Packer, Hungerbuhler, Reivich, Obrist, Amarnek, and Sackeim (1980) *Science, 207* (14 March), 1226-1228. Copyright ©1980 by the AAAS.

the morphological development of the hemispheres. Specifically, in the frontal motor and premotor regions, and in primary sensory regions, right hemisphere functions should mature earlier than the left hemisphere functions in homologous areas. This is supported by the Best et al. (1982) finding of a right-hemisphere advantage for memory-based discrimination of music notes by 2 months, but no evidence of the homologous left-hemisphere ability for discrimination of speech syllables until 3 months[2] (also compare the speech REA in Entus' $2\frac{1}{2}$ month olds, 1977, with the

---

[2] This is not meant to imply that music discrimination depends on *frontal* cortex in the right hemisphere. Presumably, it would rely, in part, on primary auditory cortex (Heschl's gyrus) in the right hemisphere (see Shankweiler, 1966); the growth vector model does assume rightward bias in development of primary sensory areas. Indeed, a double Heschl's gyrus is more often encountered in the right hemisphere than in the left (Geschwind & Levitsky, 1968; but see also Witelson & Pallie, 1973). In addition, recall that the experimental paradigm required that the infants use short-term memory in order to make their discriminations. Short-term memory ability is largely dependent on hippocampus, which is an early-emerging, ventrally-located structure within the telencephalic hemispheres (Gilles, Leviton & Dooling, 1983). The growth vector model would also, thus, predict a rightward bias in a hippocampal maturation.

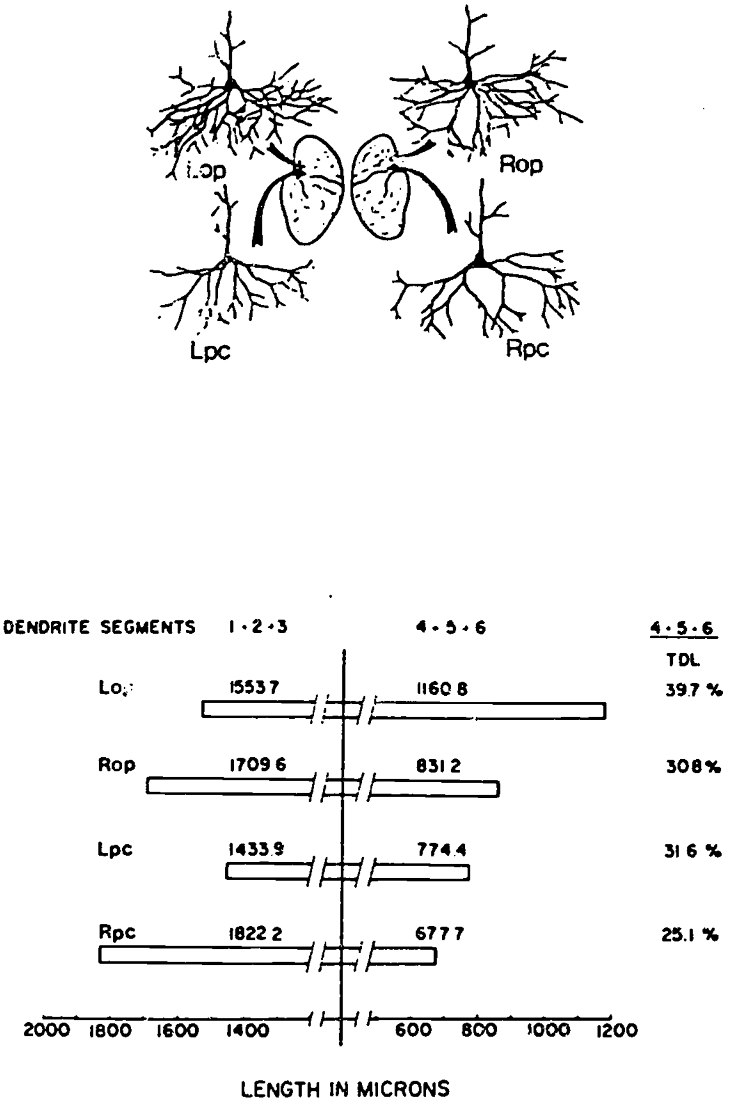


Figure 18. The top of the figure shows a somewhat schematicized drawing of typical dendritic ensembles from cells of the left and right frontal operculum and precentral regions in human cortex. Note the increased number of higher-order segments in left operculum (Lop) compared to the other three areas, and the relatively greater length of second- and third-order branches in the right operculum (Rop) and right precentral (Rpc) areas. The bottom of the figure shows dendritic length and proportion of the dendritic ensemble made up of lower-order (1,2,3) and higher-order (4,5,6) dendritic segments in left opercular (Lop), left precentral (Lpc), right opercular (Rop), and right precentral (Rpc) areas. The column of figures on the extreme right shows the percentage of total dendritic length (Tdl) occupied by higher-order dendrites in each region. Reprinted with permission of the publisher from Scheibel (1984). A dendritic correlate of human speech. In Geschwind and Galaburda (Eds.), *Cerebral dominance:*

lack thereof in Vargha-Khadem & Corballis' 2 month olds, 1979). Moreover, it is consistent with numerous reports in the language acquisition literature that children comprehend and produce the emotional intonational properties of language earlier than they comprehend and produce words and word combinations (e.g., Lewis, 1936). Research with both brain-damaged and neurologically intact adults indicates that the right hemisphere is specialized for affective, or emotional, prosodic aspects of spoken utterances both in perception (e.g., Haggard & Parkinson, 1971; Heilman, Scholes & Watson, 1975; Ley & Bryden, 1982; Papanicolaou, Levin, Eisenberg, & Moore, 1983; Safer & Leventhal, 1977; Tucker, Watson & Heilman, 1976; Wechsler, 1973) and in production (e.g., Benowitz, Bear, Rosenthal, Mesulam. Zaidel, & Sperry, 1983; Kent & Rosenbeck, 1982; Ross, 1981, 1984; Ross & Mesulam, 1979; Tucker et al., 1976). In addition, according to the neuroembryological principle that later-maturing brain structures generally show greater functional plasticity than earlier-maturing structures (Jacobson, 1978), the relative degree of plasticity of various cortical regions should be inversely correlated with their rate of maturation.

The proposed pattern of asymmetry in functional development should correspond to a greater plasticity of the later maturing, left frontal-motor, -premotor and primary sensory regions, relative to the plasticity of the homologous right-hemisphere regions. Conversely, there should be later development, and greater plasticity, on the right relative to the left side for posterior tertiary association area functions. Unfortunately, there are no published data relevant to this issue. Moreover, it will be difficult, to say the least, to match right- and left-hemisphere skills in terms of cortical areas involved, as well as for level of cognitive complexity and/or difficulty, in order to compare their development in normal children and their plasticity in brain-damaged children. For example, one might wish to compare the development and plasticity of reading ability, which depends on a tertiary association region of left hemisphere in adults (angular gyrus), versus that of complex spatial and face-recognition abilities, which depend in adult's upon right tertiary association areas (parietal and inferotemporal). In our society, normal children begin reading around 6 years of age, but do not develop the ability to solve complex mazes, or a right-hemisphere configurational-processing superiority for recognition of unfamiliar faces, until about 10 years of age (e.g., Kohn & Dennis, 1974; Levine, 1985). If we could assume that the criteria of comparability were met by these findings, the model would then predict greater ability of the *right* hemisphere to acquire reading skills, relative to the left hemisphere's ability to acquire spatial/facial skills, following early unilateral damage. However, there are several inherent problems. The comparability of cognitive levels for these skills is uncertain: the onset ages may be, at least in part, artifacts of our educational system. Finally, acquisition of a skill may call upon different cortical regions in the child than those that underly the execution of the already-acquired skill in the adult (see Kirk, 1985).

Individual differences in functional asymmetries

The growth vector is very likely influenced by hormonal and genetic factors, given the current understanding in neuroembryology that developmental gradients involve some sort of gradient(s) in biochemical influences on the prenatal guidance of neuronal migration and development of neuronal connections (Jacobson, 1978; Sperry, 1963). The possibility of hormonal influences on growth gradients may aid our understanding of the development of sex differences in brain organization and cognitive functions[3] (see also the chapter by Netley, this volume*), which may

---

[3] The effect of sex differentiation upon the growth vector may also be related to sex differences in morphological asymmetries for other body parts, e.g., the hands and feet (Levy & Levy, 1978;

be mediated by sex differences in maturation rates (see Newcombe & Bandura, 1983; Waber, 1976, 1977, 1979a; but see also Rovet, 1983; Waber, Mann, Merola, & Moylan, 1985). These sex differences are most apparent in the extreme cases of sexual anomalies such as Turner's syndrome (XO). Turner's syndrome is associated with large deficits in spatial abilities (e.g., Waber, 1979b), an increase in the rate of prenatal development (Netley & Rovet, 1981), and maldevelopment of tertiary association areas in the right hemisphere (Christensen & Nielsen, 1981). Sex differences in brain organization and function are also apparent in anomalies such as supernumerary-X syndrome (XXX and XXY), which is linked with low verbal relative to spatial abilities, and slow prenatal growth rates (Netley, this volume*; Netley & Rovet, 1981, 1982, 1983; Rovet & Netley, 1983).

Hormonal and genetic influences on the growth vector may also be involved in heritable learning disorders. For example, Galaburda and colleagues (Galaburda & Eidelberg, 1982; Galaburda & Kemper, 1979; Galaburda, Rosen, Sherman, & Assal, 1986; Galaburda, Sherman, Rosen, Aboitiz, & Geschwind, 1985) found symmetrical plana temporale in the brains of all four male dyslexics and the one female dyslexic that they have studied post-mortem. The same brains showed abnormalities in cortical neuronal organization, which predominated in the left hemisphere perisylvian regions. The nature of the latter abnormalities led them to posit a neuroembryological disturbance in the prenatal migration of neurons during mid-gestation. Hormonal/genetic effects on the growth vector may be particularly relevant to understanding observed sex and handedness biases in the incidence of learning disabilities, for example, the suggestion that testosterone's effect on brain development forms the basis for the higher incidence of learning disabilities among males and left-handers (Geschwind & Behan, 1982, 1984; Marx, 1982).

Conclusion

So now, what of the original question: Is the pattern of cerebral asymmetry developmentally invariant, or do functional asymmetries develop? Whether we refer to evidence of functional and/or structural asymmetries, even in very early development the extant data support "developmental invariance" (see also Kinsbourne, 1975; Witelson, 1985). Yet this does not necessarily imply that *nothing* is changing or developing. The timeless constancy of cerebral asymmetries coexists with continuous developmental change at many levels. Lateralized perceptual and cognitive functions *do* undergo developmental change (e.g., the child's language and spatial skills change both qualitatively and quantitatively), plasticity of function also undergoes developmental change (see Witelson, 1985), and the cerebral hemispheres supporting these abilities undergo change themselves (e.g., increases in dendritic arborization, neuronal connections, neurotransmitter functions, glial support cells, and myelinization).

According to developmental biology, change and constancy are co-determinants of developmental growth in a biological system. The structural and functional properties of the two cerebral hemispheres do change developmentally, but always in different manners because they develop within the context of an ever-present lateralization of functions, which is continuous with a lateralized gradient of neuronal differentiation and maturation. The argument presented in this chapter is that the normal direction of this lateralizing gradient is from right to left, and that it interacts with gradients in the other main axes of embryologic development to result in a 3D

Means & Walters, 1982); the latter asymmetries in turn appear to be traceable to prenatal development (Mittwoch, 1977).

diagonal growth vector from right frontal-motor and primary sensory areas to left-posterior and tertiary association areas. This growth pattern has implications for hemispheric and regional differences in gross morphology and neuronal organization, as well as for differences in plasticity and in maturation of perceptual and cognitive functions.

References

Best, C. T. (1978). *The role of consonant and vowel acoustic features in infant cerebral asymmetries for speech perception.* Unpublished doctoral dissertation, Michigan State University.

Best, C. T., Hoffman, H., & Glanville, B. B. (1982). Development of infant ear asymmetries for speech and music. *Perception & Psychophysics, 31*, 75-85.

Benowitz, L. I., Bear, D. M., Rosenthal, R., Mesulam, M. M., Zaidel, E., & Sperry, R. W. (1983). Hemispheric specialization and nonverbal communication. *Cortex, 19*, 5-12.

Bever, T. G. (1978). Broca and Lashley were right: Cerebral dominance is an accident of growth. In D. Caplan (Ed.), *Biological studies of mental processes* (pp. 186-230). Cambridge, MA: MIT Press.

Broca, P. (1865). Sur la faculté du langage articulé. *Bulletin of Social Anthropology, 6*, 493-494.

Brown, J. W., & Jaffe, J. (1975). Note: Hypothesis on cerebral dominance. *Neuropsychologia, 13*, 107-110.

Chi, F. G., Dooling, E. C., & Gilles, F. H. (1977). Left-right asymmetries of the temporal speech areas of the human fetus. *Archives of Neurology, 34*, 346-348.

Christensen, A-L., & Nielsen, J. (1981). A neuropsychological investigation of 17 women with Turner's syndrome. In W. Schmid & J. Nielsen (Eds.), *Human behavior and genetics* (pp. 151-166). Amsterdam: Elsevier/North-Holland.

Corballis, M. C., & Morgan, M. J. (1978). On the biological basis of human laterality: I. Evidence for a maturational left-right gradient. *The Behavioral and Brain Sciences, 2*, 261-267.

Cowan, M. W. (1979). The development of the brain. *Scientific American, 241*, 112-133.

Crowell, D. H., Jones, R. H., Kapuniai, L. E., & Nakagawa, J. K. (1973). Unilateral cortical activity in newborn humans: An early index of cerebral dominance? *Science, 180*, 205-208.

Cutting, J. E. (1974). Two left hemisphere mechanisms in speech perception. *Perception & Psychophysics, 16*, 601-612.

Darwin, C. J. (1971). Ear differences in the recall of fricatives and vowels. *Quarterly Journal of Experimental Psychology, 23*, 46-62.

Dooling, E. C., Chi, J. G., & Gilles, F. H. (1983). Telencephalic development: Changing gyral patterns. In F. H. Gilles, A. Leviton, & E. C. Dooling (Eds.), *The developing human brain: Growth and epidemiologic neuropathy* (pp. 94-104). John Wright, PSG Inc.

Emde, R. N., & Robinson, J. (1979). The first two months: Recent research in developmental psychobiology and the changing view of the newborn. In J. Noshpitz & J. Call (Eds.), *Handbook of child psychiatry. Vol. 1: Development*. New York: Basic Books.

Entus, A. K. (1977). Hemispheric asymmetry in processing of dichotically presented speech and nonspeech sounds by infants. In S. Segalowitz & F. A. Gruber (Eds.) *Language development and neurological theory* (pp. 63-73). New York: Academic Press.

Falzi, G., Perrone, P., & Vignolo, L. A. (1982). Right-left asymmetry in anterior speech region. *Archives of Neurology, 39*, 239-240.

Friedes, D. (1977). Do dichotic listening procedures measure lateralization of information processing or retrieval strategy? *Perception & Psychophysics, 21*, 259-263.

Galaburda, A. M. (1984). Anatomical asymmetrics. In N. Geschwind & A. M. Galaburda (Eds.), *Cerebral dominance: The biological foundations* (pp. 11-25). Cambridge, MA: Harvard University Press.

Galaburda, A. M., & Eidelberg, D. (1982). Symmetry and asymmetry and asymmetry in the human posterior thalamus: II. Thalamic lesions in a case of developmental dyslexia. *Archives of Neurology, 39*, 333-336.

Galaburda, A. M., & Kemper, T. L. (1979). Cytoarchitectonic abnormalities in developmental dyslexia: A case study. *Annals of Neurology, 6*, 94-100.

Galaburda, A. M., LeMay, M., Kemper, T. L., & Geschwind, N. (1978). Right-left asymmetries in the brain. *Science, 199*, 852-856.

Galaburda, A. M., Rosen, G. D., Sherman, G. F., & Assal, F. (1986). Neuropathological findings in a woman with developmental dyslexia. *Annals of Neurology, 20*, 170.

Galaburda, A. M., Sanides, F., & Geschwind, N. (1978). Cytoarchitectonic left-right asymmetries in the temporal speech region. *Archives of Neurology, 35*, 812-817.

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F., & Geschwind, N. (1985). Developmental dyslexia: Four consecutive patients with cortical anomalies. *Annals of Neurology, 18*, 222-233.

Geschwind, N., & Behan, P. O. (1982). Left-handedness: Association with immune disease, migraine, and developmental learning disorder. *Proceedings of the National Academy of Science, 79*, 5079-5100.

Geschwind, N., & Behan, P. O. (1984). Laterality, hormones, and immunity. In N. Geschwind & A. M. Galaburda (Eds.), *Cerebral dominance: The biological foundations* (pp. 211-224). Cambridge, MA: Harvard University Press.

Geschwind, N., & Levitsky, W. (1968). Human brain: Left-right asymmetries in temporal speech region. *Science, 161*, 136-189.

Gilles, F. H., Leviton, A., & Dooling, E. C. (1983). *The developing human brain: Growth and epidemiologic neuropathology*. Boston: John Wright, PSG, Inc.

Glanville, B. B., Best, C. T., & Levenson, R. (1977). A cardiac measure of cerebral asymmetries in infant auditory perception. *Developmental Psychology, 13*, 54-59.

Graham, F. K., Leavitt, L. A., Strock, B. D., & Brown, J. W. (1978). Precocious cardiac orienting in a human anencephalic infant. *Science, 199*, 322-324.

Gur, R. C., Packer, I. K., Hungerbuhler, J. P., Reivich, M., Obrist, W. D., Amarnek, W. S., & Sackeim, H. A. (1980). Differences in the distribution of gray and white matter in human cerebral hemispheres. *Science, 207*, 1226-1228.

Haggard, M. P., & Parkinson, A. M. (1971). Stimulus and task factors as determinants of ear advantages. *Quarterly Journal of Experimental Psychology, 23*, 168-177.

Heilman, J. M., Scholes, R., & Watson, R. T. (1975). Auditory affective agnosia: Disturbed comprehension of affective speech. *Journal of Neurology, Neurosurgery and Psychiatry, 38*, 69-72.

Jacobson, M. (1978). *Developmental neurobiology*. New York: Plenum Press.

Keller, R. (1942). The asymmetry of the human body. *CIBA Symposium, 3*, 1126-1127.

Kemper, T. L. (1984). Asymmetrical lesions in dyslexia. In N. Geschwind & A. M. Galaburda (Eds.), *Cerebral dominance: The biological foundations* (pp. 75-89). Cambridge, MA: Harvard University Press.

Kent, R. D., & Rosenbeck, J. C. (1982). Prosodic disturbance and neurologic lesion. *Brain and Language, 15*, 259-291.

Kinsbourne, M. (1975). The ontogeny of cerebral dominance. *Annals of the New York Academy of Sciences, 263*, 244-250.

Kirk, V. (1985). Hemispheric contributions to the development of graphic skill. In C. T. Best (Ed.), *Hemispheric function and collaboration in the child* (pp. 193-228). New York: Academic Press.

Kohn, B., & Dennis, M. (1974). Selective impairments of visuospatial abilities in infantile hemiplegics after right cerebral hemi-decortication. *Neuropsychologia, 12*, 505-512.

LeMay, M. (1976). Morphological cerebral asymmetries of modern man, fossil man, and nonhuman primate. *Annals of the New York Academy of Science, 280*, 349-366.

LeMay, M. (1977). Asymmetries of the skull and handedness. *Journal of Neurological Science, 32*, 243-253.

LeMay, M. (1984). Radiological, developmental, and fossil asymmetries. In N. Geschwind & A. M. Galaburda (Eds.), *Cerebral dominance: The biological foundations* (pp. 26-42). Cambridge, MA: Harvard University Press.

LeMay, M., & Geschwind, N. (1978). Asymmetries of the human cerebral hemispheres. In A. Caramazza & E. B. Zurif (Eds.), *Language acquisition and language breakdown* (pp. 311-328). Baltimore, MD: Johns Hopkins Press.

LeMay, M., & Kido, D. K. (1978). Asymmetries of the cerebral hemispheres on computed tomograms. *Journal of Computer-Assisted Tomography, 2*, 471-476.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Levine, S. C. (1985). Developmental changes in right hemisphere involvement in face perception. In C. T. Best (Ed.), *Hemispheric function and collaboration in the child* (pp. 157-191). New York: Academic Press.

Levy, J., & Levy, J. M. (1978). Human lateralization from head to foot: Sex-related factors. *Science, 200*, 1291-1292.

Lewis, M. M. (1936). *Infant speech: A study of the beginnings of language*. New York: Harcourt Brace.

Ley, R. G., & Bryden, M. P. (1982). A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content. *Brain and Cognition, 1*, 3-9

MacKain, K. S., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left hemisphere function. *Science, 219*, 1347-1349.

Marx, J. L. (1982). Autoimmunity in left-handers. *Science, 217*, 141-144.

McHenry, L. C., Merory, J., Bass, E., Stump, D. A., Williams, R., Witcofski, R., Howard, G. & Toole, J. F. (1978). Xenon-133 inhalation method for regional cerebral blood flow measurements: Normal values and test-results. *Stroke, 9*, 396-399.

Means, L. W., & Walters, R. E. (1982). Sex, handedness and asymmetry of hand and foot length. *Neuropsychologia, 20*, 715-719.

Meyer, J. S., Ishihara, N., Deshmukh, V. D., Naritomi, H., Sakai, F., Hsu, M-C., & Pollack, P. (1978). Improved method for noninvasive measurement of regional cerebral blood flow by 133Xenon inhalation. Part I: Description of method and normal values obtained in healthy volunteers. *Stroke, 9*, 195-205.

Mittwoch, U. (1977). To be born right is to be born male. *New Scientist, 73*, 74-76.

Molfese, D. L., Freeman, R. B., & Palermo, D. S. (1975). The ontogeny of brain lateralization for speech and nonspeech stimuli. *Brain and Language, 2*, 356-368.

Molfese, D., & Molfese, V. J. (1985). Electrophysiological indices of auditory discrimination in newborn infants: The basis for predicting later language development? *Infant Behavior and Development, 8*, 197-211.

Netley, C., & Rovet, J. (1981). *Prenatal growth rate and hemispheric organization*. Paper presented at meeting of the International Neuropsychology Society. Atlanta, GA, February.

Netley, C., & Rovet, J. (1982). Verbal deficits in children with 47,XXY and 47,XXX karyotypes: A descriptive and experimental study. *Brain and Language, 17*, 58-72.

Netley, C., & Rovet, J. (1983). Relationships among brain organization, maturation rate, and the development of verbal and nonverbal ability. In S. Segalowitz (Ed.), *Language functions and brain organization* (pp. 245-265). New York: Academic Press.

Newcombe, N., & Bandura, M. (1983). Effect of age at puberty on spatial ability in girls. *Developmental Psychology, 19*, 215-224.

Orton, S. T. (1937). *Reading, writing, and speech problems in children*. New York: W. W. Norton & Co.

Papanicolaou, A. C., Levin, H. S., Eisenberg, H. M., & Moore, B. D. (1983). Evoked potential indices of selective hemispheric engagement in affective and phonetic tasks. *Neuropsychologia, 21*, 401-405.

Rabinowicz, T. (1979). The differentiate maturation of the human cerebral cortex. In F. Falkner & J. M. Tanner (Eds), *Human Growth. Vol. 3: Neurobiology and nutrition* (pp. 97-123). New York: Plenum Press.

Rakic, P. (1981). Developmental events leading to laminar and areal organization of the neocortex. In F. O. Schmitt, F. G. Worden, G. Adelman, & S. G. Dennis (Eds.), *The organization of cerebral cortex* (pp. 7-28). Cambridge, MA: MIT Press.

Ratcliff, G., Dila, C., Taylor, L., & Milner, B. (1980). The morphological asymmetry of the hemispheres and cerebral dominance for speech: A possible relationship. *Brain and Language, 11*, 87-88.

Rose, S. A. (1984). Developmental changes in hemispheric specialization for tactual processing in very young children: Evidence from cross-modal transfer. *Developmental Psychology, 20*, 568-574.

Ross, E. D. (1981). The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology, 38*, 561-569.

Ross, E. D. (1984). Right hemisphere role in language, affective behavior and emotion. *Trends in Neuroscience*, 342-346.

Ross, E. D., & Mesulam, M. M. (1979). Dominant language functions of the right hemisphere: Prosody and emotional gesturing. *Archives of Neurology, 36*, 144-146.

Rovet, J. (1983). Cognitive and neuropsychological test performance of persons with abnormalities of adolescent development: A test of Waber's hypothesis. *Child Development, 54*, 941-950.

Rovet, J., & Netley, C. (1983). The triple-X chromosome syndrome in childhood: Recent empirical findings. *Child Development, 54*, 831-845.

Safer, M. A., & Leventhal, H. (1977). Ear differences in evaluating emotional tone of voice and of verbal content. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 75-82.

Scheibel, A. B. (1984). A dendritic correlate of human speech. In N. Geschwind & A. M. Galaburda (Eds.), *Cerebral dominance: The biological foundations*. Cambridge, MA: Harvard University Press.

Schwartz, J., & Tallal, P. (1980). Rate of acoustic change may underlie hemispheric specialization for speech perception. *Science, 207*, 1380-1381.

Semmes, J. (1968). Hemispheric specialization: A possible clue to mechanism. *Neuropsychologia, 6*, 11-26.

Shankweiler, D. (1966). Effects of temporal-lobe damage on perception of dichotically presented melodies. *Journal of Comparative and Physiological Psychology, 62*, 115-119.

Sidtis, J. J. (1980). On the nature of the cortical function underlying right hemisphere auditory perception. *Neuropsychologia, 18*, 321-330.

Sperry, R. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proceedings of the National Academy of Science, 50*, 703-710.

Studdert-Kennedy, M., & Shankweiler, D. (1970). Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America, 48*, 579-594.

Studdert-Kennedy, M., & Shankweiler, D. (1980). Hemispheric specialization for language processes. *Science, 211*, 960-961.

Taylor, D. C. (1969). Differential rates of cerebral maturation between sexes and between hemispheres. *Lancet, 2*, 140-142.

Tucker, D. M., Watson, R. G., & Heilamn, K. M. (1976). Affective discrimination and evocation in patients with parietal disease. *Neurology, 26*, 354.

Vargha-Khadem, F., & Corballis, M. C. (1979). Cerebral asymmetry in infants. *Brain and Language, 8*, 1-9.

Waber, D. P. (1976). Sex differences in cognition: A function of maturation rate. *Science, 192*, 572-574.

Waber, D. P. (1977). Sex differences and rate of physical growth. *Developmental Psychology, 13*, 29-38.

Waber, D. P. (1979a). Cognitive abilities and sex-related variations in the maturation of cerebral cortical functions. In M. Wittig & A. Peterson (Eds.) *Sex-related differences in cognitive functioning*. New York: Academic Press.

Waber, D. P. (1979b). Neuropsychological aspects of Turner syndrome. *Developmental Medicine and Child Neurology, 231*, 58-70.

Waber, D. P., Mann, M. B., Merola, J., & Moylan, P. M. (1985). Physical maturation rate and cognitive performance in early adolescence: A longitudinal examination. *Developmental Psychology, 21*, 666-681.

Wada, J. A., Clarke, R., & Hamm, A. (1975). Cerebral hemispheric asymmetry in humans. *Archives of Neurology, 32*, 239-246.

Wechsler, A. F. (1973). The effect of organic brain disease on recall of emotionally charged versus neutral narrative text. *Neurology, 23*, 130-135.

Weinberger, D. R., Luchins, D. J., Morihisa, J., & Wyatt, R. J. (1982). Asymmetrical volumes of the right and left frontal and occipital regions of the human brain. *Annals of Neurology, 11*, 97-100.

Weiss, M. S., & House, A. S. (1973). Perception of dichotically presented vowels. *Journal of the Acoustical Society of America, 53*, 51-53.

Witelson, S. F. (1977). Early hemispheric specialization and interhemispheric plasticity: An empirical and theoretical review. In S. Segalowitz & F. A. Gruber (Eds.), *Language development and neurological theory* (pp. 213-287). New York: Academic Press.

Witelson, S. F. (1983). Bumps on the brain: Neuroanatomical asymmetries as a basis for functional asymmetries. In S. Segalowitz (Ed.), *Language functions and brain organization* (pp. 117-144). New York: Academic Press.

Witelson, S. F. (1985). Hemisphere specialization from birth: Mark II. In C. T. Best (Ed.), *Hemispheric function and collaboration in the child* (pp. 33-85). New York: Academic Press.

Witelson, S. F., & Pallie, W. (1973). Left hemisphere specialization for language in the newborn: Neuroanatomical evidence of asymmetry. *Brain, 96*, 641-646.

Yakovlev, P. I., & Lecours, A. R. (1967). The myelogenetic cycles of regional maturation of the brian. In A. Minkowski (Ed.), *Regional development of the brain in early life* (pp. 3-70). Oxford: Blackwell.

THE ROLE OF COARTICULATORY EFFECTS IN THE PERCEPTION OF FRICATIVES BY CHILDREN AND ADULTS[*]

Susan Nittrouer and Michael Studdert-Kennedy[†]

Abstract. *Adult listeners are sensitive to the acoustic variations that result from a speaker's coarticulating (or coproducing) phonetic segments. The present study charts the development of such sensitivity in young children by examining their responses to coarticulatory effects in fricative-vowel syllables. Children at each of the ages 3, 4, 5, and 7 years, and adults identified tokens from a synthetic /ʃ/-/s/ continuum followed by one of four natural vocalic portions: /i/ or /u/, spoken with transitions appropriate for either /ʃ/ or /s/. Children demonstrated larger shifts in fricative phoneme boundaries as a function of vocalic transition than adults, but relatively smaller shifts as a function of vowel quality; responses were less consistent for children than for adults; and differences between children and adults decreased as children increased in age. Overall, these results indicate that perceptual sensitivity to certain coarticulatory effects is present at as young as three years of age. Moreover, the decrease in the effect of the vocalic transitions with age suggests that, contrary to a commonly held view, the perceptual organization of speech may become more rather than less segmental as the child develops.*

---

[*] *Journal of Speech and Hearing Research*, in press

[†] Also Queens College and Graduate Center, City University of New York

Aknowledgment. The work reported here was part of a thesis submitted by the first author to the City University of New York in partial fulfillment of the requirements for the doctoral degree in Speech and Hearing Sciences. The authors are grateful to Arthur Boothroyd and Lawrence Raphael, who served as thesis advisors. Bruno Repp and D. H. Whalen offered valuable suggestions at every stage of this work; Carol Fowler reviewed an earlier draft of the paper; Denise Wright, Melanie Campbell, Lise Jensen, Karen Selvaggi, Patricia Nittrouer, and Gloria Schlisselberg helped with various aspects of the data collection and analysis. We thank each of these individuals. In addition, we are grateful to the children who participated in this study, and to their parents who permitted their participation. This work was supported in part by NICHD Grant HD-01994 to Haskins Laboratories. The writing was completed while the first author was a postdoctoral fellow at Haskins Laboratories on NIH Grant NS-07237, and while the second author was on sabbatical leave as a Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, CA: the support of the City University of New York and the Spencer Foundation is gratefully acknowledged.

HASKINS LABORATORIES: Status Report on Speech Research **SR-88 (1986)**

Introduction

Coarticulation refers to overlapping movements in the production of neighboring or near-neighboring phonetic segments. The acoustic consequences of coarticulation are clearly evident on spectrograms. As early as 1948, Joos described the two major effects. First, lines, vertical to the time axis, cannot be clearly drawn between individual segments. While the center of a segment may be discernible in the acoustic pattern, regions exist between segments that do not seem to belong uniquely to one or the other. Joos termed this difficulty the problem of perceptual segmentation. The second difficulty is that the acoustic attributes of both the center and the adjoining regions of a given segment vary with phonetic context. This finding has come to be known as the problem of perceptual invariance (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Pisoni, 1985).

Despite many dozens of studies in the years since Joos wrote, we still have no generally agreed upon solutions to these problems, and no firm understanding of the function of coarticulation (if any) in listening. Is coarticulation necessary and intrinsic to production, and must a listener therefore draw on the contextually variable information that it carries to recover the phonetic message? Or is coarticulation simply a result of a speaker becoming rapid and skillful? If so, are the acoustic consequences of coarticulation merely noise that a listener filters out?

One way of gaining insight into the perceptual function of coarticulatory effects might be to trace its development in children. This was the goal of the present study. Two different outcomes, with different implications, initially seemed possible. If very young children, whose phonetic and phonological skills were still developing, proved sensitive to the acoustic consequences of coarticulation, we might suspect that coarticulatory information contributed to the recovery of phonetic form. Alternatively, if young children proved relatively insensitive to the acoustic consequences of coarticulation and if their sensitivity increased with age, then we might infer, following Stevens and Blumstein (1978), that speech perception is initially based on invariant characteristics of the signal, and that the ability to use aspects of the signal that vary across contexts is learned through extensive association of this varying information with specific invariant properties.

Researchers interested in listeners' sensitivity to the acoustic consequences of coarticulation have usually varied the phonetic context in which segments are presented for identification. The assumption is then that, if identification varies with phonetic context, listeners are using their knowledge of coarticulation and its acoustic consequences for perception. Although several contextual effects have been investigated, the present experiment focused on just one: the influence of vocalic segments on voiceless fricative identification. We chose fricatives for study, partly because fricative spectra are often said to provide relatively steady-state and invariant perceptual cues, and partly because contextual effects on fricative perception have been quite thoroughly studied in adults.

Adult Studies

Kunisaki and Fujisaki (1977) investigated listeners' abilities to use the acoustic consequences of coarticulation by measuring the effects of vowel context on fricative perception. They prepared a ten-step synthetic fricative continuum (from /ʃ/ to /s/), followed by synthetic /a/ and /u/. These vocalic portions contained first and second formant transitions, but it is unclear from the description whether the transitions were appropriate for /ʃ/ or /s/. Nonetheless, these authors

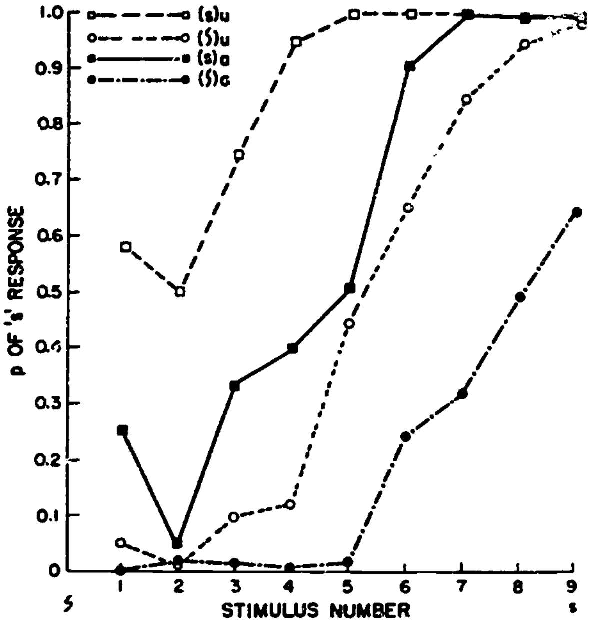


Figure 1. Mean identification functions for adult subjects: the probability estimate of an 's' response to fricative-vowel syllables as a function of fricative pole values in a synthetic series ranging from /ʃ/ to /s/, followed by naturally spoken vowels. Parameters of the curves are vowels and the fricatives (in parentheses) after which they were originally spoken (From data of Mann & Repp, 1980).

demonstrated that the phoneme boundary shifted to a lower value of fricative spectrum (that is, listeners perceived more stimuli as /s/), when the fricative noise was followed by the rounded vowel /u/ than when it was followed by the unrounded vowel /a/. The explanation proposed by Kunisaki and Fujisaki was that, in natural speech, anticipatory liprounding lowers the pole values of the fricative, so that listeners come to expect that /s/ will have lower pole values when the following vowel is rounded than when it is unrounded. In other words, listeners "know" the acoustic results of liprounding and use this tacit knowledge in their judgments of a synthetic series. The contextual effect in perception is thus linked to coarticulation in production.

An alternative explanation for this effect might be based strictly on properties of general audition. For example, fricative noises with relatively low frequency values might be perceived as relatively higher (i.e., more /s/-like) in the /u/ context than in the /a/ context, due to auditory contrast induced by the lower spectrum for /u/ than for /a/.

Mann and Repp (1980) replicated the vowel context effect of Kunisaki and Fujisaki (1977). They also demonstrated an effect of the initial vocalic transition by appending natural vocalic segments /a/ or /u/, originally spoken after either /ʃ/ or /s/, to members of a synthetic /ʃ/-/s/ continuum. Identification functions drawn from data obtained in this experiment are shown in Figure 1. It can be seen that phoneme boundaries shifted as a function of both vowel context and coarticulated formant transition: the likelihood of an 's' response increased both with shifts in the vowel from /a/ to /u/, and with shifts in the fricative from /ʃ/ to /s/ after which the vowel was originally spoken.

However, in a further experiment with flat vocalic formants following the fricative noise, Mann and Repp (1980) found that the vowel context effect virtually disappeared. Thus the question arose as to the actual source of the effect. Did the transition merely provide the coherence between fricative and vowel segments necessary for them to be heard as a single syllable to which both fricative and vowel contribute? Or was the vowel effect the result of the listener using information in the vocalic transitions rather than vowel quality *per se*? Or are transition and vowel context effects separate phenomena?

Evidence existed to suggest that the first possibility provided at least a partial explanation for why the vowel context effect disappeared when transitions were removed. Work by Darwin and Bethell-Fox (1977) and by Dorman, Raphael, and Liberman (1979) had demonstrated the importance of signal continuity in providing phonetic coherence. For example, in the Darwin and Bethell-Fox experiment the perception of a liquid-vowel syllable was changed into the perception of a stop-vowel syllable when fundamental frequency shifted in the region of the segmental boundary. Moreover, Mann and Repp (1980) noted that, in the transitionless stimuli, the fricative and vowel segments sounded segregated.

The second possibility, that the transition was solely responsible for the context effect, was eliminated both by the first experiment of Mann and Repp (1980) and by Whalen (1981). Both studies demonstrated a vowel context effect when all stimuli provided transitional information appropriate for the same fricative. This effect can be seen in Figure 1 (from Mann & Repp, 1980).

Whalen (1981) offered support for the third possibility, that both transitions and vowel quality independently affect the placement of the phoneme boundary. Using synthetic fricative noises from an /ʃ/-/s/ continuum and vocalic segments appropriate for /i,u,ü,ï/ with both /ʃ/ and /s/ transitions, Whalen demonstrated the transition and the vowel context effects with both synthetic and naturally produced vowel segments. Thus any concern that these effects were artifacts of synthetic speech or of a quirk in the particular natural samples used could be dismissed. The fact that these effects were obtained for English speakers using non-English vowels (/ü/ and /ï/) suggests that the effects are not related to specific linguistic experience, but rather to some more basic characteristic of auditory or phonetic perception such as sensitivity to the effects of liprounding on fricative spectra.

Child Studies

Only one study has investigated a phonetic context effect in children (Mann, Sharlin, & Dorman, 1985). In this study, the /ʃ/-/s/ continuum of Mann and Repp (1980) was combined with naturally produced vocalic segments taken from utterances of /ʃeIv/ and /ʃu/. Consequently, although the vowel context effect could be studied, the transition effect could not. Subjects were a) adults, b) 5-year-olds, c) 7-year-olds with age-appropriate speech production skills, and d) 7-year-olds judged to misarticulate fricatives. The resulting identification functions (pooled for each subject group) were similar in one respect for the adults and all three groups of children: more 's' responses were given in the /u/ context. The sizes of the shifts in mean phoneme boundary as a function of vowel context were also similar for all groups (a shift of approximately one stimulus on the continuum). Although values for slopes were not given, visual inspection of the identification functions suggests that they were equally steep for both groups of 7-year-olds and for adults, but may have been somewhat less steep for the 5-year-olds: In other words, the 5-year-olds may have been less consistent in their responses.

These apparent differences in slopes between the 5-year-olds and the 7-year-olds and adults are of particular interest. The simplest explanation for these differences is that they reflect the variability usually associated with a partially learned (or not fully established) response. However, the degree of control exerted by formant transitions on responding was not assessed because the same formant transitions (those appropriate to /ʃ/) were used for all stimuli. If the responses of the 5-year-olds were more strongly controlled by formant transitions than the responses of 7-year-olds and adults, the fixed transition stimuli of this study would probably have made these stimuli particularly ambiguous for 5-year-olds. Thus, the age-related slope differences may have reflected a substantive difference in perceptual processing rather than simple variability of response. It would therefore be of interest to investigate the transition effect in children's speech perception to see if children do indeed learn to use different kinds of coarticulatory informtion at different ages. The chance of discovering such differences would presumably be increased by extending the age range of the study to include children younger than 5 years.

One other study deserves mention. Morrongiello, Robson, Best, and Clifton (1984) investigated the ability of 5-year-olds to use multiple cues in recognizing a stop between an initial consonant and following vowel. Briefly, a natural /s/ segment was placed before each of two synthetic patterns identical in all aspects except for the extent of an initial F_1 transition. One pattern, heard by adults equivocally as either 'ay' or 'day,' had a short F_1 transition of 181 Hz; the other pattern, always heard by adults as day, had a long F_1 transition of 381 Hz. An acoustic continuum from 'say' to 'stay' was then constructed by inserting a silent interval, varying in duration from 0 ms to 104 ms, between the /s/ and the synthetic patterns. For adults, the crossover point between the identification of 'say' and 'stay' varied with the extent of the F_1 transition: significantly more 'stay' responses were given to tokens with the longer transition (Best, Morrongiello, & Robson, 1981). Morrongiello et al. (1984) found that, although children also gave more 'stay' responses to tokens with the longer F_1 transition, the shift in phoneme boundary was smaller for children than for adults. This decreased boundary shift was entirely due to the children giving more 'stay' responses to tokens with the shorter vocalic transition. Morrongiello et al. proposed that the children may have weighted the transitional information relatively more heavily, and the temporal information relatively less heavily, than adults do. That is, perhaps the children were more sensitive to transitional cues than adults, a possibility encouraged by the finding that any transition—even a brief one—was sufficient to elicit some 'stay' responses from the children. This suggestion provides a third possible outcome for the present experiment: that children would prove even more sensitive to some kinds of coarticulatory effects than adults.

Present Study

The main purpose of the present experiment was to trace the development of perceptual sensitivity to the acoustic consequences of coarticulation by comparing children's shifts in phoneme boundaries as a function of phonetic context with those of adults. If the same sized shifts were seen, this could imply that perceptual use of contextually variable properties does not develop gradually from an initial sensitivity to putatively invariant properties of the speech signal, but is, in some sense, intrinsic to the process of speech perception. If children showed smaller shifts or no shifts, this would imply that the ability to use contextually variable information is an acquired skill (Stevens & Blumstein, 1978)—perhaps of particular utility in following the "reduced" patterns of casual adult speech.

Finally, if children demonstrated relatively greater sensitivity to coarticulatory effects than adults, this could imply that they are more dependent on information specifying syllable coherence: Perhaps young children are not as adept as adults at recovering the individual phonemes from the syllable, but instead tend to perceive syllables as relatively undifferentiated wholes. If children do not recover individual segments to the same extent as adults, it might be predicted that responses would be less consistent for children than for adults, indicating that relatively less weight is being assigned to the spectrum of the fricative noise *per se*, and more weight to the syllabic structure specified, for example, by formant transitions.

Method

Subjects

One group of twelve adults and four groups of eight children of the ages 3, 4, 5, and 7 years participated in this experiment. All adults were 20- or 21-years-old, and all children were within −1 and +5 months of their designated age. No age group had greater than a 5 to 3 ratio between sexes. All subjects were native speakers of American English, speaking a dialect typical of the Middle Atlantic States, and all the children had age-appropriate articulation skills, as judged independently by two speech pathologists from recordings of spontaneous speech.

Each subject also passed audiometric and tympanometric screenings. Pure tones of 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz were presented at 20 dB HL for the audiometric evaluation. Tympanometric results for all subjects displayed normal pressure peaks between +100 daPa and −150 daPa.

Finally, all subjects were required to identify accurately the endpoints of the /ʃ/-/s/ continuum, followed by /i/ and /u/ with appropriate vocalic transitions, at greater than 90% accuracy during an initial practice trial. In addition, any subject failing to show greater than 80% correct identification for these good exemplars during the acual testing was eliminated. This requirement served as a partial check on the attentiveness of the subjects. There was no way to measure subjects' attentiveness (either for children or for adults) to the more ambiguous stimuli, and so it was necessary to assume that attention remained constant across stimuli.

One 3-year-old was eliminated because he failed to identify the endpoint stimuli with appropriate transitions at greater than 90% accuracy during the practice session. One 4-year-old was eliminated because she did not respond with greater than 80% accuracy to these tokens during the actual testing. Five children were eliminated because their speech was judged to be less well-developed than expected for their ages (three 3-year-olds, and one child at each of the ages of 4 and 5).

Stimuli

For the present study, we needed stimuli that would permit us to observe the effects on frictive perception of fricative spectrum, vowel context, and vocalic transition. We therefore prepared hybrid syllables composed of synthetic fricative noises and natural vowels, analogous to those of Mann and Repp (1980) and of Whalen (1981). The fricatives varied along a nine-step continuum from /ʃ/ to /s/, and the vocalic portions were excerpted from naturally spoken /ʃi/, /si/, /ʃu/, and /su/ syllables, so that formant transitions were appropriate for either /ʃ/ or /s/.

The stimuli were constructed on the basis of pilot tests, and consisted of tokens from an /ʃ/ to /s/ continuum created on the serial software synthesizer at Haskins Laboratories. Each token contained 210 ms of fricative noise with a single pole value specified. There were nine tokens on this continuum, with pole values varying from 2200 Hz to 3800 Hz in 200-Hz steps. Each token also had a zero at .75 x the pole value, creating a dip in the spectrum. Thus we synthesized, in effect, two-pole fricatives, more similar to natural speech than might have sufficed for the perception of /ʃ/ and /s/. For example, Heinz and Stevens (1961) demonstrated that a single-pole stimulus is sufficient for listeners to perceive fricatives; however, we wanted stimuli as close to natural speech as possible because children were serving as subjects. Thus, any concern that children may have responded differently from adults because they had difficulty hearing synthetic stimuli as speech was mitigated. The amplitude of these tokens rose gradually by 20 dB over the first 170 ms, remained constant for the next 20 ms, then fell by 5 dB over the final 20 ms. Fricative noises were synthesized with a sampling rate of 10 kHz, and were low-pass filtered at 4.9 kHz.

The vocalic portions of the stimuli were the vowels /i/ and /u/ produced by a male speaker in the syllables /ʃi/, /si/, /ʃu/, and /su/. The unrounded vowel /i/ was chosen instead of /a/ because second formant trajectories are more distinct between /i/ and /u/ than between /a/ and /u/ (Soli, 1981). The speaker originally produced five samples of each syllable, but because Mann and Repp (1980) had found little token-to-token variability in results, it seemed sufficient to use a single sample of each. Tokens were selected so that they matched one another as closely as possible in duration and intonation contour. These syllables were digitized on a VAX computer using a 10 kHz sampling rate, and low-pass filtering with an upper cut-off of 4.9 kHz. The vowels, including the vocalic transitions, were isolated from the fricative nose with a waveform editing program. The vocalic segments were between 350 ms and 370 ms in duration. Spectrographic displays and LPC analysis were used to obtain frequency values for formant transitions. These values are given in Table 1 along with the transition values from Whalen (1981) for comparison. Thus, there were nine fricative noises combined with four vocalic segments, resulting in a total of 36 unique stimuli. Each stimulus was presented 10 times, making a total of 360 test tokens.

Four test tapes were prepared: two using only the /i/ context, and two using only the /u/ context. Ten practice items began each tape (the two endpoints with appropriate vocalic segments presented five times each in random order). The 90 test items per tape were then presented in five randomized blocks of 18. Before each group of 10 stimuli, a female voice announced the number of the next group of stimuli to be heard (numbers 1 through 9). The interstimulus interval was 2 s. Tapes were made on Scotch Audio Tape 208 at a tape speed of 19 cm per second (7.5 ips) and at a -5 VU recording level.

Table 1

Formant transition characteristics of vocalic portions used in the present study and by Whalen (1981)

	F2			F3		
	onset, (Hz)	dur, (ms)	steady-state (Hz)	onset, (Hz)	dur, (ms)	steady-state (Hz)
Present study						
/(ʃ)i/	2070	0	2070	2530	60	2580
/(s)i/	1870	80	2050	2430	140	2600
/(ʃ)u/	1940	300	1000	2350	60	2150
/(s)u/	1620	300	1000	2340	170	2130
Whalen, 1981						
/(ʃ)i/	2100	130	2300	2750	130	2900
/(s)i/	1700	130	2300	2500	130	3000
/(ʃ)u/	1800	300	850	2100	0	2100
/(s)u/	1600	300	850	2750	130	2100

Materials and Equipment

Stimuli were presented over a Uher Model 4200 tape recorder using Sennheiser earphones. A group listening station permitted the adults to listen in groups of four, and one of the experimenters to listen simultaneously when a child was being tested. A Beltone Model 9D portable audiometer and a Grason-Stadler Model 27 portable tympanometer were used to screen pure-tone threshold and middle ear function.

For the children, four hand-drawn pictures were prepared corresponding to each of the four possible responses. One picture was of a shoe; one was of a girl named "Sue"; one was of a boy pointing and saying "see"; and the last was of a girl referred to by the pronoun "she." Several board games were also prepared. These games consisted of brightly colored squares and circles (nine on each board) with the numbers 1 through 9 written on the spaces. A small plastic animal served as a marker, moving to the next space each time the female voice was heard announcing the number of the next group of 10 stimuli. Toy stickers, selected before the start of each game, were given as prizes at the end of the game.

Procedure

Each subject took the hearing screening before testing. Next, the children were shown the two pictures to be used in the first perceptual test (either 'see' and 'she,' or 'Sue' and 'shoe').

Practice was provided using live voice to familiarize the child with the labeling procedure. This procedure consisted of having the child point to the picture illustrating the stimulus perceived and say the label associated with that picture. If there was a discrepancy between what the child seemed to be saying and the picture to which he/she was pointing, clarification was requested by the experimenter. After the child correctly responded to 10 practice items with live voice, the 10 taped practice items were presented.

If the child correctly responded to 9 out of 10 of these correctly, the board game was introduced, and the actual testing started. Stimuli were presented over earphones at approximately 75 dB SPL. Two experimenters were needed during the testing of the 3- and 4-year-olds: one to work with the child, listening and watching for discrepancies between his/her pointing and verbal responses as well as just maintaining attention with constant eye contact and occasional verbal praise, and one to stop the tape between stimulus presentations when necessary and to record the child's responses. The first experimenter working with the child was not able to hear the stimulus presentations. The second experimenter recording the children's responses wore the earphones around her neck so that she could hear when a stimulus had been presented, but could not hear it well enough to form her own impression of whether it was /s/ or /ʃ/. Thus, there was no danger of the response recorded being influenced by her own perception. Only one experimenter was needed for testing 5- and 7-year-olds.

Testing was divided into three sessions for the children; two sessions of 20-30 minutes duration, and one of 10-15 minutes duration. Activities during the sessions were arranged as follows:

Session 1: Hearing screening and first half of first perceptual test (1 test tape).

Session 2: Second half of first perceptual test, and first half of second perceptual test (2 test tapes).

Session 3: Second half of second perceptual test (1 test tape). All testing was done in one day with breaks between sessions. Testing took place either in the home or in a daycare facility. The order of presentation of the perceptual tapes was counterbalanced across subjects within each age group. Adults simply responded by writing whether they heard 's' or 'sh.' In addition, adults listened in groups of four, and all tapes were presented during one session with the /i/ and /u/ tapes alternated.

Results

Probit analyses were done on individual and group response proportions (probability estimates) for each context. (The general term "context" is used here to refer to the four vocalic portions used: /(ʃ)i/, /(s)i/, /(ʃ)u/, and /(s)u/.) Probit analysis fits a cumulative normal curve to probability estimates as a function of stimulus level by the method of least squares (Finney, 1971), estimating the mean (phoneme boundary) and standard deviation for each distribution. The slope is obtained by taking the reciprocal of the standard deviation, and therefore serves as an index of consistency. Figures 2 through 6 display the identification functions obtained for each age group from probit analyses of mean probability estimates. Table 2 provides mean phoneme boundaries for each age group in each context. The smaller (or lower) the phoneme boundary value, the more 's' responses were given. Results should be compared (1) between different vocalic transitions for the same vowel, /(ʃ)i/ vs. /(s)i/ and /(ʃ)u/ vs. /(s)u/, and (2) between different vowel contexts for the same vocalic transition, (ʃ)i/ vs. /(ʃ)u/ and /(s)i/ vs. /(s)u/. From the adult studies reviewed above, we would expect lower phoneme boundaries for tokens with /s/ transitions in the first case, and lower boundaries for /u/ contexts in the second.

Inspection of the mean phoneme boundaries given in Table 2 reveals that the predicted transition effect is apparent within both vowel contexts (lower phoneme boundaries for tokens with /s/ transitions), but the size of this effect appears greater for /u/ than for /i/. The expected vowel effect occurred only for tokens with /s/ transitions (a lower phoneme boundary for the /(s)u/ context than for the /(s)i/ context). For token with /ʃ/ transitions, the obtained vowel effect was opposite to what would be predicted from previous studies: a lower phoneme boundary was obtained in the /(ʃ)i/ context than on the /(ʃ)u/ context. In fact, all age groups demonstrated the highest phoneme boundary in the /(ʃ)u/ context, indicating more 'sh' responses in this context than in any other.

Table 2

Mean phoneme boundaries for each age group (computed from individual phoneme boundaries), with group standard deviations in parentheses

	/(ʃ)i/	/(s)i/	/(ʃ)u/	/(s)u/
Adults	5.46	5.14	5.94	3.84
	(0.98)	(1.03)	(1.23)	(1.04)
7-year-olds	5.21	4.63	5.81	3.74
	(0.61)	(0.66)	(0.39)	(0.90)
5-year-olds	5.53	4.85	6.26	4.32
	(0.57)	(0.48)	(0.99)	(1.53)
4-year-olds	5.95	5.22	7.89	4.51
	(0.55)	[illegible]	(1.70)	(0.58)
3-year-olds	5.20	4.60	7.87	3.91
	(0.83)	(0.55)	(2.16)	(2.38)

A 3-way ANOVA (Age × Vowel × Transition) performed on individual phoneme boundaries revealed a significant main effect of transition, $F(1,39) = 171.64$, $p < .001$, and significant interaction effects for Transition × Vowel, $F(1,39) = 75.30$, $p < .001$, for Transition × Age, $F(4,39) = 3.15$, $p = .02$, and for Vowel × Age, $F(4,39) = 2.61$, $p = .05$. Planned comparisons indicated a significant, overall effect of age on the placement of phoneme boundaries for the 3-, 4-

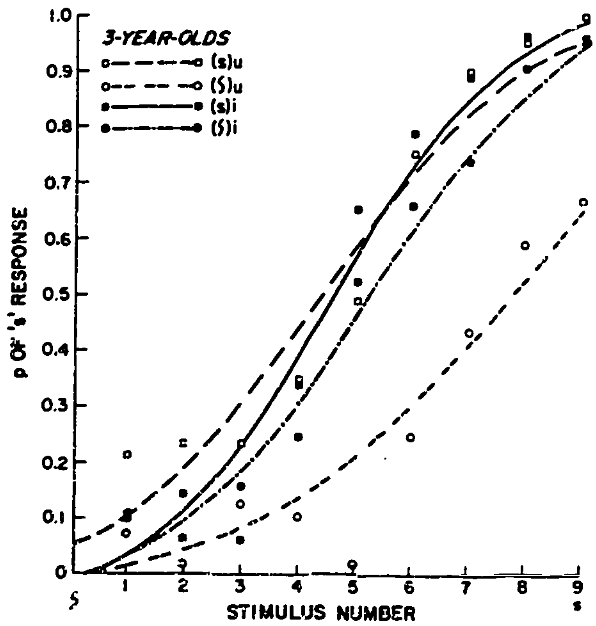


Figure 2. Mean identification functions for 3-year-olds: Symbols represent mean probability estimates at each stimulus step for each context (i.e., the vowels /i/ or /u/, with formant transitions appropriate for /ʃ/ or /s/). Lines are the cumulative normal curves fitted to these mean probability estimates with probit analysis.

and 5-year-olds vs. 7-year-olds and adults, $F(1,39) = 4.67$, $p = .04$.[1] The comparison between 3-, 4-, and 5-year-olds vs. 7-year-olds and adults also showed significant results for the vowel effect, $F(1,39) = 6.75$, $p = .01$, and for the transition effect, $F(1,39) = 6.04$, $p = .02$. Thus, there was a difference between the younger children as compared to the 7-year-olds and adults in the relative placement of phoneme boundaries. Specifically, phoneme boundaries were affected differently by both vowel and transition for the younger children than for the 7-year-olds and adults. Finally, the planned comparison between 3- and 4-year-olds vs. 5-year-olds, for the transition effect showed a significant difference, $F(1,39) = 5,82$, $p = .02$. The size of the transition effect decreased with increasing age.

---

[1] The following orthogonal, planned comparisons were performed on data for each factor (i.e., Age, Vowel, and Transition):

Group 1	vs.	Group 2
7-year-olds	vs.	Adults
3-year-olds	vs.	4-year-olds
3- and 4-year olds	vs.	5-year-olds
3-,4-, and 5-year olds	vs.	7-year-olds and Adults

These comparisons were intended to locate, if possible, breaks in the developmental sequence from child to adult. Only those comparisons that resulted in statistical significance are reported here.

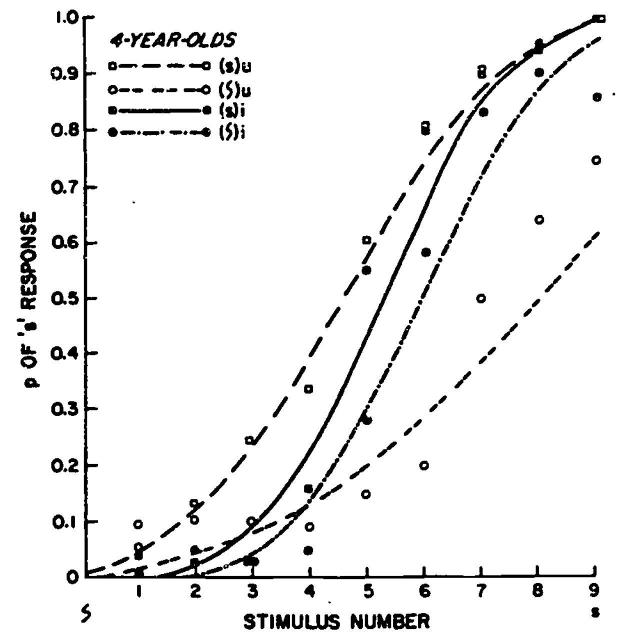


Figure 3. Mean identification functions for 4-year-olds. See Figure 2 caption for additional details.

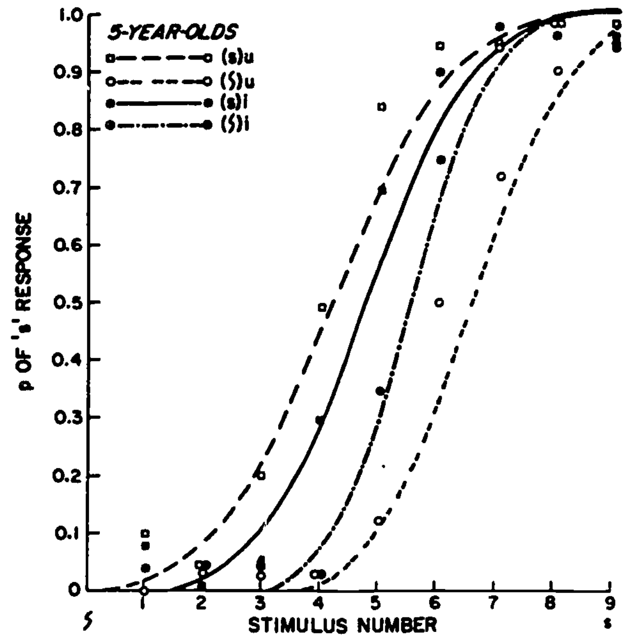


Figure 4. Mean identification functions for 5-year-olds. See Figure 2 caption for additional details.

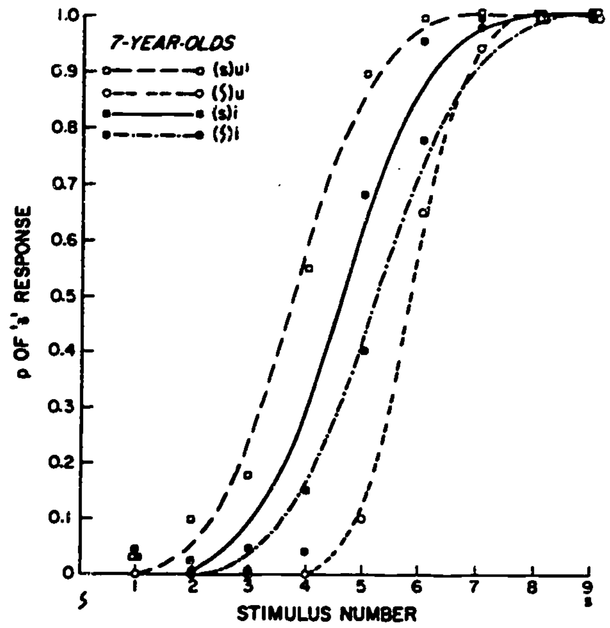


Figure 5. Mean identification functions for 7-year-olds. See Figure 2 caption for additional details.

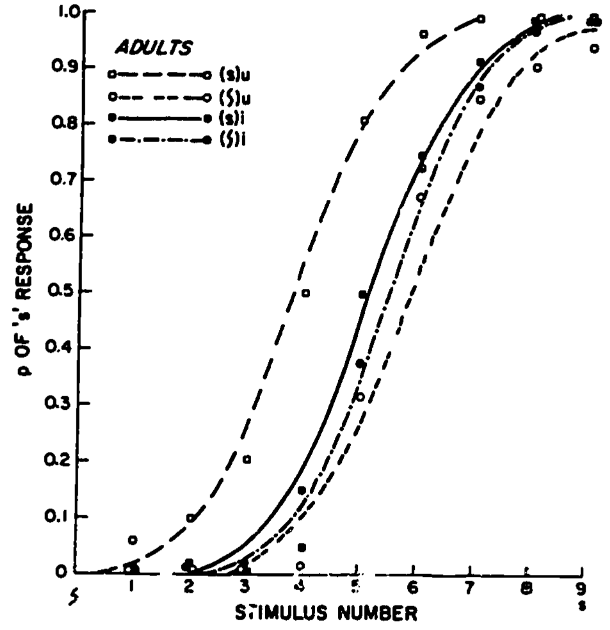


Figure 6. Mean identification functions for adults. See Figure 2 caption for additional details.

Inspection of Table 3, listing mean slopes for each age group, reveals that slope values were generally greater for older subjects. Adults and 7-year-olds demonstrated mean slopes above 1.00; slopes for 5-year-olds were closer to 1.00; and, with one exception, slopes for 4- and 3-year-olds were below 1.00. Looking across contexts within age groups it can be seen that slope values were fairly consistent. Only 4-year-olds seem to display different slope values as a function of context, with the values for /i/ higher than the values for /u/. However, this difference is not statistically significant. A 3-way ANOVA (Age x Vowel x Transition) performed on individual slope values revealed a significant main effect of age, $F(4,39) = 13.79$, $p < .001$. Planned comparisons indicated significant differences in slope values for 3-, 4-, and 5-year-olds vs. 7-year-olds and adults, $F(1,39) = 38.30$, $p < .001$, and for 3- and 4-year-olds vs. 5-year-olds, $F(1,39) = 14.31$, $p < .001$. Therefore, consistency of responses increased with increasing age for these subjects.

Table 3

Mean slope values[a] for each age group (computed from individual slope values)

	/(ʃ)i/	/(s)i/	/(ʃ)u/	/(s)u/
Adults	1.37	1.33	1.28	1.24
7-year-olds	1.30	1.48	1.73	1.58
5-year-olds	1.40	1.09	1.01	1.08
4-year-olds	0.86	1.00	0.42	0.60
3-year-olds	0.47	0.50	0.37	0.52

[a] Slope values are the reciprocals of the standard deviations of the fitted normal curves. They estimate the changes in probability of an 's' response per unit change in the fricative noise.

Discussion

The "Reversed" Vowel Effect

The first finding to be discussed is an apparent discrepancy between results of this study and earlier studies. The subjects in the present experiment gave more 's' responses to the unrounded vowel than to the rounded vowel for tokens with /ʃ/ transitions. Not only does this finding differ from those of earlier studies, it is contrary to the suggestion made by Kunisaki and Fujisaki (1977) that more 's' responses would be expected before rounded than before unrounded vowels, regardless of transition, because fricative poles are lower before rounded vowels.

This "reversed" vowel effect found for tokens with /ʃ/ transitions is probably related to the inequality in the relative size of transition effects found between vowel contexts. The predicted transition effect for these stimuli was that more 's' responses would be given to tokens with /s/

vocalic transitions. This effect was obtained, but to a greater extent for /u/ than for /i/. Both of these findings (i.e., the "reversed" vowel effect for tokens with /ʃ/ transitions and unequal transition effects for /i/ and /u/) may be explained by comparing the acoustic characteristics of the formant transitions used in this and previous studies.

The extents of F_2 transitions are more similar for /a/ and /u/ than for /i/ and /u/ (Soli, 1981). This follows from the relatively shorter distance that the tongue must travel from the point of fricative constriction to the front vowel /i/ than to the back vowels /a/ and /u/. The pattern of results obtained by Mann and Repp (1980) for tokens with /ʃ/ transitions may therefore have largely been an effect of the following vowel. That is, since /a/ and /u/ are similar in the extent of their transitions from the fricative constriction, differences between vocalic portions with transitions appropriate for the same fricative would be entirely due to vowel quality.

The /i/ portions used by Whalen (1981) exhibited greater transitional differences than the /i/ portions used in the present study: Differences in F_2 onset values were twice as large between /(ʃ)i/ and /(s)i/, and the extents of /i/ transitions were greater. Moreover, the acoustic differences between the /(ʃ)u/ and /(s)u/ portions of Whalen's study were not as great as the corresponding differences between the /u/ portions of this study. In particular, the /(ʃ)/ portion used by Whalen does not provide as much transitional information as the /(ʃ)u/ portion used here because F_3 shows no change in Whalen's /(ʃ)u/. In short, the extent of the transitions for /i/ and /u/ tokens seems to have been more similar in Whalen's syllables than in those of the present study. Therefore, differences in placement of /(ʃ)i/ and /(ʃ)u/ boundaries in Whalen's study would have been a function more of vowel context than of differences in the extents of the transitions in the two contexts.

We thus hypothesize that the perceptual weight given to transitions is proportional to their extents. In fact, precisely this result has been reported in a study of the relative perceptual weights assigned to bursts and transitions in naturally spoken stop-vowel syllables (Dorman, Studdert-Kennedy, & Raphael, 1977). In the present study, the transitions of the /(ʃ)i/ portion displayed very little formant movement and may therefore have roughly represented a null condition; /(s)i/ had weak transitions favoring 's'; /(s)u/ had somewhat more perceptually salient transitional information favoring 's'; but /(ʃ)u/ provided very strong transitional information favoring 'sh.' Superimposed on this pattern seems to be an independent vowel effect that tends to move the two /u/ functions in the direction favoring 's' responses—at least for adults, who demonstrate sensitivity to vowel quality.

Clearly, the relation between the size of transitional differences found in acoustic measurements and the size of boundary shifts observed in perception warrants more systematic investigation. In most previous studies of phonetic context effects, vocalic transitions were manipulated as nominal categories: that is, they were treated as either /s/ or /ʃ/ transitions. This may be too coarse a measure. Vocalic transitions might better be regarded as continuous variables, relatively more /s/-like or more /ʃ/-like. In any event, the finding that subjects in the present experiment gave more 's' responses to the /(ʃ)i/ context than to the /(ʃ)u/ context argues against explanations of the vowel effect based primarily on auditory contrast. If the predicted vowel effect were a consequence of the fricative noise sounding higher when the vowel spectrum is lower, as in /u/ compared to /i/, this effect should have been found regardless of transition.

Comparisons Between Children and Adults

In considering whether children demonstrated sensitivity to coarticulatory effects to a greater or lesser extent than adults, three factors must be discussed: the vowel effect, the transition effect, and the extent to which fricative identification was based on the fricative spectrum itself.

Vowels and transitions. The expected vowel effect for these stimuli was that more 's' responses would be given to /u/ tokens than to /i/ tokens due to the lowered fricative poles associated with rounded vowels. Children in this study demonstrated a weaker effect of this sort than adults. Visual inspection of the mean identification functions for each age group clearly shows a greater separation between the (s)i and (s)u functions for adults than for children, and for 7-year-olds than for younger age groups.

The predicted transition effect (more 's' responses to tokens with /s/ vocalic transitions) was obtained for all age groups for both vowels, but the size of this effect was larger for /u/ than for /i/. There was also an effect of age on the size of the transition effect: Younger subjects demonstrated greater transition effects than older subjects. This age effect appears to be largely (though not entirely) due to younger subjects giving substantially more 'sh' responses to the /(ʃ)u/ context than older subjects.

For these /(ʃ)u/ tokens two opposing factors presumably influenced the locations of phoneme boundaries: perceptually salient transitional information biased responses toward 'sh,' while the rounded vowel biased responses toward 's.' The fact that younger subjects gave substantially more 'sh' responses to /(ʃ)u/ tokens than older subjects seems to be another indication that younger subjects were more influenced by transitional information than older subjects, and less influenced by vowel quality. While /(s)u/ tokens provided perceptually salient transitional information concerning fricative identity, vowel quality biased responses in the same direction (toward 's' responses). Therefore, we cannot assess the relative contributions of vowels and transitions to responses in this context.

Based on these results, it appears that vowel effects and transition effects are additive.[2] The negative boundary shift obtained for the /(ʃ)i-(ʃ)u/ difference evidently reflects the overwhelming cancellation of the vowel effect by the /(ʃ)u-(s)u/ transition effect: vowel and transition effects are evidently orthogonal. The fact that the 3- and 4-year-olds demonstrate a large negative vowel effect for /ʃ/ tokens is thus a result of their demonstrating a very large transition effect for /u/.

Fricative spectrum. Another purpose of the present experiment was to investigate the extent to which fricative identification was associated with the fricative spectrum itself. To make this determination, we must look at the slopes of the identification functions. The finding that 3-, 4-, and 5-year-olds exhibited shallower slopes than 7-year-olds and adults shows that they were less consistent in responding. This decrement in consistency exhibited by young children might simply reflect an inability to pay attention to the task at hand. However, the fact that all subjects

---

[2] On first consideration, the statistically significant Vowel × Transition interaction may belie this statement; however, the statistical analysis treats values on these factors as nominal categories. In fact, the /u/ transitions provided more information concerning fricative identity than the /i/ transitions, and it is precisely this inequality of information that we believe helps support the claim of orthogonality of effects.

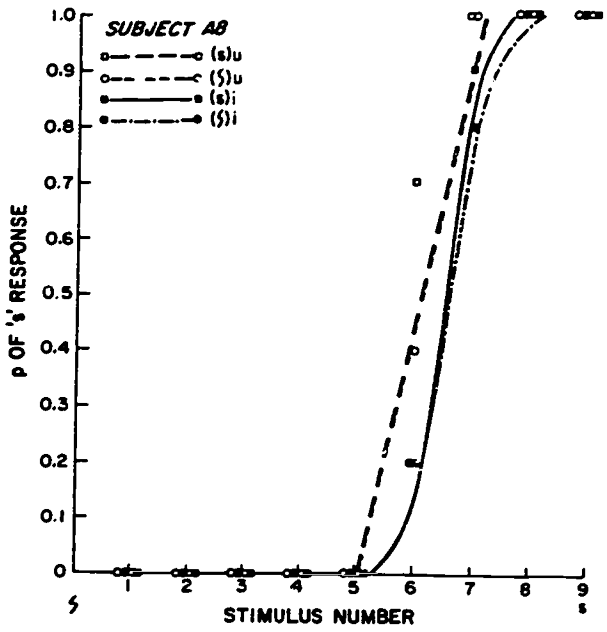


Figure 7. Identification functions for Subject A8. See Figure 2 caption for additional details.

responded with greater than 80% accuracy to the endpoints presented with the appropriate vocalic transition argues against this interpretation. An alternative account is that the lack of consistency reflected the children's relatively lower sensitivity to the fricative spectrum. One adult subject who appeared to attend selectively to the frictive noise, and to make decisons about category based primarily on its spectrum, was Subject A8. whose identification functions are shown in Figure 7. These functions show very consistent responses and no effect of vocalic transition. Yet they do exhibit a vowel effect.

This last result is important because it suggests that attention to the fricative spectrum is associated with a vowel effect. Vowel context effects are found because the fricative spectrum differs as a function of the quality of the following vowel. If a listener did not attend to the fricative spectrum to some extent, then he/she would not be able to determine whether this spectrum differed as a function of the following vowel.

This relation between attending to the fricative spectrum and showing vowel effects might explain why young children did not demonstrate vowel effects to the same extent as adults in this study. Morrongiello et al. (1984) reported that the children in their study appeared to weight the transitional information relatively more heavily than the silent interval, as compared to adults. That is, children seemed to be relatively more sensitive to the transitional pattern than to the silent interval. The results obtained in the present experiment show a similar pattern: children were apparently more sensitive to the transitional pattern than to the fricative spectrum, as compared to adults.

The shapes of the identification functions also indicate the increased reliance on transitional information of the children in this study. If categorical decisions were more heavily weighted by vocalic transitions than by fricative spectrum, we would expect responses to be biased in favor of the fricative associated with that transition, and many of the stimuli to be ambiguous. Therefore, phoneme boundaries would be either extremely low or extremely high, and slope values would be low.

An example of a 3-year-old subject whose responses to /u/ tokens seem to have been controlled primarily by transitional information is Subject 32, whose identification functions are shown in Figure 8. Phoneme boundaries are extreme for /(ʃ)u/ and /(s)u/, and slope values are low for these contexts. Similar patterns (extreme phoneme boundaries and low slope values) can be found in results obtained for several of the other 3- and 4-year-olds, but only for /u/ contexts. No subject displayed a pattern of results indicating strong influence of transitions in either of the /i/ contexts, probably because /i/ transitions provided less salient information about fricative identity than did /u/ transitions. However, many of the 3-year-olds displayed very low slope values, simply indicating inconsistent responses for the /i/ contexts. Apparently, even when transitional information was greatly reduced, 3-year-olds were not able to increase their use of the cues provided by the fricative spectrum for fricative identification.

Four-year-olds seem to have been able to switch response patterns to some extent depending on how much information was provided by the transitions. For the /u/ contexts in which the transitions provided perceptually salient information concerning fricative identity, responses seem to have been based largely on vocalic transitions, as indicated by the low slope values and extreme phoneme boundaries. For the /i/ contexts in which transitions provided little information, four-year-olds seem to have been able to use the fricative spectrum to make decisions about fricative identity to at least some extent, as indicated by higher slope values.

A good example of this response pattern is Subject 44, whose identification functions are shown in Figure 9. Responses for the /(ʃ)u/ context are strongly influenced by the vocalic transitions, as indicated by the shallow slope and extremely high phoneme boundary, but responses for the /(ʃ)i/ context are clearly based on the fricative spectrum, as indicated by the very steep slope.

Another developmental trend found in these data is that 7-year-olds performed similarly to adults in most respects. Their responses were as consistent as those of adults, indicating strong reliance on the fricative spectrum for fricative identification, and the magnitude of the transition effect for individual 7-year-olds was equivalent to that obtained for individual adults. The only difference between adults and 7-year-olds was in the magnitude of the vowel effect: 7-year-olds demonstrated slightly reduced vowel effects. In general, though, 7-year-olds appear to have developed speech perception strategies similar to those of adults.

Conclusions

Contextual effects on young children's speech perception were initially studied with the expectation that they would be less than or equal to the effects observed in older children and adults. For vowel quality the effects were indeed reduced: Young children seemed less able than older children and adults to attend to those portions of a syllable primarily associated with the individual segments (that is, with the steady state friction and vocalic formants). On the other

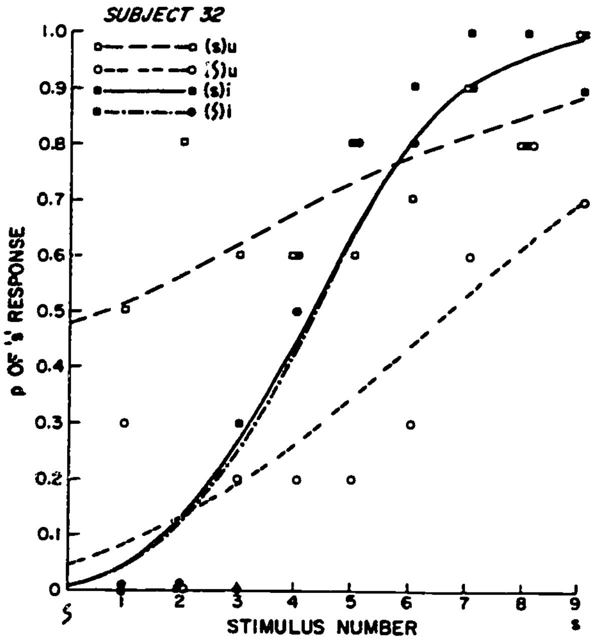


Figure 8. Identification functions for Subject 32. See Figure 2 caption for additional details.

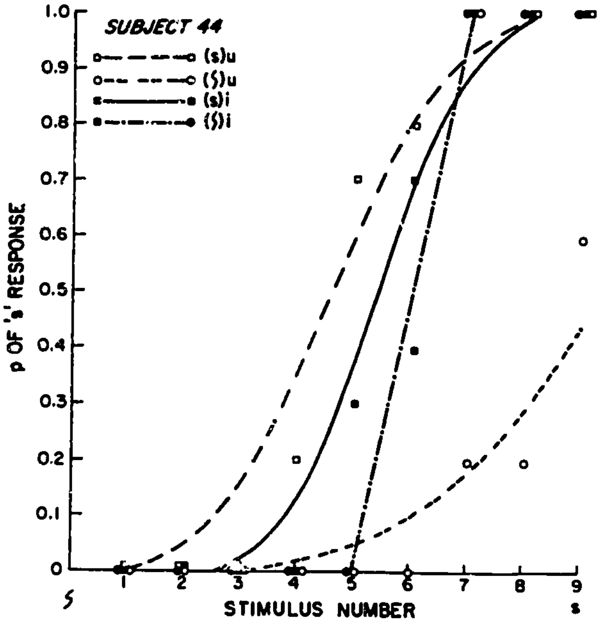


Figure 9. Identification functions for Subject 44. See Figure 2 caption for additional details.

hand, for the vocalic transitions, the reverse was found: younger children were more sensitive than older children and adults to those portions of a syllable that ensure its perceptual coherence.

These results suggest that the younger children were still listening for whole words or syllables, much as (we may infer from their productions) the 1- to 2-year-olds described by Menyuk and Menn (1979). Of course, it is unlikely that each of the many hundreds of entries in a 3- to 4-year-old's lexicon is still a whole word, quite lacking in segmental structure. But it is not implausible to suppose that 3- to 4-year-olds are still discovering how segments are packaged in the syllable and are not yet as adept at recovering them as they will soon become. In fact, this suggestion receives support from many studies indicating that preliterate children do not have access to the phonemic structure of speech (e.g., Fox & Routh, 1975). However, this interpretation runs contrary to the commonly held view (e.g., Kent, 1983) that the perceptual (and *a fortiori* articulatory) organization of speech becomes less rather than more segmental as the child develops.

Such an interpretation also runs counter to the claim of Stevens and Blumstein (1978) that sensitivity to coarticulation in adult speech perception is a secondary effect, learned by association with a primary invariant. Our results suggest rather that perceptual sensitivity to certain forms of coarticulation is present from a very early age and therefore may be intrinsic to the process of speech perception. The child does not use segments to discover coarticulation, but rather coarticulation to discover segments. The adults, though more skilled at segmental recovery than the child, may still do much the same.

References

Best, C. T., Morrongiello, B. A., & Robson, R. C. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics, 28*, 191-211.

Darwin, C. J., & Bethell-Fox, C. E. (1977). Pitch continuity and speech source attribution. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 665-672.

Dorman, M. F., Raphael, L. J., & Liberman, A. M. (1979). Some experiments on the sound of silence in phonetic perception. *Journal of the Acoustical Society of America, 65*, 1518-1532.

Dorman, M. F., Studdert-Kennedy, M., & Raphael, L. J. (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics, 22*, 109-122.

Finney, D. J. (1971). *Probit analysis*. Cambridge: Cambridge University Press.

Fox, B., & Routh, D. K. (1975). Analyzing spoken language into words, syllables, and phonemes: A developmental study. *Journal of Psycholinguistic Research, 4*, 331-342.

Heinz, J. M., & Stevens, K. N. (1961). On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America, 33*, 589-596.

Joos, M. (1948). Acoustic phonetics. *Language, 48 (Suppl. 2)*, 1-136.

Kent, R. D. (1983). The segmental organization of speech. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 57-89). New York: Springer-Verlag.

Kunisaki, O., & Fujisaki, H. (1977). On the influence of context upon perception of voiceless fricative consonants. *Annual Bulletin, 11*, 85-91. (Tokyo: Research Institute for Logopedics and Phoniatrics.)

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431-461.

Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the /ʃ/-/s/ distinction. *Perception & Psychophysics, 28*, 213-228.

Mann, V. A., Sharlin, H. M., & Dorman, M. (1985). Children's perception of sibilants: The relation between articulation and perceptual development. *Journal of Experimental Child Psychology, 39*, 252-264.

Menyuk, P., & Menn, L. (1979). Early strategies for the perception and production of words and sounds. In P. Fletcher & M. Garman (Eds.), *Language acquisition* (pp. 49-70). Cambridge: Cambridge University Press.

Morrongiello, B. A., Robson, R. C., Best, C. T., & Clifton, R. K. (1984). Trading relations in the perception of speech by 5-year-old children. *Journal of Experimental Child Psychology, 37*, 231-250.

Pisoni, D. B. (1985). Speech perception: Some new directions in research and theory. *Journal of the Acoustical Society of America, 78*, 381-388.

Soli, S. D. (1981). Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *Journal of the Acoustical Society of America, 70*, 976-984.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America, 64*, 1358-1368.

Whalen, D. H. (1981). Effects of vocalic transitions and vowel quality on the English /s/-/ʃ/ boundary. *Journal of the Acoustical Society of America, 69* 275-282.

HEMISPHERIC ASYMMETRIES IN PHONOLOGICAL PROCESSING*

G. Lukatela,** Claudia Carello,† M. Savić,** and M. T. Turvey‡

Abstract. *In Serbo-Croat lexical decision to phonologically bivalent letter strings is slowed relative to their phonologically unique counterparts. This feature was exploited in order to assess the linguistic capacity of the two hemispheres. Lexical decision to laterally presented words and pseudowords revealed a right visual field advantage for both males and females in words and pseudowords. But the demands of phonological processing were not met in the same way by the two hemispheres, the two hemispheres did not respond in the same way in the two sexes, and the pattern of these differences was opposite for words and pseudowords.*

Introduction

Although it is generally accepted that the left hemisphere is specialized for language, there is some question as to what language capacity, if any, exists in the right hemisphere. For example, among split-brain patients (i.e., those in whom the corpus callosum has been severed so that exchange of information between the two hemispheres is precluded), right-hemisphere language competence appears to vary considerably. With a majority of such patients, left visual field (LVF) presentation (which projects only to the right hemisphere) of verbal stimuli reveals no right-hemisphere language of any kind. In contrast, a few of these commissurotomy patients appear to have a rich semantic system, and a few–who demonstrate syntax and phonology in addition to semantics–appear linguistically sophisticated (Gazzaniga, 1983; Patterson & Besner, 1984). Whether such differences indicate different degrees of lateralization or whether they correspond to varying degrees of early presurgical left-hemisphere damage for which the right hemisphere took over, or whether they suggest that only the intermediate level of competence represents a "normal" right hemisphere, is a subject of considerable debate (Gazzaniga, 1983; Levy, 1983; Zaidel, 1983).

---

* *Neuropsychologia* 1986, *24*, 341-350.

** University of Belgrade

† State University of New York at Binghamton

‡ Also University of Connecticut

Acknowledgment. The research was supported in part by National Institute of Health Grants HD-08495 to the University of Belgrade, Yugoslavia, and HD-01994 to Haskins Laboratories.

With respect to patients with left-hemisphere damage, Coltheart argues that the acquired reading disorder known as deep dyslexia provides several parallels to the intermediate level reading performance of two split-brain patients, N.G. and L.B. (Coltheart, 1980, 1983). These include failures to read a word that is, nonetheless, understood (Saffran & Marin, 1977); an inability to generate a phonological code for written pseudowords; and a deficiency in judging the homophony of two printed pseudowords (Coltheart, 1983; Zaidel & Peters, 1981). Such results have been taken to indicate a right-hemisphere deficit in deriving phonology from orthography (Patterson & Besner, 1984).

Indications of lateralization within normal populations are sought in visual field advantages. With an intact corpus callosum, each hemisphere has access to the other's information, but the "specialist" ought to demonstrate a superiority in its domain. In lexical decision tasks, for example, in which a subject must simply decide whether or not a given letter string is a word, a right visual field (RVF) advantage is typically found (Bradshaw & Gates, 1978; Bradshaw, Gates, & Nettleton, 1977). Unfortunately, this may be the only straightforward result. When interactions with visual field are investigated, the findings tend to be equivocal. We will review these and suggest how the disparate observations might be clarified.

For example, the RVF superiority may be greater for accepting words than rejecting pseudowords. This interaction is found when the pseudowords are chosen to preserve the "structure and shape of the words employed" (Bradshaw & Gates, 1978) but not when the pseudowords are homophonic with real words (Bradshaw et al., 1977). In the latter case, the error analysis of the right-handed subjects revealed the same tendency, however–fewer errors were made on words when they appeared in the RVF while fewer errors occurred for pseudowords in the LVF. But even this error pattern changed when nonwords (orthographically illegal letter strings) were used, yet they seem as suited to rejection by right hemisphere visual analysis as pseudohomophones.

There is also some suggestion that RVF superiority may be greater for males than females, at least for right-handed subjects (Bradshaw et al., 1977). This difference disappears when the letter strings reoccur in the second half of the experiment, suggesting that the right hemisphere of females may be involved in language processing of unfamiliar material only (Bradshaw & Gates, 1978). But in a modified lexical decision task in which subjects had to decide whether or not a pseudoword was homophonic with a real word, the interaction of sex and visual field was not significant (although Bradshaw et al., 1977, noted that 75% of their subjects of both sexes showed visual field differences in the expected direction in Experiment 2). In their Experiment 1, Bradshaw and Gates (1978) found a sex by lexicality by visual field interaction in the error analysis (with the RVF advantage for words being greater in males than females). If different types of pseudowords and nonwords show differential degrees of RVF superiority (contrast Bradshaw & Gates, 1978, with Bradshaw et al., 1977) and if the lexicality-visual field interaction differs for the sexes (suggested in Bradshaw & Gates, 1978, Experiment 1), then some types of pseudowords might be expected to obscure the interaction of visual field and sex. Orthographically regular, nonhomophonic pseudowords may be important to revealing sex by visual field interactions.

Of course, if one considers the lexical decision task to represent a fairly unsophisticated linguistic skill, then any hemispheric differences so revealed may be irrelevant to the language capacity of the right hemisphere. Indeed, the ability to perform lexical decision has been found in left-hemisphere stroke patients who are globally aphasic (Glass, Gazzaniga, & Premack, 1973). Investigations of the phonological capacity of the two hemispheres, therefore, may be more to the point. In this regard, a RVF advantage (in terms of accuracy of report) has been found for pronounceable (defined as orthographically legal) letter strings but not for unpronounceable (orthographically illegal) strings, even though both were reported as a series of letters (Turvey, Feldman, & Lukatela, 1984). In a naming task (in which response latency to pronounce a letter string is measured) the RVF advantage was greater for pseudohomophones than for either words or nonhomophonic pseudowords (Bradshaw & Gates, 1978, Experiment 4). No such interaction was found, however, for the lexical decision-like task that compared homophonic pseudowords with nonhomophonic pseudowords (despite its mention in the general discussion [Bradshaw & Gates, 1978]). Finally, in lexical decision tasks some investigators find a pseudohomophone effect for RVF presentation only (Cohen & Freeman, 1978), but others find such an effect for both RVF and LVF presentations (Barry, 1981).

Although phonology is an important property in defining language ability, the complex script-sound relationship in English makes it a tricky property to manipulate (Parkin, 1982). Pseudohomophones are somewhat suspect as an adequate test of whether or not phonological analysis is taking place (Feldman, Lukatela, & Turvey, 1985; Martin, 1982). Investigations of phonological effects are more straightforward in Serbo-Croat, a Slavic language used by the majority of Yugoslavs. Where English uses word order as its primary grammatical device, Serbo-Croat uses inflection. There is another interesting contrast with English: Due to deliberate alphabet reform in the last century, the Serbo-Croatian orthography is phonologically shallow—a given grapheme has only one pronunciation. In addition, the single spoken language is transcribed in two distinct scripts, both of which can be read by educated Yugoslavs. The Cyrillic script (predominant in the East) and the Roman script (predominant in the West) were modified to map onto the same 30 sounds. In the uppercase, printed form, most of the letters are unique to one or the other alphabet, seven are common (i.e., receive the same phonetic interpretation in Roman and Cyrillic) and, importantly, four are ambiguous (i.e., receive one interpretation in Roman but a different interpretation in Cyrillic). Combinations of unique and common letters produce phonologically unambiguous letter strings (e.g., ЛОЗА /loza/ means "vineyard" in Cyrillic but cannot be read in Roman; NOGA /noga/ means "leg" in Roman but cannot be read in Cyrillic). Combinations of ambiguous and common letters yield phonologically ambiguous or bivalent letter strings (e.g., POTOP means "motor" in Cyrillic /rotor/ but "flood" in Roman /potop/; BEHA means "vein" in Cyrillic /vena/ but is a nonword—though orthographically legal—in Roman /bexa/).

A comparison of lexical decision time for a phonologically ambiguous string that is a word in Cyrillic (BEHA), for example, with the latency for that same word written in unique Roman (VENA) provides a measure of the effect of phonological ambiguity (with lexicality, frequency, meaningfulness, orthographic structure, and so on, controlled). Where more than one phonological code is assembled, it is argued that this retards either lexical search or lexical activation time (see Feldman & Turvey, 1983, and Lukatela et al., 1980, for a summary of the arguments). Indeed, the effect has been shown to be quite robust (on the order of 300 ms) as phonological analysis appears to be nonoptional for readers of Serbo-Croat (Carello & Turvey, in press; Feldman & Turvey, 1983; Turvey et al., 1984).[1]

---

[1] An ambiguous letter string takes longer to decide about than an unambiguous letter string when the former is (i) a word, though different, in both readings; (ii) a pseudoword, though different, in both readings; and (iii) a word in one reading and a pseudoword in the other (where the affirmative lexical decision dictates that the unambiguous control be a word) (Lukatela et al., 1978, 1980). The effect is more pronounced with words than pseudowords (Feldman & Turvey, 1983; Lukatela et al., 1978). The greater the number of ambiguous letters in the string, the longer lexical decision takes (Feldman et al., 1983; Feldman & Turvey, 1983). While attempts

This cleaner manipulation of phonology would make it easier to evaluate phonological involvement in the two hemispheres. The logic is similar to that used to evaluate reading strategies employed by good and poor readers: Those who are more reliant on phonology ought to be slowed more by phonological bivalence. With respect to reading skill, this turns out to be the good readers (Feldman et al., 1985). With respect to language specialization, we would expect a larger effect in the left hemisphere (RVF). Moreover, if females show less of a difference between the hemispheres than males in general, then the phonological ambiguity effect in particular ought to be larger in the females' LVF than the males' LVF.

The present study uses Serbo-Croatian letter strings in a lexical decision task in an effort to clarify a number of issues raised here. In general, we expect shorter latencies for RVF presentations. But this will be examined in light of the lexical status of the items (i.e., is the RVF advantage more dramatic for words than pseudowords?), and sex of the subject (i.e., are males more lateralized than females?). In addition, irrespective of lexicality we will look at the influence of phonological ambiguity on processing in the two hemispheres. Response latency to a phonologically ambiguous Cyrillic word (e.g., BEHA /vena/), which is a pseudoword in Roman (/bexa/), ought to be longer than latency to that same word written in pure Roman (VENA /vena/).[2] Similarly, a phonologically ambiguous pseudoword (HABA), which has both a Cyrillic (/nava/) and a Roman (/xaba/) interpretation ought to take longer to reject than that same pseudoword written in pure Roman (NAVA /nava/). Of interest is whether or not the phonological ambiguity effect is larger for RVF presentations. It should be underscored that we are looking at sensitivity to phonological ambiguity as indexed by the degree of increase in latency *relative to the unambiguous base line*. Thus, if phonologically ambiguous letter strings show less of a RVF advantage than unambiguous letter strings, it does not mean that the right hemisphere is more involved in processing the former. Rather, it suggests that the left hemisphere is more sensitive to ambiguity in phonological interpretation.

Method

Subjects

Fifty-six male undergraduates from the Faculty of Electrical Engineering at the University of Belgrade and 56 female high school seniors from the Sixth Belgrade Gymnasium served voluntarily

---

to bias subjects toward a Roman reading by instructions or task (i.e., uniquely Cyrillic letters never appear) did not eliminate the effect, the presence of a single unique character did (Feldman et al., 1983; Lukatela et al., 1978). Finally, the effect is more pronounced in good readers than in poor readers (Feldman et al., 1985), suggesting that those who more effectively exploit the phonologically analytic strategy are harmed more by ambiguity.

[2] Phonologically ambiguous letter strings (OBOJICA) that are a word in Roman ("both" /obojitsa/) but a pseudoword in the Cyrillic reading (/ovojisa/) do not occur in sufficient numbers to allow experimentation with appropriate control on frequency and structure. It should be emphasized that the phonological ambiguity effect does not simply reflect alphabet differences, however. Comparisons of a word (e.g., "leg") written in pure Roman (NOGA /noga/) or pure Cyrillic (HOГA /noga/) reveal little, if any, difference (Feldman & Turvey, 1983). Moreover, a phonological ambiguity effect is found completely within alphabet–that is, when phonologically bivalent Cyrillic words are compared to different words written in pure Cyrillic (Lukatela et al., 1978).

as subjects. All had normal or corrected to normal vision, and none had had previous experience with visual processing experiments.

Procedure

A subject sat with his/her head in a chin rest at the viewer of a two-channel, wide field-of-view tachistoscope specially constructed by Dr. M. Gurjanov of the Faculty of Electrical Engineering at the University of Belgrade. Subjects were instructed to fixate a small, bright, centrally located point of light, which was visible throughout the experiment. Each trial was preceded by a ready signal from the experimenter. Stimuli were presented unilaterally (varied pseudorandomly) for 150 ms with an intertrial interval of approximately 3 s. The subjects' task was to decide, as rapidly as possible, whether or not a letter string was a word by either its Roman or Cyrillic reading. Decisions were indicated by depressing a telegraph key with both thumbs for a "No" response or by depressing a slightly further key with both forefingers for a "Yes" response. Latency was measured from the onset of a slide. A blank field immediately preceded and followed the display interval. Each subject viewed 76 slides, which included 16 practice trials.

Stimuli

For the critical comparison, defined across two groups of subjects, two types of words and pseudowords were included. Phonologically ambiguous items (10 words and 10 pseudowords) consisted of CVCVC strings in which vowels and the middle consonant were common to the two alphabets but the initial and final consonants could receive different phonetic interpretations in Roman and Cyrillic (Group 1). Unambiguous items were pure Roman versions (10 words and 10 pseudowords) of the bivalent letter strings (Group 2). In addition, subjects saw two types of filler items. Pure Cyrillic letter strings (10 words and 10 pseudowords) were included in Group 2 to balance the number of Cyrillic items shown to the two groups (Group 1 saw the pure Roman version of these). Additional pure Roman letter strings (20 words and 20 pseudowords) were included in both groups in an attempt to bias the reading of the bivalent items and thereby enhance the phonological ambiguity effect. The available evidence suggests, however, that such long term bias (i.e., defined over the course of the experiment) has little influence on the size of the effect (compare Lukatela et al., 1980, with Lukatela, Savić, Gligorijević, Ognjenović, & Turvey, 1978).

All words were selected from the mid-frequency range of word frequencies for Serbian elementary school children (Lukić, 1983). Pseudowords were derived from the words by replacing one or two consonants in the letter string. All letter strings were printed in uppercase Roman or Cyrillic (IBM Gothic). Each string of five letters was arranged horizontally to the left of center of a 35 mm slide for RVF presentation and to the right of center (on a complementary set) for LVF presentation. Stimuli subtended a horizontal visual angle of 2 deg, commencing 1 deg to the left or right of the midline of the slide.

Design

The major constraint of the design of the experiment was that a given subject never encountered a given word or pseudoword more than once. This was achieved by a further subdivision of groups with respect to hemifield presentation. Group 1a saw half of the words and half of the pseudowords in the LVF and the other halves in the RVF. For Group 1b, the reverse was true. The same pattern occurred for Groups 2a and 2b. Equal numbers of males and females were in each subgroup.

Results

Minimum and maximum acceptable latencies were set at 350 and 3500 ms, respectively (these are fairly long to allow for the difficulty of judging phonologically ambiguous stimuli presented unilaterally). The data (Figures 1 and 2) were first subjected to an analysis of variance by Sex, Visual Field, and Lexicality, collapsing the phonologically ambiguous letter strings and their unambiguous equivalents. The main effect of Field was significant, $F(1,110) = 20.53, MSe = 303212, p < .001$, with the right visual field (936 ms) being faster than the left (988 ms). Words were accepted faster (885 ms) than pseudowords were rejected (1038 ms), $F(1,110) = 48.20, MSe = 2606988, p < .001$. But neither the main effect of Sex, $F(1,110) < 1$, nor any interaction was significant (all $F < 1$ except Field × Lexicality, $F(1,110) = 1.81, MSe = 26893, p > .15$).

A second analysis focused on the relationship between phonological ambiguity and lexicality in the two visual fields. Again, the main effects of Field, $F(1,110) = 20.59, MSe = 303212, p < .001$, and Lexicality, $F(1,110) = 54.51, MSe = 260988, p < .001$, were significant, as was Phonology, $F(1,110) = 60.13, MSe = 9943366, p < .001$, with ambiguous strings being slower (1111 ms) than their unambiguous controls (813 ms). In addition, the Lexicality × Phonology interaction was significant, $F(1,110) = 14.39, MSe = 688367, p < .001$, with pseudowords suffering more (376 ms difference) than words (319 ms). No other interactions were significant: Field × Lexicality, $F(1,110) = 1.81, Mse = 26893, p > .15$, Field × Phonology and Field × Phonology × Lexicality, $F < 1$.

The words and pseudowords were then analyzed separately for Field by Phonology by Sex effects. Since variances were not homogenous in a first analysis, a square root transform was performed on the latency data. For Words, significant main effects were obtained for phonology, $F(1,108) = 52.89, MSe = 739, p < .001$ (ambiguous words averaged 995 ms, unambiguous averaged 775 ms) and Field, $F(1,108) = 17.03, MSe = 66, p < .001$ (LVF = 919 ms, RVF = 852 ms), but not for Sex, $F(1,108) = 1.26, MSe = 18, p > .25$. The Field × Phonology × Sex interaction was marginally significant, $F(1,110) = 3.63, MSe = 14, p < .056$ (the phonological ambiguity effect in the LVF was larger for females than males but approximately equivalent for the two sexes in the RVF, as shown in Table 1). This interaction was significant in the error analysis, $F(1,108) = 5.28, MSe = 3, p < .02$, but in the opposite direction (the increase in errors for bivalent items relative to their Roman controls was greater in the RVF for females but in the LVF for males. See Table 2.)

Table 1

Difference (in ms) between phonologically ambiguous words and their unambiguous equivalents in the two visual fields for males and females.

	Visual Field	
Sex	Left	Right
Male	160	236
Female	260	218

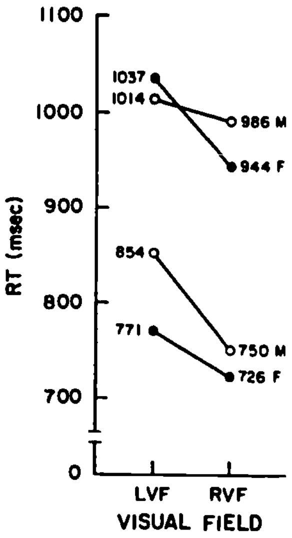


Figure 1. Average lexical decision times (in ms) to phonologically ambiguous words (upper lines) and their unambiguous equivalents (lower lines) presented to the left and right visual fields for males and females.

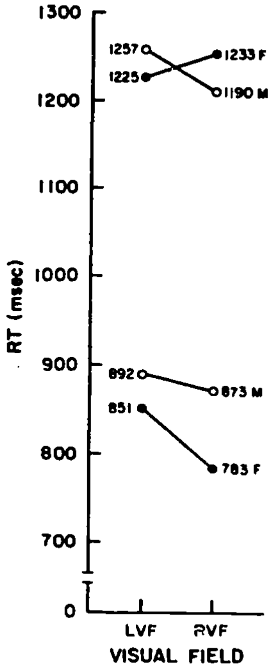


Figure 2. Average lexical decision times (in ms) to phonologically ambiguous pseudowords (upper lines) and their unambiguous equivalents (lower lines) presented to the left and right visual fields for males and females.

Table 2

Difference (in average number of errors per 5 items) between phonologically ambiguous words and their unambiguous equivalents in the two visual fields for males and females.

	Visual Field	
Sex	Left	Right
Male	1.0	0.4
Female	0.5	0.8

The error analysis also revealed main effects of Phonology, $F(1,108) = 20.73, MSe = 26, p < .001$ (on average, 1.2 errors—out of 5 per visual field—were made on ambiguous words, .5 on unambiguous words), and Sex, $F(1,108) = 4.15, MSe = 5, p < .05$ (females committed fewer errors [0.7] than males [1.0]). There was also a field by sex interaction, $F(1,108) = 7.03, MSe = 4, p < .01$ (females committed fewer errors in the LVF than the RVF, males committed more LVF errors than RVF errors).

For Pseudowords, as with Words, significant main effects were obtained for Phonology, $F(1,108) = 54.71, MSe = 1768, p < .001$ (ambiguous pseudowords averaged 1226 ms, unambiguous averaged 850 ms) and Field, $F(1,108) = 6.11, MSe = 17, p < .01 (LVF = 1056$ ms, $RVF = 1020$ ms), but not for Sex, $F < 1$. And, once again, the Field × Phonology × Sex interaction was significant, $F(1,108) = 5.02, MSe = 14, p < .025$. This time, the direction of the interaction was reversed—the ambiguity effect in the RVF was larger for females than males and approximately equivalent in the LVF (Table 3). Due to the small number of errors on pseudowords, an error analysis could not be performed.

Table 3

Difference (in ms) between phonologically ambiguous words and their unambiguous equivalents in the two visual fields for males and females.

	Visual Field	
Sex	Left	Right
Male	365	317
Female	374	450

Discussion

The experiment was designed in an effort to clarify some disparate findings with regard to the differential language capacity of the two hemispheres in normal subjects with undivided commissures. With laterally presented verbal material, a left hemisphere (RVF) advantage was found irrespective of an item's lexical status (word or pseudoword) or phonological interpretation (ambiguous or unambiguous). This was as true for females as it was for males. We did not find an interaction of visual field with either sex or lexicality. These results are consistent with the patterns found with naming latencies (Bradshsw & Gates, 1978, Experiments 3 and 4) and identification proportions (Young, Ellis & Bion, 1984) but differ somewhat from other lexical decision studies. Bradshaw and Gates (1978) found the RVF advantage to be greater for words than structurally similar pseudowords (Experiment 1) but this was really only true of their female subjects. A lexicality by field interaction did not obtain for word-pseudohomophone (Bradshaw et al., 1977) or pseudohomophone-nonhomophone comparisons (Bradshaw & Gates, 1978, Experiment 2). Word-nonword (consonant string) comparisons were mixed, depending on sex and handedness (Bradshaw et al., 1977).

Bradshaw and his colleagues suggest that the RVF advantage is greater for males than females but, in fact, this difference was not absolute—it was true of right handers but not left handers (Bradshaw et al., 1977) and it disappeared when stimuli were presented a second time (Bradshaw & Gates, 1978). Although their field by sex interaction was not significant, Bradshaw and Gates (1978) attribute this to a few subjects since 75% of the males showed a RVF superiority and 75% of the females showed a LVF superiority. We did not find this trend, though, as 69% of the males and 65% of the females showed a RVF superiority. Again, this is consistent with the naming and identification studies that did not obtain field differences for males and females. This is not to dispute hemispheric differences between the sexes but to point out that their appearance in three-way interactions with selected tasks may indicate that we are looking at quite subtle but potentially more interesting nuances of suspected differences in language skill.

In the present study, these interesting differences begin to appear in an examination of the phonological ambiguity effect. The general slowing of responses to ambiguous strings relative to their unambiguous equivalents (Feldman, Kostić, Lukatela, & Turvey, 1983; Feldman & Turvey, 1983) was replicated. The effect was larger for pseudowords, however, which is counter to the usual finding with centrally presented letter strings (Feldman & Turvey, 1983; Lukatela et al., 1978).

Of particular interest, of course, are the Sex × Phonology × Field interactions. If one looks just at the latency differences for words (Table 1), two cognate descriptions can be given. First, the phonological ambiguity effect in the left hemisphere (RVF) is comparable for the two sexes but in the right hemisphere (LVF) it is greater for females. Second, the phonological ambiguity effect in females is larger in the right hemisphere than in the left; in males, it is larger in the left hemisphere than in the right. This pattern is consistent with suggestions in the literature that if the right hemisphere is linguistically competent (particularly with respect to something as sophisticated as phonology), then it is more likely to be so in females. It is then conjectured that this "extra" language capacity in the right hemisphere underlies the traditional female verbal superiority (and spatial inferiority) as spatial processing space has been usurped by language (Bradshaw et al., 1977).

Turning to pseudowords, the opposite pattern is found. For males, phonological ambiguity is felt more strongly in the right hemisphere. A possible explanation of this outcome for male subjects is to be found in suggestions that the right hemisphere lexicon does not contain all words (Bradshaw & Gates, 1978; Day, 1977; Ellis & Shepherd, 1974; Hines, 1976, 1977; but see Patterson & Besner, 1984, for a summary of contrary results). To the extent that it does not, unfound letter strings would have to be shuttled to the left hemisphere for further search before negative lexical decisions could be made (Patterson & Besner, 1984). Since two or more phonological codes are assembled in the present experiment, a combined search-and-shuttle for each code could be expected to take longer than the more straightforward left-hemisphere search. The contrary result that is observed for females echoes the opposition displayed by the two sexes with regard to words, demonstrating some consistency to their differences.

We must, however, introduce a note of caution in interpreting the word latencies if they are viewed in conjunction with the error data (Table 2). A speed-accuracy trade off becomes apparent for both males and females: Smaller latency differences between phonologically ambiguous words and their baseline are accompanied by larger error differences. The possibility remains, therefore, that an emphasis on accuracy might reverse the hemispheric pattern of the phonological ambiguity effect in males and females. For pseudowords (Table 3), in contrast, the error rate was quite low (and positively correlated with latency). Does the lower error rate for pseudowords relative to words mean that the former provide a purer look into language differences in the hemispheres? While this possibility cannot be ruled out, we would claim, more conservatively, that strategies used to distinguish words from pseudowords are not homogeneous. (As examples, serial search models distinguish self-terminating and exhaustive searches for words and pseudowords, respectively; parallel activation models contrast exceeding the threshold of a single lexical entry with violating a general deadline for exceeding any item's threshold.) Further understanding of the data presented here awaits an elucidation of the mechanisms underlying lexical decision—recall, for example, that this task can be performed by some global aphasics (Glass et al., 1973). But returning to the major question of phonological sensitivity, it can be concluded unequivocally that the demands of phonological processing are not met in the same way by the two hemispheres nor do the two hemispheres respond in the same way in the two sexes.

References

Barry, C. (1981). Hemispheric asymmetry in lexical access and phonological encoding. *Neuropsychologia 19*, 473-478.

Bradshaw, J. L., & Gates, E. A. (1978). Visual differences in verbal tasks: Effects of task familiarity and sex of subject: *Brain and Language, 5*, 166-187.

Bradshaw, J. L., Gates, E. A., & Nettleton, N. C. (1977). Bihemispheric involvement in lexical decisions: Handedness and a possible sex difference. *Neuropsychologia, 15*, 277-286.

Carello, C., & Turvey, M. T. (in press). Dependence of reading on orthography: Investigations in Serbo-Croatian. In H. Kao & O. Tzeng (Eds.), *Psychological studies in the Chinese language II*. Hong Kong: Hong Kong University Press.

Cohen, G., & Freeman, R. (1978). Individual differences in reading strategies in relation to handedness and cerebral asymmetry. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, NJ: Erlbaum.

Coltheart, M. (1980). A right hemisphere hypothesis. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia*. London: Routledge & Keegan Paul.

Coltheart, M. (1983). The right hemisphere and disorders of reading. In A. Young (Ed.), *Functions of the right cerebral hemisphere*. London: Academic Press.

Day, J. (1977). Right hemisphere language processing in normal righthanders. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 518-528.

Ellis, H. D., & Shepherd, J. W. (1974). Recognition of abstract and concrete words presented in left and right visual fields. *Journal of Experimental Psychology, 103*, 1035-1036.

Feldman, L. B., Kostić, A., Lukatela, G., & Turvey, M. T. (1983). An evaluation of the "Basic Orthographic Syllabic Structure" in a phonologically shallow orthography. *Psychological Research, 45*, 55-72.

Feldman, L. B., Lukatela, G., & Turvey, M. T. (1985). Effects of phonological ambiguity on beginning readers of Serbo-Croatian. *Journal of Experimental Child Psychology, 29*, 492-510.

Feldman, L. B., & Turvey, M. T. (1983). Visual word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 288-298.

Gazzaniga, M. S. (1983). Right hemisphere language following brain bisection: A 20-year perspective. *American Psychologist, 38*, 525-537; and Reply to Levy and to Zaidel, 547-549.

Glass, A. S., Gazzaniga, M. S., & Premack, D. (1973). Artificial language training in global aphasics. *Neuropsychologia, 11*, 95-103.

Hines, D. (1976). Recognition of verbs, abstract nouns and concrete nouns from left and right visual half fields. *Neuropsychologia, 14*, 211-216.

Hines, D. (1977). Differences in tachistoscopic recognition between abstract and concrete words as a function of visual half-field and frequency. *Cortex, 13*, 66-73.

Levy, J. (1983). Language, cognition, and the right hemisphere: A response to Gazzaniga. *American Psychologist, 38*, 538-541.

Lukatela, G., Popadić, D., Ognjenović, P., & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition, 8*, 124-132.

Lukatela, G., Savić, M., Gligorijević, B., Ognjenović, P., & Turvey, M. T. (1978). Bi-alphabetical lexical decision. *Language and Speech, 21*, 142-165.

Lukić, V. (1983). *Frequency dictionary of children*. Belgrade.

Martin, R. C. (1982). The pseudohomophone effect: The role of visual similarity in non-word decision. *Quarterly Journal of Experimental Psychology, 34A*, 395-409.

Parkin, A. (1982). Phonological recoding in lexical decision: Effects of spelling-to-sound regularity depend on how regularity is defined. *Memory & Cognition, 10*, 43-53.

Patterson, K., & Besner, D. (1984). Is the right hemisphere literate? *Cognitive Neuropsychology, 1*, 315-341.

Saffran, E. M., & Marin, O. S. (1977). Reading without phonology: Evidence from aphasia. Quarterly Journal of Experimental Psychology, 29, 515-525.

Turvey, M. T., Feldman, L. B., & Lukatela, G. (1984). The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy. In L. Henderson (Ed.), *Orthographics and reading*. London: Erlbaum.

Young, A. W., Ellis, A. W., & Bion, P. J. (1984). Left hemisphere superiority for pronounceable nonwords, but not for unpronounceable letter strings. *Brain and Language, 22*, 14-25.

Zaidel, E. (1983). A response to Gazzaniga: Language in the right hemisphere, convergent perspectives. *American Psychologist, 38*, 542-546.

Zaidel, E., & Peters, A. M. (1981). Phonological encoding and ideographic reading by the disconnected right hemisphere: Two case studies. *Brain and Language, 14*, 205-234.

PROCESSING LEXICAL AMBIGUITY AND VISUAL WORD RECOGNITION IN A DEEP ORTHOGRAPHY*

Shlomo Bentin† and **Ram Frost**‡

Abstract. *We investigated the effect of semantic and phonemic ambiguity on lexical decision and naming performance in the deep Hebrew orthography. Experiment 1 revealed that lexical decisions for ambiguous consonant strings are faster than those for any of the high- or low-frequency voweled alternative meanings of the same strings. These results suggested that lexical decisions for phonemically and semantically ambiguous Hebrew consonant strings are based on the ambiguous orthographic information. However, a significant frequency effect for both ambiguous and unambiguous words suggested that if vowels are present, subjects do not ignore them completely while making lexical decisions. Experiment 2 revealed that naming low-frequency voweled alternatives of ambiguous strings took significantly longer than naming the high-frequency alternatives or the unvoweled strings without a significant difference between the latter two string types. Voweled and unvoweled unambiguous strings, however, were named equally fast. We propose that semantic and phonological disambiguation of unvoweled words in Hebrew is achieved in parallel to the lexical decision, but is not required by it. Naming Hebrew words usually requires a readout of phonemic information from the lexicon.*

This study examined the influence of phonemic and semantic ambiguity on lexical decision and naming performance in Hebrew. Most lexical ambiguity research has concentrated primarily on the way homographs (polysemous words) are disambiguated within a semantic context (for a recent review see Simpson, 1984). Little has been said about lexical access of isolated ambiguous words or about ambiguity effects on the process of word recognition.

Examination of lexical ambiguity in relation to lexical access and models of word recognition is of special interest in Hebrew orthography because of its special way of conveying phonemic

---

* *Memory & Cognition*, 1987, *15*, 13-23.

† Also, Aranne Laboratory of Human Psychophysiology, Department of Neurology, Hadassah Hospital, and Department of Psychology, Hebrew University, Jerusalem, Israel.

‡ Also, Department of Psychology, Hebrew University, Jerusalem, Israel.

Aknowledgment. We thank Drs. Morris Moscovitch, Leonard Katz, and James Chumbley for useful comments on earlier drafts of this paper, and Tammi Rossman for her skillful help. This work was supported in part by National Institute of Child Health and Human Development Grant HD-01994 to Haskins Laboratories.

information. In Hebrew, the vowels are signified by small diacritical dots and dashes, and letters carry mostly consonantal information. In most printed texts (e.g., books, magazines, newspapers, etc.), the diacritical vowel marks are omitted. (For a detailed description of Hebrew orthography, see Navon & Shimron, 1984). Because several words may share an identical consonantal structure, an unequivocal lexical (and phonemic) representation can be attributed to a letter string only through a top-down process, which is usually affected by semantic context. In other words, homographs are the rule rather than the exception in Hebrew, and processing lexical ambiguity is a routine procedure for the reader. The nature of ambiguity in Hebrew consonantal strings is different, however, from that in English homographs. Ambiguity in Hebrew print results from missing vowel information. Thus, in contrast to English, where, with a few exceptions (see, e.g., Carpenter & Daneman, 1981), most homographs are also homophones, Hebrew consonant strings are not only semantically but also phonemically ambiguous. Moreover, the number of different etymological derivations represented by one letter string is usually much higher than in English. Consequently, there is greater uncertainty about which word is represented by a given consonantal string in Hebrew than there is about an English homograph.

It is because of this high level of uncertainty that we assumed that lexical decisions for Hebrew homographs might be made without phonological disambiguation. Rather, it is possible that an abstract orthographic representation, common to all alternative meanings of a consonant string, provides, in most cases, the necessary and sufficient information for the lexical decision.

Previous research has already suggested that lexical decisions in Hebrew are based primarily on graphemic/orthographic cues (Bentin, Bargai, & Katz, 1984; Koriat, 1984), but the nature of this process has never been elaborated. One possible assumption is that although lexical access is mediated primarily by orthographic codes, the decision is delayed until one phonological alternative is determined. Thus, a specific word must be deciphered from print before the positive decision can be generated. A second possibility is that a positive decision can be based on the phonologically ambiguous homographic cluster common to all alternatives, and that lexical disambiguation is a subsequent process. We attempted to investigate the validity of each of these assumptions by comparing lexical decision performance on Hebrew consonant strings in the unvoweled ambiguous form and those disambiguated by the vowel marks.[1]

Early studies revealed that, in English, lexical decisions for homographs are faster than for nonhomographs (Rubenstein, Garfield, & Millikan, 1970; Rubenstein, Lewis, & Rubenstein, 1971). One common explanation for this effect of homography is that words with multiple meanings have multiple entries and, therefore, during lexical search, the probability of encountering one of many entries of a word with a large number of meanings is greater than the probability of encountering the single entry of a word with only one meaning. Recently, Jastrzembski (1981) elaborated this model on the basis of Morton's logogen theory (1970, 1979) assuming that words having multiple meanings are represented by one logogen for each meaning. Jastrzembski assumed that logogens accumulate evidence in a probabilistic fashion. Thus, the more logogens a word has, the more likely one of them will reach threshold. Accordingly, a word with many meanings, and therefore many logogens, will be more likely to have one of its logogens reach threshold

---

[1] If the vowel marks are presented in conjunction with the consonants (usually placed below the letters), they disambiguate the phonology and, in most cases, unequivocally determine one word. Thus, in a fully vowelized system, the reader may use the phonemic cues provided by the vowel marks to aid lexical access.

sooner than a word with few logogens. Note, however, that all the probabilistic models of lexical access that were proposed to account for the effect of homography were based on responses to homographs that were also homophones. In contrast, it has been reported that, in English, the latency of naming homographs that have different pronunciations for different meanings (e.g., WIND) is significantly *longer* than that of naming homographs with a single pronunciation (e.g., FALL) (Kroll & Schweickert, 1978).

These data suggest that, in English, phonological ambiguity must be resolved before the lexical decision is generated. If the same strategy is employed in Hebrew, lexical decisions about unvoweled, phonologically ambiguous strings should be slower than decisions about any of their explicitly voweled, phonologically disambiguted alternatives. On the other hand, faster responses to the ambiguous homograph than to any of its disambiguated alternatives would suggest that phonological disambiguation is not required for lexical decision. We will claim that such a result supports the possibility that lexical decisions for phonologically ambiguous consonant strings are based primarily on the abstract orthographic representation, which is common to all phonological alternatives. In other words, such a result might suggest that the different phonological representations of a Hebrew consonant string are related to only one visual logogen.

Experiment 1

Experiment 1 compared lexical decision performance for unvoweled ambiguous consonant strings and for their voweled disambiguated alternatives. Each ambiguous consonant string could represent either a high-frequency or a low-frequency word. The same voweled words were presented in two task conditions that were determined by the nature of the nonwords employed. In the optional condition, the nonwords were legal but meaningless permutations of the consonant letters of real words. In this condition, words and nonwords were distinguishable simply on the basis of the consonant pattern, so that there was no logical necessity for the subject to attend to the vowels in order to make a lexical decision. In the obligatory condition, special nonwords were used. The consontal pattern of these nonwords corresponded to real words in Hebrew, but they were presented in an inappropriately voweled manner. Thus, the special nonwords employed in the obligatory condition were in fact words when unvoweled or when voweled correctly; therefore, in this condition, the subject was forced to attend to the vowels, without which discrimination between words and nonwords could not have been made. Recall that the difference between the high- and the low-frequency alternatives of each homograph was indicated in print only by the vowel marks. Therefore, the relative sizes of the frequency effect in the optional and obligatory condition may reflect the extent to which the information provided by the vowels was processed in each of these conditions by subjects while generating the lexical decision.

In addition to comparing performance for ambiguous strings, Experiment 1 also compared lexical decision performance for voweled and unvoweled unambiguous words. Unambiguous words were consonant strings that represented only one legal phonological derivation. Thus, although the phonology presented in print was not complete if vowel marks were missing, access to these words did not require a choice between different phonological alternatives. Therefore, any difference between voweled and unvoweled unambiguous words should reflect only the effect of missing phonemic information, but no effects of ambiguity. Comparison of lexical decision performance for ambiguous and unambiguous words in the optional and obligatory conditions is particularly important because, regardless of the ambiguity of the words, different nonwords might affect performance on low- and high-frequency words differently (Duchek & Neely, 1984, cited by Balota & Chumbley, 1984; James, 1975).

If lexical decisions for unvoweled and for voweled words with ambiguous consonant strings in the optional condition are based primarily on the ambiguous consonantal information, then (1) the reaction times (RTs) for unvoweled ambiguous words should not be longer than RTs for any of their voweled high- and low-frequency disambiguated alternatives, and (2) word frequency should have a larger effect in the obligatory than in the optional condition (since subjects in the obligatory condition were forced to process the vowels in order to discriminate between words and nonwords). On the other hand, if phonological disambiguation is mandatory for lexical decisions and if positive responses for Hebrew homographs are based on successful access to an entry uniquely related to one alternative meaning, then (1) RTs for the unvoweled words should be longer than for the voweled words (since vowels simplify the phonological disambiguation of the ambiguous consonant string), and (2) word frequency should have a similar effect on lexical decisions for ambiguous and unambiguous words in the optional and in the obligatory conditions.

Method

Subjects. Ninety-six undergraduate students participated, either as part of the requirements of an introductory psychology course, or were paid for their participation. They were all native speakers of Hebrew with normal or corrected-to-normal vision.

Stimuli. The ambiguous words were 16 unvoweled consonant strings. Each string represented a pair of different nouns with different meanings, different pronunciations, and different word frequencies. Lexical decision performance for each of these unvoweled strings was compared with lexical decision performance for its two voweled alternate representations (Figure 1). The 16 pairs of nouns were selected from a pool of 50 similar pairs on the basis of frequency evaluation, as follows. The 100 words were printed with the vowel marks on a single page in random order. Fifty undergraduates were asked to rate each word on a 5-point scale ranging from *least frequent* (1) to *most frequent* (5). Estimated frequency for each word was calculated by averaging the rating across all 50 judges. Words rated above 3.5 and below 2.5 were included in the high- and low-frequency groups, respectively. The 16 word pairs selected for this study were those in which one member was rated high and the other was rated low in frequency. The mean rating of the high-frequency group was 4.23 and of the low-frequency group was 1.69. Although in the voweled form these consonant strings were unequivocally specified, because their consonantal structures were shared by different words, we labeled the critical pairs as ambiguous and differentiated between voweled and unvoweled ambiguous words.

The unambiguous words were 32 consonant strings, each of which represents in print only one word. Sixteen of the 32 ambiguous words were high-frequency and 16 were low-frequency, according to the same criteria as above. Rating was performed by a different group of 50 judges in a similar manner (Frost, Katz, & Bentin, 1987). The mean rating was 4.18 for the high-frequency unambiguous group and 1.71 for the low-frequency unambiguous group.

The nonwords in the optional condition were 32 pronounceable but meaningless permutations of consonants arbitrarily assigned with vowel marks. We shall label these stimuli *regular nonwords*. Note that regular nonwords could not be read as words even if presented unvoweled. In contrast, in the obligatory condition, the nonwords were in fact words when unvoweled or when

	High Frequency Alternative	Low Frequency Alternative	Unvoweled Print
Graphemic representation	דָבָר	דֶבֶר	דבר
Phonetic transcription	davar	dɛvɛr	?
English translation	thing	pest	?

Figure 1. Example of ambiguity in Hebrew print. Note that the two words represented by the same consonant string are not homophones.

voweled correctly; the consonantal patterns corresponded to words, but in this study they were inappropriately voweled. These nonwords were labeled *special nonwords*. The special nonwords were generated from words of average frequency.

All stimuli were generated by a PDP 11/34 computer and displayed at the center of a Tektronix CRT. The size of each letter was 1.2 X 1.2 cm, and the length of the whole word was between 4.5 and 6.5 cm (three to five letters), subtending a visual angle of approximately 4.5 degrees.

Design. Six test lists were assembled. In Lists A to D, all stimuli were presented with the vowel marks. In Lists E and F, the stimuli were presented in the regular unvoweled manner. Lists A and B were presented in the optional condition. List A comprised the 16 high-frequency alternatives of the ambiguous pairs, the 16 low-frequency unambiguous words, and the 32 regular nonwords. List B comprised the 16 low-frequency alternatives of the ambiguous pairs, the 16 high-frequency unambiguous words, and the same 32 nonwords used in List A. Lists C and D were presented in the obligatory condition. List C comprised the same words as List A, but the set of 32 special nonwords was employed. List D comprised the same words as List B and the special nonwords. Lists E and F were similar to Lists A and B, respectively, but words were presented without the vowel marks. Therefore, Lists E and F included identical ambiguous words (because the different alternatives of the ambiguous words were indistinguishable without the vowel marks) but different unambiguous words. Different groups of 16 subjects each were randomly assigned to each of these lists.

Procedure. The experiment took place in a semidarkened sound-treated room. Subjects sat approximately 70 cm from the screen. They were instructed to press one of two alternative microswitch buttons, according to whether the stimulus on the screen was or was not an actual Hebrew word. The dominant hand was always used for "yes" (i.e., word) responses and the other hand for the "no" (i.e., nonword) responses. Subjects presented with Lists C and D were warned about the special nature of the nonwords.

Following the instructions, 32 practice trials (16 words and 16 nonwords) were presented. Words/nonwords on the practice trials were prepared in congruence with those on the test list that followed. The 64 test trials were presented next at a rate of one stimulus every 2.5 sec. The subject's response terminated the stimulus exposure. If no response was given within 2 sec, the stimulus was removed and an error was marked. The subject started the test trials by pressing a ready button. RTs were measured in milliseconds and errors were marked.

Results

The RTs for correct responses were averaged for each word over the 16 subjects who were exposed to it, and for each subject over the 16 words in each frequency group. RTs that were above or below two standard deviations from a subject's or a word's mean were excluded, and the mean was recalculated. Less than 1.5% of the RTs were outliers.

We will describe first the comparison between the ambiguous consonant strings in the unvoweled and the two voweled presentations. The RTs to the high-frequency words in List A, to the low-frequency words in List B, and to the unvoweled ambiguous words in list E[2] were compared by a one-way ANOVA[3] and the Tukey-A post hoc procedure. RTs to unvoweled consonant strings were faster than to any of the two voweled alternatives, and RTs to the high-frequency alternatives were faster than to the low-frequency alternatives (Figure 2). The ANOVAs revealed that all differences were significant: for the stimulus analysis, $F(2,30) = 22.09, MSe = 4,521, p < 0.0001$; for the subject analysis, $F(2,45) = 5.15, MSe = 14,709, p < 0.01$; and $minF'(2,63) = 4.18, p < 0.05$.

The effect of the special nonwords on lexical decisions for high- and low-frequency voweled ambiguous words was assessed by comparing the RTs to ambiguous words in the optional condition (List A and List B) and in the obligatory condition (List C and List D) in a two-way (condition X frequency) ANOVA. RTs to ambiguous words were slower in the obligatory condition than in the optional condition, $minF'(1,24) = 10.96, p < 0.01$. Across conditions, RTs to high-frequency words were faster than to low-frequency words, $minF'(1,49) = 9.97, p < 0.01$. However, the

---

[2] Note that the ambiguous consonant strings in Lists E and F were identical. Therefore, there were twice as many subjects who responded to unvoweled ambiguous words than to each of the voweled alternatives. The RTs to unvoweled consonant strings in Lists E and F were not significantly different (616 ms and 621 ms, respectively). Therefore, in order to obtain an equal number of observations in each cell, we used only the RTs of the subjects presented with List E.

[3] Whenever appropriate, both stimulus and subject analyses were performed, and minF' was calculated. The stimulus analysis used a within-stimulus design and the subject analysis used a between-subjects design.

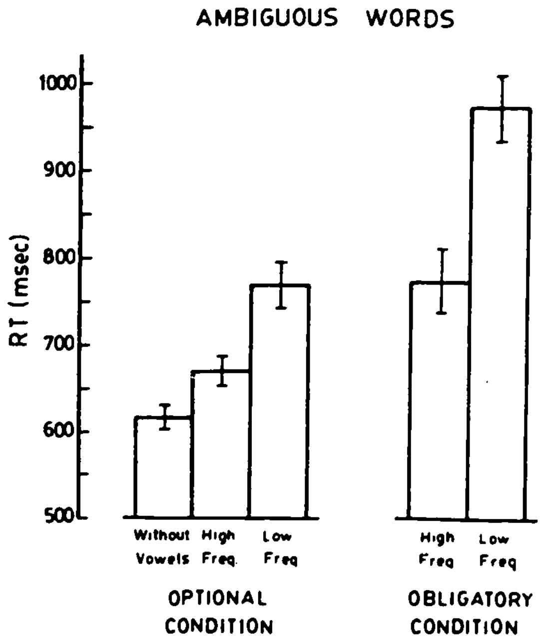


Figure 2. Mean RTs and SEs to ambiguous words in the lexical decision task. In the obligatory condition nonwords would become real words if presented without the vowel marks.

most important result of this comparison was that the frequency effect was twice as large in the obligatory condition as in the optional condition (Figure 2). This interaction was significant both for the stimulus analysis, $F(1,15) = 11.91, MSe = 3,346, p < 0.004$, and for the subject analysis, $F(1,60) = 4.25, MSe = 92,036, p < 0.05, minF'(1,73) = 3.13, p < 0.06$.

The effect of the vowel marks on high- and low-frequency unambiguous words was assessed by comparing the voweled high-frequency unambiguous words in List B and the voweled low-frequency unambiguous words in List A, with their unvoweled presentations in Lists E and F, respectively (Figure 3). A two-way (vowel condition × frequency) ANOVA revealed a significant frequency effect, $F(1,60) = 7.44, MSe = 12,334, p < 0.01$, but no effect of the vowel marks, and no interaction.

The effect of the special nonwords on lexical decisions for voweled unambiguous words was assessed by comparing the RTs to voweled high- and low-frequency unambiguous words in the optional condition (List B and List A), and in the obligatory condition (List D and List C). A condition X frequency two-way ANOVA revealed that lexical decisions were slower in the obligatory than in the optional condition, $F(1,60) = 9.48, MSe = 18,471, p < 0.005$, and slower for low-frequency than for high-frequency words, $F(1,60) = 3.84, MSe = 18,471, p < 0.05$, but in contrast to ambiguous words, there was no interaction between the two factors, $F(1,60) = 0.31$. The different effect of nonword condition on ambiguous and unambiguous words was verified by an ambiguity (ambiguous, unambiguous) × condition (optional, obligatory) X frequency (high, low) ANOVA, revealing a marginally significant three-way interaction, $F(1,120) = 3.57, MSe = 20067, p < 0.06$.

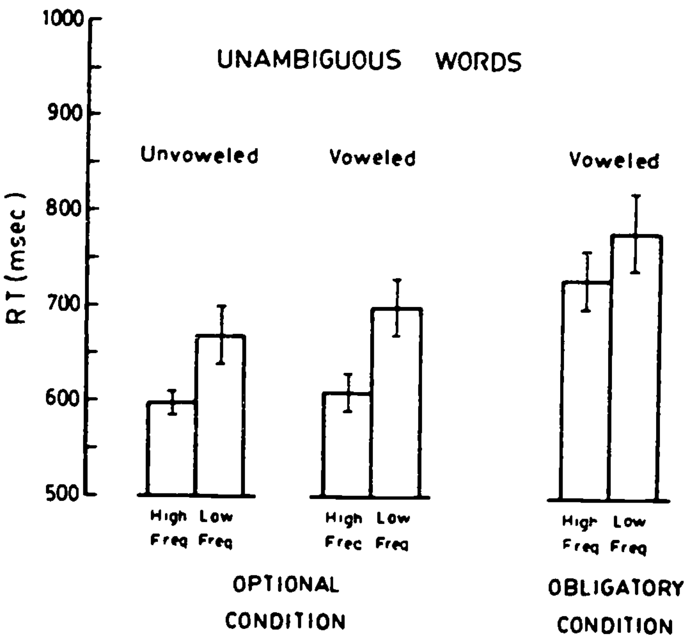


Figure 3. Mean RTs and SEs to unambiguous words in the lexical decision task. In the obligatory condition nonwords would become real words if presented without the vowel marks.

The nonwords in Lists A and B were identical, as were the nonwords in Lists C and D, and those in Lists E and F. Therefore, the RTs to nonwords in the optional condition (Lists A and B) were compared with the RTs to nonwords in the obligatory condition (Lists C and D) and with RTs to unvoweled nonwords (Lists E and F) by a one-way ANOVA and Tukey-A post-hoc procedure.[4] RTs to voweled nonwords in the optional condition and unvoweled nonwords were not significantly different (764 ms and 739 ms, respectively), and both were faster than the RTs to nonwords in the obligatory condition (1,055 ms), $F(2,93) = 36.16, MSe = 27,255, p < 0.0001$.

The distribution of errors between the different stimulus groups and conditions is presented in Table 1. Although the uneven distribution of errors within cells precluded a significant statistical analysis, the overall pattern does not suggest a speed-accuracy trade-off between conditions.

Discussion

The comparison of responses to voweled and unvoweled ambiguous words revealed that explicit presentation of the vowel marks that disambiguated the consonant strings did not facilitate lexical decisions relative to the decisions about ambiguous, unvoweled strings. In fact, an opposite effect was found. Apparently, this result simply replicates previous findings that suggested that lexical decisions for homographs are faster than for nonhomographs. However, we have obtained this result with consonant strings that represented in print not only different meanings, but also

---

[4] Initial analysis revealed that, indeed, the RTs to nonwords in List A were similar to those in List B, the RTs to nonwords in List C were similar to those in List D, and the RTs in List E were similar to those in List F.

different words (i.e., different phonological representations). Therefore, our results are in contrast to results obtained in similar studies about English (Kroll & Schweickert, 1978).

There are several ways to explain our data. The first explanation is based on the hypothesis that a lexical decision for an ambiguous unvoweled consonant string is based on accessing at least one of the words it represents. The existence of several possible entries increases the probability that one of them will be accessed. This explanation is unlikely for two reasons: (1) in English this effect was found only for homographs that represented different meanings but only one pronunciation; (2) the probabilistic explanation is based on the assumption that the distribution of the RTs to the different meanings represented by the homograph partially overlap. Although this is probably true in our study as well, we tried to diminish this overlap by selecting only homographs that represented both very high-frequency and very low-frequency words.

The second explanation of our data is based on the assumption that Hebrew readers are more familiar with unvoweled than with voweled words. This explanation, however, is contradicted by the absence of an effect of vowel marks on lexical decisions for unambiguous words and for nonwords.

The third explanation is congruent with our hypothesis. If lexical decisions for ambiguous consonant strings do not require phonological disambiguation, addition of vowel marks is unnecesary. Moreover, vowel marks add information that probably cannot be totally ignored and thus increases the word processing time. Therefore, we suggest that these data support the hypothesis that lexical decisions for unvoweled Hebrew homographs may be based on information that is common to all lexical alternatives that are represented by one consonant string. Decisions based on the common information should be at least as fast as those based on accessing specific entries, for the additional reason that the frequency of the common consonant string is higher than the frequency of each of its individual lexical realizations. This suggestion is supported by previous results, which revealed that lexical decisions for low-frequency ambiguous consonant strings are faster than for low-frequency unambiguous consonant strings (Bentin et al., 1984).

Word frequency affected lexical decision performance in both the optional and the obligatory conditions, suggesting that subjects did not ignore the differences between the words even if such a detailed analysis was not absolutely necessary for lexical access. Apparently, this result suggests that when the element of ambiguity is eliminated (even by adding unfamiliar vowels), lexical decisions are based on a full analysis of the graphemic and the phonemic codes. However, the enhanced effect of word frequency in the obligatory condition suggests that such a simple conclusion might be premature. Recall that in the optional condition discrimiantion between words and nonwords could have been accomplished even if the vowels were ignored. Therefore, a relatively reduced effect of frequency may have resulted if some of the decisions were based only on the consonant strings (which were identical in the high- and low-frequency groups), whereas other decisions involved processing of the full phonological code. However, an alternative explanation is possible. The overall slower RTs in the obligatory than in the optional condition suggest that the word/nonword discrimination was more difficult in the former than in the latter condition. Obviously, the words and the nonwords in the obligatory condition were more alike, reducing the certainty level of the subjects. The uncertainty could have been greater for the low- than for the high-frequency words because the former were subjectively more similar to nonwords. This interpretation is supported by previous studies that revealed that low-frequency word decisions are facilitated more than high-frequency word decisions by the presence of unpronounceable nonwords

(Duchek & Neely, 1984, cited by Balota & Chumbley, 1984; James, 1975). Thus, the interaction between the effect of frequency and the nature of the nonwords may be explained as a postaccess decision factor rather than a different manner of lexical access.

An insight into the origin of the interaction between the nonword type and the word frequency effect can be achieved by comparing the condition effect on the ambiguous words with the effect on the unambiguous words. Recall that, for unambiguous words, the word frequency effect did not interact with the effect of the nonword type. If the difference in the magnitude of the frequency effect in the obligatory and optional conditions was not related to word ambiguity, it should have emerged with unambiguous words as well. Therefore, we conclude that the interaction between the nonword type and the frequency effect was related to the ambiguity factor. Although other explanations are possible, we propose to consider this interaction as corroborative evidence that lexical decisions for ambiguous Hebrew words do not require phonological disambiguation.

Experiment 2

In contrast to lexical decision, naming a phonemically ambiguous string of consonants necessarily requires the selection of only one of the alternative phonological representations. The main purpose of Experiment 2 was to investigate the nature of this selection in an attempt to enhance our understanding of the process of disambiguation of Hebrew unvoweled consonant strings.

Previous studies in Hebrew employed only unambiguous words, that is, consonant strings that represented only one lexical item and could be pronounced correctly in only one manner. It has been reported that naming voweled words was delayed when the vowel marks were incompatible with the correct sound of the word, even if the subjects were instructed to ignore the vowels, but it was equally fast if unvoweled words were compared with correctly voweled words (Navon & Shimron, 1981). In a more recent study, however, vowels were found to speed up naming, although they had no effect on lexical decisions (Koriat, 1984).

Recent data from our laboratory revealed that in contrast to more shallow orthographies (such as Serbo-Croatian and English), in Hebrew, making lexical decisions for unvoweled unambiguous consonant strings was faster than reading the same words aloud (Frost et al., 1987). Furthermore, significant semantic priming was found for naming in Hebrew but not in Serbo-Croatian. These results suggest that although naming voweled Hebrew words can, in principle, be based on phonetic cues generated via a process of grapheme-to-phoneme transformation, naming unvoweled words is always mediated by the lexicon. Therefore, lexical ambiguity would influence naming because it probably requires a choice among several lexical representations. In the present experiment, we compared the naming of ambiguous and unambiguous, voweled and unvoweled stimuli. This comparison, we hoped, would shed additional light on naming words in a deep orthography in general, and on the rules of disambiguation of Hebrew homographs in particular.

An additional aspect of word processing that might be influenced by lexical ambiguity is the word frequency effect on naming. Several studies suggested that frequency effects are smaller in naming than in lexical decision tasks (Andrews, 1982; Balota & Chumbley, 1984; Frederiksen & Kroll, 1976). The same relationship was recently found in Hebrew (Frost et al., 1987). In that study, however, only unambiguous words were employed. Naming ambiguous words, on the other hand, might be affected by frequency both during lexical access, and at postlexical processing states (Balota & Chumbley, 1985; Forster & Bednall, 1976; Simpson, 1981). Therefore, it is possible that naming voweled ambiguous words might be affected more strongly by the relative frequency of their alternative phonological representations than naming voweled high- and low-frequency unambiguous words. A secondary purpose of Experiment 2 was to test this hypothesis.

Method

Subjects. The subjects were 64 undergraduates who were paid for their participation. They were all native speakers of Hebrew, with normal or corrected-to-normal vision. None of them had participated in Experiment 1.

Stimuli and design. The stimuli were identical to those used in Experiment 1. Only List A (voweled high-frequency ambiguous words, low-frequency unambiguous words, and regular nonwords), List B (voweled low-frequency ambiguous words, high-frequency unambiguous words, and regular nonwords), and Lists E and F (unvoweled replicas of Lists A and B, respectively) were used, each presented to a different group of 16 subjects. The assignment of subjects to lists was random.

Procedure. The conditions of Experiment 1 were repeated in this experiment. The subjects were instructed to read aloud as fast as possible words and nonwords that were presented to them on the CRT screen. In the voweled condition (Lists A and B), subjects were told to read the stimuli as voweled. In the unvoweled condition (Lists E and F), subjects were told to read aloud as fast as possible the words and the nonwords that were presented on the screen. Since, in Hebrew, reading without vowels is the rule rather than the exception, no additional instructions were given (or solicited) for reading the words. Subjects were told, however, that there was no correct or incorrect way to read the nonwords, and that they could pronounce them by arbitrary assignment of vowels to the consonants.

Subjects' responses were recorded by a Mura DX-118 microphone, which was connected to a Colbourne Instruments voice-key. RTs were measured in milliseconds from stimulus onset by the computer; responses were recorded on a magnetic tape for offline analysis.

Results

As in Experiment 1, the RTs were averaged across subjects for each word, and across words for each subject. All RTs were normalized by excluding responses that were above or below two standard deviations from the subject's or the word's mean. Less than 1.5% of the RTs were excluded.

The time to initiation of naming ambiguous consonant strings in the high- and low-frequency presentations (Lists A and B, respectively) and in the unvoweled condition (List E)[5] was analyzed by one-way ANOVAs across stimuli and across subjects (see Note 3). In contrast to lexical decision performance, the unvoweled consonant strings were named as fast as the high-frequency voweled alternatives, but both groups were named faster than the low-frequency alternatives (Figure

---

[5] The RTs to ambiguous words in List E and List F were not significantly different (653 ms and 665 ms, respectively). For technical reasons, the RTs to unvoweled strings were collapsed across the different phonological realizations that were indeed produced. However, since the great majority of responses were high-frequency words, we assume that the average results are not biased significantly.

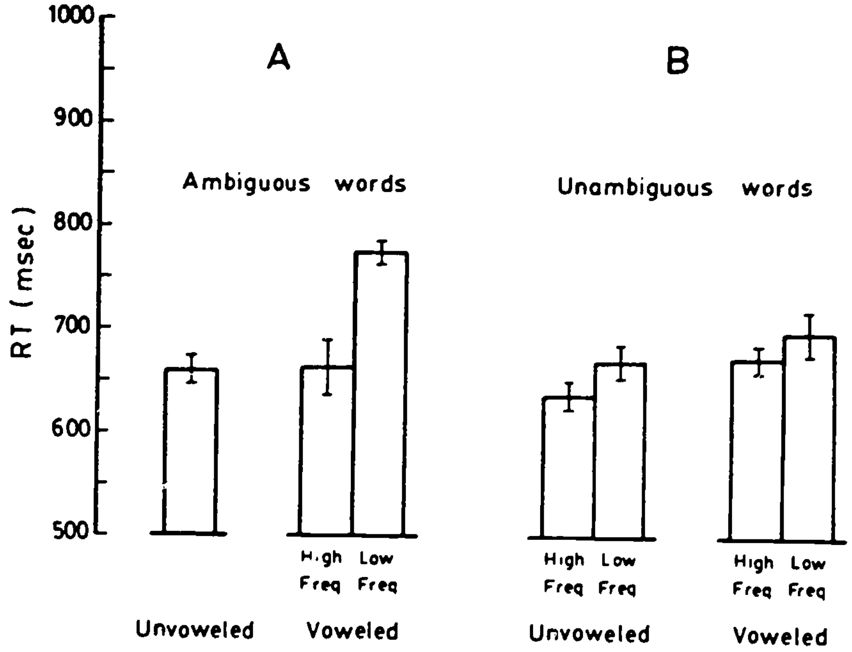


Figure 4. Mean RTs and SEs to voweled and unvoweled words in the naming task: A: Ambiguous words. B: Unambiguous words.

4A). This pattern of performance was supported by the ANOVAs followed by Tukey-A post hoc comparisons: for the stimulus analysis, $F(2,30) = 25.79, MSe = 2,268, p < 0.0001$; for the subject analysis, $F(2,45) = 5.40, MSe = 12,679, p < 0.008$; $minF'(2l,65) = 4.46, p < 0.025$).

The comparison between voweled and unvoweled unambiguous words is presented in Figure 4B. Statistical significance of the differences was assessed by vowels (voweled, unvoweled) X frequency (high, low) ANOVAs across stimuli and across subjects. Unvoweled words were named 30 ms faster than voweled words. This difference was significant for the stimulus analysis, $F(1,30) = 16.47, MSe = 13,718, p < 0.001$, but not for the subject analysis, $F(1,30) < 1$. Consequently, the minF' was not significant. The high-frequency words were named 26 ms faster than the low-frequency words. This difference was not significant either for the subject or for the stimulus analysis. The interaction between frequency and vowel conditions was not significant.

None of the unvoweled words was erroneously read as a nonword, and except four occasional errors, all voweled words were read correctly. (All four errors were low frequency alternatives.) For technical reasons errors made while reading voweled nonwords could not be recovered from the raw data.

The analysis of the words that were actually pronounced by each subject, when unvoweled ambiguous words were presented, revealed that, for 4 of 16 consonant strings, the high-frequency alternative was unanimously chosen by 32 subjects, and for 8 strings the high-frequency alternative was pronounced by more than 75% of the subjects, and in no case was the high-frequency alternative chosen by less than 45% of the subjects. In contrast, seven low-frequency alternatives were not chosen by any subject, and seven were pronounced by less than 10% of the subjects, while none of the low-frequency alternatives was chosen by more than 40% of the subjects.

The nonwords in the voweled condition (List A and List B) and in the unvoweled condition (List E and List F) were compared by a t-test. In contrast to lexical decisions, naming of unvoweled nonwords was significantly faster than naming of voweled nonwords, 757 ms vs. 843 ms; $t(62) = 2.24, p < 0.03$).

Discussion

The naming-time data revealed that the only significant difference between naming voweled and unvoweled words was the slower naming of the voweled low-frequency alternatives of ambiguous words relative to naming of the unvoweled alternatives. In addition, the frequency effect for naming ambiguous voweled words (99 ms) was as large as the frequency effect in the lexical decision task, and considerably larger than that for naming unambiguous voweled words (16 ms). Note that for voweled unambiguous words, the relationship between frequency effects in naming and lexical decision tasks replicates our findings with Hebrew unvoweled unambiguous words, as well as findings from previous studies in English.

These data suggest that vowel marks did not facilitate the naming of printed words. Moreover, vowel marks interfered with pronunciation when they imposed an unexpected interpretation of the grapheme. Therefore, we propose that the subjects might have initially generated a phonological code on the basis of the consonantal information. In the case of ambiguous words, the most frequent alternative was probably activated. A subsequent consideration of the vowel marks had no significant effect on the processing time if they were congruent with the subject's initial response tendency (as was the case with the high-frequency alternatives or with the unambiguous words), but vowel marks required a time-consuming revision of the output pattern if they were incongruent with the initial response. Supporting this hypothesis, all subjects chose the high-frequency alternative in most trials while naming unvoweled consonant strings.

This interpretation assumes that the enhancement of the frequency effect for naming ambiguous words, relative to that for naming unambiguous words, originates from postaccess processing. In agreement with Forster (1981; see also Kinoshita, 1985), we suggest that the delay in naming voweled low-frequency ambiguous words reflects the time spent in evaluating the initially generated high-frequency phonology vis-à-vis the presented vowels and rejecting this alternative in favor of the low-frequency phonology.

General Discussion

In this study, we investigated the process of disambiguation of phonemically and semantically ambiguous Hebrew printed words and the effects of this deep orthography on lexical decision and naming performance. The results suggest that when unvoweled consonant strings are presented, lexical decisions are based primarily on the ambiguous grapheme; lexical disambiguation is achieved in parallel but has little influence on the decision processes per se. In contrast, phonemic (and therefore semantic) disambiguation must precede naming of unvoweled consonant strings, and we suggest that the process of disambiguation is based on a postaccess race of the different phonemic/semantic lexical representations to which the specific consonant string is related. The result of this race (i.e., the word that is pronounced) is determined to a great extent by the relative frequency of the alternative lexical representations. These conclusions are based on the following pattern of observations.

Stimulus Ambiguity and Lexical Decision.

Presentation of the vowel marks in conjunction with the consonant letters disambiguates the printed Hebrew word. Even though word perception should have been facilitated by exclusion of the ambiguity factor, our data revealed that addition of vowel marks significantly *delayed* lexical decisions for ambiguous words. The direct implication of this result is that, in Hebrew, the information provided by the vowels is not absolutely necessary for lexical decisions. A possible explanation is that the lexical decision for unvoweled homographs is normally based on the ambiguous string of letters that may be recognized as a word (or rejected) without access to any of the alternative meanings: phonological disambiguation is, according to this hypothesis, a postaccess process. This hypothesis implies that lexical decisions for unvoweled ambiguous Hebrew words are made without reference to their meaning and, therefore, apparently contradicts the well-established effects of word meaning on lexical decisions (e.g., the semantic priming effect). We can, however, account for this apparent contradiction by assuming, along with Chumbley and Balota (1984), that the effect of word meaning in lexical decision is attributable to a decision stage following lexical access. We will elaborate our view within the framework of a slightly modified version of the two-stage model of lexical decision performance proposed by Balota and Chumbley (1984). Briefly, this model assumes that letter strings (words and nonwords) differ on a familiarity/meaningfulness (FM) dimension. The value of a particular letter string on the FM dimension is determined by its orthographic and phonological similarity to real words. Strings with very high or very low FM values are classified as words and as nonwords, respectively, during a first stage of the decision process. If the computed FM value is not extreme, a second stage, in which a more detailed analysis of the stimulus is accomplished, determines the decision. Obviously, the distribution of the FM values for words and nonwords overlaps. The amount of overlap is related to the discriminability between words and nonwords in the particular stimulus list.

The relative contributions of the phonological/semantic meaningfulness and orthographic familiarity to the computation of the FM value was not specified by the two-stage model. In agreement with Balota and Chumbley (1984), we suggest that the analyses of the orthographic familiarity and of the meaningfulness of the stimulus overlap in time to a great extent. However, the determination of an FM value need not wait until both analyses are exhausted; whenever enough information accumulates, regardless of its source (orthographic or phonologic/semantic), a value is set. The relative contribution of each type of analysis depends on the familiarity of the orthographic cluster and on the availability of its meaning. The orthographic familiarity of homographs is enhanced because the cluster of consonants is encountered in several semantic contexts; on the other hand, their meaning is ambiguous and, therefore, not immediately available. Consequently, we suggest that the FM value for homographs is based mostly on the orthographic, rather than on the semantic/phonological, familiarity. Therefore, whenever the lexical decision is based on the first-stage computation of the FM value, it is done before the phonology and meaning of the homograph are disambiguated. Furthermore, we suggest that this strategy is employed for most unvoweled words, since even for single-meaning unvoweled words, the phonology is not

immediately available in print. Empirical support for this last hypothesis was provided in a previous study (Bentin et al., 1984).[6]

The interference effect of vowel marks on lexical decision suggests that, when vowels are present, subjects cannot ignore them. This hypothesis is strongly supported by the significant frequency effect observed when the voweled high- and low-frequency alternatives of the ambiguous consonant strings were compared. It is possible that the addition of vowels reduced the orthographic familiarity of the stimuli and, at the same time, guided the retrieval of the meaning. Therefore, the computation of FM values was based on both orthographic and phonologic/semantic analysis. Alternatively, the FM value was determined by the orthographic analysis, and the meaning was determined during the second-stage analysis. The second possibility implies that lexical decisions for voweled words require the complete two-stage analysis in many more cases than required for unvoweled words. This hypothesis might also explain the observed difference between voweled and unvoweled words. Note, however, that even if this explanation is true, it does not invalidate our claim that for unvoweled words, and particularly for homographs, the lexical decision is based mostly on the first-stage analysis, which does not include phonological disambiguation.

The only difference between the optional and the obligatory conditions was that the orthographic similarity between words and nonwords was significantly larger in the obligatory condition. If, indeed, the FM value is determined primarily by orthographic familiarity, this manipulation should have shifted the nonword distribution along the FM value to the right, increasing its overlap with the distribution of words. Following the logic underlying Balota and Chumbley's (1984) model, this manipulation should have forced the subjects to increase the upper criterion above which a word is accepted without further analysis. Because most low-frequency words are located below the original criterion, raising this criterion should have affected the high- more than the low-frequency words and, therefore, the net effect should have been an attenuation rather than an amplification of the frequency effect. This trend was indeed observed with unambiguous words; the frequency effects in the obligatory condition (42 ms) were smaller than those in the optional condition (67 ms) (see Table 2). However, the same manipulation with ambiguous words yielded opposite results: the frequency effect in the obligatory condition (201 ms) was twice as large as that in the optional condition (102 ms) (Table 2). One possible explanation of this result is that vowel marks have a larger effect when they indicate a low-frequency rather than a high-frequency alternative of an ambiguous consonant string. Thus, a new aspect of the disambiguation process is disclosed: while processing a letter string, subjects might automatically generate possible lexical representations of the grapheme. When the letter string is voweled, generation of phonology might be initially based only on consonants, independently of the specific vowels employed, because the subjects have little experience with reading vowels, and because the consonants are visually more salient. At some stage, however, the top-down-generated lexical candidate (which provides unequivocal meaning and phonology) is confronted with the bottom-up analysis of the vowels. We suggest that, at this stage, vowels that indicate a high-frequency lexical alternative have a different effect than those that indicate a low-frequency word. We assume that top-down

---

[6] Note, however, that according to our model. stimulus analysis is not terminated by the lexical decision. As we suggested elsewhere (Bentin & Katz, 1984), words are exhaustively analyzed to the "deepest" lexical level, provided that the task does not interfere with this analysis. Thus, even though the lexical decision was made, the phonological disambiguation continues until at least one (and possible all) meanings are accessed.

generation of meanings is influenced by frequency. High-frequency words are more readily available and, therefore, are generated first in a sequential or cascade-type process. Therefore, the lexical candidate that is first confronted with the bottom-up vowel information is a high-frequency alternative. If the vowels indicate a different word, the subject must reject his or her first hypothesis, and generate (or at least consider) another one. This hypothesis, which is not basically different from the postaccess inhibition model suggested for naming by Forster (1981), is also supported by the naming performance in Experiment 2.

Stimulus Ambiguity and Naming

In the naming task, we were able to know which alternative was chosen when an unvoweled ambiguous consonant string was presented. The data revealed that the high-frequency alternatives were indeed chosen by the great majority of subjects. We have no immediate explanation for those few cases in which low-frequency alternatives were selected, but we tend to believe that these selections were caused by coincidental circumstances, such as unusual individual preferences or phonetic priming by the previous random stimulus. At any rate, we consider the analysis of the overt responses as supporting the word-frequency-guided order of meaning generation for ambiguous letter strings.

The effect of vowel marks on naming was very different from their effect on lexical decision. Unvoweled ambiguous words were named significantly faster than the low-frequency alternatives, but were only 16 ms faster than the voweled high-frequency alternatives. Unambiguous words were named 30 ms faster if they were presented without the vowel marks than if they were voweled. This difference was small relative to the inter-subject variability and, therefore, was not significant. However, the direction of the difference conflicts with the results of Koriat (1984). One difference between the two studies is that Koriat employed only unambiguous words. We do not have a simple model to explain how this difference might have affected naming performance, but it seems to us that the reason for the discrepant results should be related to this difference between the stimulus lists.

The main difference between lexical decision and naming unvoweled strings is that naming cannot be performed unless the stimulus is phonologically disambiguated. Because the print does not provide enough phonemic cues, naming requires postaccess processes of disambiguation. Therefore, in contrast to other languages, in Hebrew, whenever lexical decision requires second-stage processes (as, for example, when vowel marks are presented), semantic ambiguity should affect both tasks in a similar way. This assumption is supported by the remarkable similarity of the frequency effect and on absolute naming time and lexical decision time in the optional condition. On the other hand, when lexical decisions can be based on an orthographically generated FM value, naming is relatively delayed.

In conclusion, we suggest that the results of this study revealed that phonological disambiguation of Hebrew unvoweled words does not occur prelexically. Furthermore, at least for Hebrew consonant strings, it appears that lexical decisions are based on the ambiguous orthographic information without reference to meaning or phonological structure. We propose that multiple meanings facilitate lexical decisions by increasing orthographic familiarity and that the decisions are therefore based on this factor alone. These processes are best explained in the context of a multi-stage model of visual word recognition, such as the two-stage model proposed by Balota and Chumbley (1984), with only slight modifications and additions.

It is difficult to comment on the generality of these hypotheses for languages other than Hebrew. Recall, however, that Jastrzembski (1981) reported that among words with an equal number of derivations and an equal number of meanings, those whose meanings tend to be associated with only one derivation were responded to faster. Furthermore, Chumbley and Balota (1984) revealed that lexical decision RTs and RTs in semantic tasks are closely related, independent of other factors. If indeed RTs in semantic association tasks and clustering of meanings around only one etymological derivation are measures of meaning availability, these results suggest that our findings in the deep Hebrew orthography are a rather extreme example of processing ambiguity in printed words.

References

Andrews, S. (1982). Phonological recording: Is the regularity effect consistent? *Memory & Cognition, 10*, 565-575.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 340-357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24*, 89-106.

Bentin, S., Bargai, N., & Katz, L. (1984). Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10*, 353-368.

Bentin, S., & Katz, L. (1984). Semantic awareness in a nonlexical task. *Bulletin of the Psychonomic Society, 22*, 134-139.

Carpenter, P. A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixation. *Journal of Verbal Learning and Verbal Behavior, 20*, 137-161.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition, 12*, 590-606.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33A*, 465-496.

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition, 4*, 53-61.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 361-379.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 104-115.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 130-136.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13*, 278-305.

Kinoshita, S. (1985). Sentence context effects on lexically ambiguous words: Evidence for postaccess inhibition process. *Memory & Cognition, 13*, 579-595.

Koriat, A. (1984). Reading without vowels: Lexical access in Hebrew. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes*. Hillsdale, NJ: Erlbaum.

Kroll, J. F., & Schweickert, J. M. (1978, November). *Syntactic disambiguation of homographs.* Paper presented at the Nineteenth Annual Meeting of the Psychonomic Society, San Antonio, Texas.

Morton, J. (1970). A functional model for memory. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press.

Morton, J. (1979). Word recognition. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics 2: Structure and processes.* Cambridge, MA: MIT Press.

Navon, D., & Shimron, J. (1981). Does word naming involve grapheme-to-phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning and Verbal Behavior, 20*, 97-109.

Navon, D., & Shimron, J. (1984). Reading Hebrew: How necessary is graphemic representation of vowels? In L. Henderson (Ed.), *Orthographies and reading.* Hillsdale, NJ: Erlbaum.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*, 487-494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Homographic entries in the internal lexicon: Effects of systematicity and relative frequencies of meanings. *Journal of Verbal Learning and Verbal Behavior, 10*, 57-62.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20*, 120-136.

Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96*, 316-340.

SOME WORD ORDER EFFECTS IN SERBO-CROAT*

Z. Urošević, Claudia Carello,† M. Savić,** G. Lukatela,** and M. T. Turvey‡**

Abstract. *In the Serbo-Croatian language, the relative order of subject (S), verb (V), and object (O) is flexible. All six of the permutations of those elements have identical words, meaning, and voice, and all six are grammatically acceptable. Nonetheless, SVO is the dominant form. The psychological reality of this dominance was assessed in three tasks. SVO was associated with the shortest latencies (and SO forms in general were faster than OS forms) when subjects were asked to evaluate the plausibility of a sentence (Experiment 1) or to initiate an utterance (Experiment 2). This advantage did not obtain in lexical decision (Experiments 3 and 4). Results were discussed in the context of linguistic universals of word order and Forster's (1979) model of the language processor.*

The relative order of subject, verb, and object is fairly flexible in Serbo-Croat, the major language of Yugoslavia. Of the six possible orders (SVO, SOV, VSO, VOS, OSV, and OVS), all have the same referential meaning and all are considered to be grammatically acceptable (Belić, 1933). Relevant grammatical information is carried by case inflections on nouns rather than word order. For example, in GRANATA POGAĐA PALATU, GRANATA PALATU POGAĐA, etc. ("the grenade strikes the palace"), the nominative (grenade) and accusative (palace) are distinct and the direction of action is clear.

With respect to what we might term word order "preferences" demonstrated by the world's languages, Greenberg (1966) has noted that the vast majority have several variants but one dominant order. Of the six possibilities, however, all are not equally likely to dominate. In fact, orders in which the object precedes the subject (VOS, OSV, and OVS) are quite rare. Of the three so-called common types, SVO is the most frequent and it is this order that dominates in the Serbo-Croatian language (Greenberg, 1966). The issue to be addressed is whether or not the universal rarity of OS constructions relative to SO constructions has a parallel in word order preferences of readers of a language in which both constructions are equally acceptable and equally

---

* *Language and Speech*, 1986, *29*, 177-195.

** University of Belgrade

† State University of New York at Binghamton

‡ Also University of Connecticut

Aknowledgment. This research was supported by NICHD Grants HD-08495 and HD-01994 to the University of Belgrade and Haskins Laboratories, respectively. The authors would like to thank Kenneth Forster and an anonymous reviewer for helpful suggestions.

meaningful. In other words, does the bias against OS orders in languages in general surface in a processing bias against OS orders in Serbo-Croat?

Among Yugoslavian children, at least, the answer appears to be yes. On a task in which 2-4 year olds were required to act out sentences with toy animals, Slobin and Bever (1982) reported that performance on OS sentences was worse than performance on SVO sentences (where 67% was considered above chance, they averaged 63% and 73% correct, respectively). In contrast, young speakers of Turkish, also an inflected language, were equally facile with both orderings (83% and 81%). This suggests that there is not an inherent difficulty of OS sentences. Slobin and Bever maintain that the difference between the two groups lies in the reliability of inflections in the two languages. In Turkish, they are always regular and explicit but in Serbo-Croat there are many irregular declensions and many instances in which cases are indistinct (e.g., the declension for STVAR, "thing," is STVAR, STVARI, STVARI, STVAR, STVAR, STVARI, STVARI for nominative, genitive, dative, accusative, vocative, instrumental, and locative, respectively). In such cases, and in the absence of logical constraints, word order is necessary to understand the sentence. Slobin and Bever argue that the convention is to follow the SVO order in ambiguous sentences and that this tendency seems to spill over to inappropriate circumstances (i.e., when inflections are explicit) resulting in poorer performance by the Yugoslavian children.

Whether or not the canonical form continues to enjoy an advantage in adult linguistic performance is the focus of the present investigation. Broadening the comparison to SO vs. OS sentences will allow us to disentangle word order effects from frequency effects. Advantages for the canonical form might be attributable to the fact that SVO is the most frequently occurring word order. But an advantage of, say, VSO over OSV would provide a stronger case for the influence of word order given that OSV is the next most frequent order in speech after SVO while verb-initial orders are the least frequent.

We will evaluate psychological consequences of varying word order in terms of three questions. First, is the speed of recovery of meaning from a printed sentence influenced by the order in which the words appear? This is addressed with a sentence verification task in Experiment 1. Second, is the speed with which an utterance is assembled affected by the ordering of the words that comprise it? A modified naming task is used to address this question in Experiment 2. And, finally, is the speed with which one lexically evaluates the last word of a sentence affected by altering the order of the two words that precede it? This will be assessed with standard lexical decision tasks in Experiments 3 and 4. One way in which to understand the differences among these tasks is in the context of Forster's (1979, 1981, 1985) characterization of the language processor. Briefly, this model posits that three relatively independent, hierarchically arranged operations are necessary for normal language comprehension: (1) the lexical processor accesses the representations of words in the lexicon; (2) the syntactic processor assigns a grammatical structure to the arrangement of words (stem with attached affixes); and (3) the message processor assigns meaning to the arrangement of words. Each operates automatically, and autonomously, accepting inputs from no source other than the next lowest level. A general problem solver integrates the outputs of these operations.

It can be supposed that in order to accommodate the variety of experimental tasks, the general problem solver can be reorganized into a variety of special purpose devices. That is to say, although the general problems to be solved in normal language are not limited to, say, binary decisions as to lexicality or plausibility, such decisions can be made. Discerning the influences on those special purposes may elucidate the organization that underlies ordinary language comprehension. For example, if a special purpose device is influenced by information that is not directly relevant to its decision, then we might suppose that the processes responsible for evaluating that information are unavoidable, automatic aspects of normal language comprehension that cannot be disengaged simply because a contrived task does not require them. We will discuss each of our tasks, in turn, in light of this strategy.

The sentence verification procedure used in Experiment 1 requires consideration of knowledge not explicitly stated in a given sentence in order to decide whether or not the situation described is semantically plausible (e.g., "father shaves beard" vs. "boy eats concrete"). Insofar as the relevant attributes are available in the lexicon,[1] the message processor could provide the requisite information to a special purpose sentence verification device. In this fashion, plausibility evaluations might be accomplished on the basis of the language processor alone, that is, without reference to general conceptual knowledge.[2] Outputs from the other processors, though not directly relevant to the decision, might be expected to speed or slow the decision as a function of whether or not they are consistent with the output of the message processor.

In assembling an utterance (Experiment 2), readers were forced to consider entire utterances: They were instructed to read the word triads smoothly, as an entity, rather than word by word. The task can be considered as somewhat akin to naming experiments in which subjects simply read a target word. Latencies to initiate pronunciation are slowed for nonwords relative to words (Theios & Muise, 1977) and, at least for English, are hastened by an associatively related prime. But they seem to be unaffected by syntactic relatedness (Seidenberg, Waters, Sanders, & Langer, 1984). A special purpose naming device, therefore, seems able to ignore the signal from the syntactic processor.

In the lexical decision task, positive output from the lexical processor indicates that a lexical entry has been found and so the decision device issues a "yes." Access occurs more quickly

---

[1] Forster (1979) points out that "the tailor made uniforms" can be assessed on the basis of information provided by the lexical entries but "the fireman sprayed the stadium" requires background knowledge (being on fire is not a typical characteristic of a stadium). Kostić (1983) offers verb valence (Starchuk & Chanal, 1963) and animacy mode requirements as lexical markers that could similarly permit the message processor to perform pragmatic evaluations in inflected languages. For example, a verb such as *govoriti* (to speak) requires, by and large, an animate noun in the dative case (GOVORI BRATU, "he speaks to the brother"). Either type of animacy violation (GOVORI BRATA, "he speaks the brother" and GOVORI MAGLI, "he speaks to the fog," respectively) is pragmatically infeasible.

[2] This construal of the message processor should be distinguished from what Fodor (1983) has described as a Quinean process, one that considers information from a variety of sources, using background knowledge not provided by the immediate circumstances. Fodor contrasts this with modular processes that are fast, mandatory, and cannot be influenced by information from sources outside their domain. In Forster's model, however, the message processor is supposed to be a module just as the lexical and syntactic processors are modules; it operates quickly and inexorably. Its domain happens to include information about real world possibility to the extent that some information useful to that designation is to be found in the lexicon.

if a lexical entry of interest has been "primed" by an associatively related one (e.g., Lupker, 1984; Meyer, Schvaneveldt, & Ruddy, 1975). It has also been shown, however, that in this task, which seems to require information from the lexical processor alone, the special-purpose decision making device cannot ignore information from the syntactic and message processors. For example, violations of the prescribed case, gender, or number agreement between context and target pairs in Serbo-Croat are revealed by the syntactic processor (which does so by paying attention to, rather than stripping, inflections), yet they slow lexical decision time relative to grammatically congruent pairs (see Gurjanov, G. Lukatela, K. Lukatela, Savić, & Turvey, 1985, for a summary of such results). Violations of pragmatic event structure have similarly been found to lengthen lexical decision [Kostić, 1983]). Such effects are not on the lexical processor itself but on the post-lexical special purpose decision making device.

The direction to be taken here considers that just as the variety of special purpose devices appear to be differentially sensitive to lexical, syntactic, and semantic aspects of language, so, too, might they be differentially sensitive to word order. While none of them requires an explicit evaluation of word order, verification and reading are somewhat global in requiring subjects to consider all three words at once. Lexical decision, in contrast, is more local because the actual decision is concerned only with the last word. It is expected that the globally oriented special purpose devices—verification and reading—will be more sensitive to word order than the locally focused lexical decision device.

To be redundant, some word order effects are expected even though changing word order does not change semantic or syntactic structure, nor does it entail a change from active to passive voice. Even though the nouns are in the same order as those in the English passive voice sentence "the man(O) was bitten(V) by the dog(S)," "coveka(O) je ujeo(V) pas(S)" is in active voice and means "the dog bit the man" (Javarek & Sudjić, 1963.) Any differences that might be obtained, therefore, can be attributed fairly confidently to the relative efficacy of putatively equivalent word orders.

Experiment 1

In classical sentence verification tasks, a reader is presented with a sentence of the type "a cat is an animal" (vs. "a cat is a mammal") or "all animals are cats" (vs. "some animals are cats") and must decide whether it is true or false. Latency is found to be a function of semantic relatedness (Smith, 1967; Wilkins, 1971), the number of shared features (Ripps, Shoben, & Smith, 1973; Smith, Shoben, & Ripps, 1974) or set overlap of the subject and predicate (Meyer, 1970). More recently, investigators have looked at the implications of propositions that require one to refer to general knowledge—that is, information that is not explicitly stated—in order to verify an assertion (Hayes-Roth & Thorndyke, 1979; Singer, 1981). For example, in the context of "the woman drove downtown to work," "the woman drove a car" is implied (Singer, 1981).

Our first experiment was of this type but did not require participants to remember information from contexts. Simple three-word sentences were semantically evaluated on their own merits. Subjects and objects were chosen to be unrelated associatively or semantically while verbs were chosen to depict a semantically plausible or implausible situation (e.g., "grenade strikes palace" vs. "boy eats concrete"). Plausible sentences were constructed so that referential constraints on a given string of words permitted no ambiguity as to the direction of the action. That is to say, in a sentence such as "grenade strikes palace" distinctive case markings provide redundant information; it is clear that "palace" is not a thing that effects striking. In contrast, for a sentence such as "girl strikes boy," either noun could, in principle, do the striking and case markings would be critical to understanding the sentence. Sentences of the "grenade..." type provide a more conservative test of the relative merits of the various word orders because there is no need (or temptation or opportunity) to resort to the dominant word order as a default interpretation. Word order should be superfluous to verification, which simply requires understanding the possible events into which the constituents could enter. The usual concerns of verification experiments for whether or not words are ambiguous (e.g., Oden & Spira, 1983; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982) or whether the properties with which the sentences are concerned are the dominant properties for those words (Ashcraft, 1976; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974) do not matter here because they are constant across word orders.

The experiment was designed to focus on the contrast between the dominant or canonical form in Serbo-Croat, SVO, and the noncanonical forms. Each of five groups of participants saw SVO type sentences and one other form. It was hypothesized that the semantic relationship implicit in these sentences would be recovered most quickly from the canonical form specifically, and from SO constructions in general.

Method

Subjects. Eighty high school seniors from the Fifth Belgrade Gymnasium served voluntarily as subjects. None had had previous experience with visual processing experiments. There were 16 subjects in each of 5 experimental groups. Each experimental group was further subdivided into 2 counterbalancing groups of 8 subjects each.

Materials. Target sentences were variants of a basic set of 46 semantically plausible and 46 semantically implausible three-word sentences. In order to generate a large enough pool with the appropriate properties (e.g., CVCV... structure, midfrequency range, unambiguous case markings), two word lengths were necessary: Twenty-four sentences of each plausibility set were fashioned from six-letter words and twenty-two of each set comprised words of four letters each. All words were singular in number. Verbs were transitive in the third person present tense. Subjects were nouns in the nominative case while objects were nouns in the accusative case. Nouns were chosen so that these case inflections were distinct. Otherwise, nouns were not restricted as to gender or animacy. The subject and object of a given sentence simply were chosen so as to exhibit no obvious semantic or associative link. Each of the 92 sentences was typed in upper-case Roman letters in each of the 6 word orders to prepare a total of 552 black on white slides. Each letter string was centered on a slide.

Design. The six word-order comparisons were distributed over five experimental groups, with SVO always compared with one other order. Two experimental lists were composed for each counterbalancing group (A and B) in each experimental group (1, 2, 3, 4, and 5). Group 1A saw half of the sentences (11 four-letter plausible, 11 four-letter implausible, 12 six-letter plausible, and 12 six-letter implausible) ordered SVO and the other half ordered SOV. Group 1B saw the same 92 sentences but with SVO and SOV now applied to the counterbalancing halves. This was true of groups 2 (SVO vs. VSO), 3 (SVO vs. OSV), 4 (SVO vs. VOS), and 5 (SVO vs. OVS). In sum, each subject saw the same sentences as every other subject but not necessarily in the same word order; no subject ever experienced the same sentence more than once (not even in a different word order); and each sentence appeared in every word order.

Procedure. On each trial, a single sentence was exposed for 2500 ms (with a 5000 ms intertrial interval) in one channel of a three-channel Scientific Prototype Model GB tachistoscope, illuminated at 10.3 cd/m^2. The subject's task was to decide as rapidly as possible whether or not the sentence described a meaningful, possible event. Subjects were instructed not to assume a poetic or philosophical attitude toward the nature of possibility but to adopt a common sense criterion. Decisions were indicated by depressing a telegraph key with both thumbs for a "No" response or by depressing a slightly further key with both forefingers for a "Yes" response. Latencies were measured from the onset of a slide. The experimental sequence was preceded by a practice sequence of 20 different sentences, which were the same for all experimental groups.

Results and Discussion

Error rates are summarized in Table 1. It can be seen that subjects were in fairly high agreement with the experimenters' categorization of plausible and implausible sentences, particularly when one considers that some proportion of the errors is due to the speeded classification task itself.

Table 1

Proportion of Errors as a Function of Word Order and Plausibility

Word Order	Semantically Plausible	Semantically Implausible
Canonical	.11	.07
Noncanonical	.14	.08

Average verification latencies are shown in Table 2. Because the six-letter word sentences contained 50% more letters than the four-letter word sentences, word length was included as a factor in the analysis to see if this extra load made a difference. A 5 (Group) × 2 (Canonical vs. Noncanonical) × 2 (Semantic Plausibility) × 2 (Word Length) analysis of variance revealed significant main effects of Group, $F(4,75) = 4.88; MSerr = 374956.06$, $p < .005$, Canonical/Noncanonical, $F(1,75) = 28.56, MSerr = 9002.50$, $p < .001$, Plausibility, $F(1,75) = 76.74, MSerr = 49377.39$, $p < .001$, and Word Length, $F(1,75) = 66.22, MSerr = 47547.10$, $p < .001$. Significant interactions were found for Group × Canonical/Noncanonical, $F(4,75) = 4.34$, $MSerr = 9002.50$, $p < .005$, Plausibility × Canonical/Noncanonical, $F(1,75) = 10.18, MSerr = 9016.78, p < .005$, and Plausibility × Word Length, $F(1,75) = 10.61, MSerr = 16998.49$, $p < .005$. All other F values were less than one except Group × Plausibility × Word Length, $F(4,75) = 2.29, MSerr = 16998.49$, which failed to meet significance at the .05 level ($p = .07$).

The effect of Group indicates that the various groups differed in how quickly they responded overall, averaging 1710, 1745, 1788, 1940, and 1977 for SOV, VSO, VOS, OVS, and OSV, respectively. The Canonical/Noncanonical effect reflects a faster response time to the SVO type (1812 ms) than to non-SVO types (1852 ms). Sentences comprising four-letter words were verified more quickly than sentences comprising six-letter words (1762 ms and 1902 ms, respectively). And semantically plausible sentences were accepted more quickly (1755 ms) than semantically implausible sentences were rejected (1909 ms).

Table 2

Average Verfiication Latencies (in ms) for Canonical and Noncanonical Forms as a Function of Word Length and Semantic Plausibility

		Semantically Plausible		Semantically Implausible	
Group	Word Order	4-letter Words	6-letter Words	4-letter Words	6-letter Words
1	SVO	1525	1707	1734	1824
	SVO	1554	1754	1755	1826
2	SVO	1579	1766	1725	1894
	VSO	1642	1766	1715	1869
3	SVO	1704	1907	2008	2091
	SVO	1846	2035	2057	2169
4	SVO	1616	1718	1836	1914
	VOS	1690	1778	1812	1941
5	SVO	1748	1960	1941	2042
	OVS	1777	2028	1972	2052

With respect to the interactions, Word Length × Plausibility indicates that the advantage for sentences of four-letter words over those of six-letter words was more pronounced in semantically plausible (a 174 ms difference) than semantically implausible (106 ms) sentences. (The three-way interaction with Group approached significance because for two of the groups, the size of the advantage for four-letter words did not differ with plausibility.) The advantage of canonical over noncanonical forms was greater in semantically plausible (64 ms) than semantically implausible (16 ms) sentences. Finally, the Group × Canonical/Noncanonical interaction suggests that the degree of difference between SVO and the noncanonical forms varied for the five comparison forms (VSO - SVO = 7 ms, SOV - SVO = 24 ms, OVS - SVO = 34 ms, VOS - SVO = 34 ms, and OSV - SVO = 99 ms).

These last differences provide an indication that among noncanonical forms, SO constructions have an advantage over OS constructions. In order to permit a comparison of all permutations, each order was scaled in units of SVO. In Group 1, for example, a given subject's average response to SOV sentences was divided by that subject's average response to SVO sentences. (It should be noted that, for counterbalancing purposes, half of the subjects saw one set of SVO sentences and half saw a different set. Nonetheless, SVO is a legitimate normalizer since each subject in Group 1 had a counterpart in Groups 2, 3, 4, and 5.) A 4 (Word Order) × 2 (Word Length) × 2 (Semantic Plausibility) analysis of variance was performed on these ratios and their means

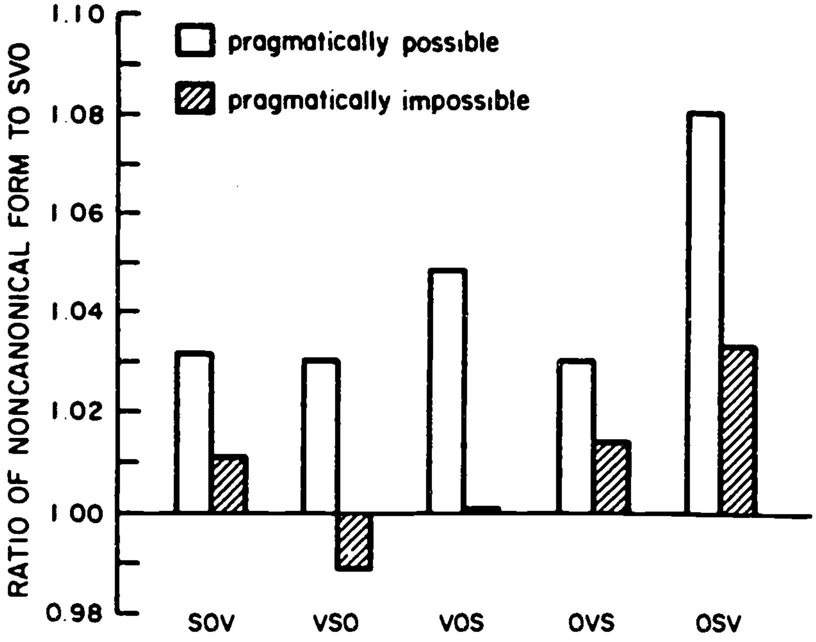


Figure 1. Verification latencies for semantically plausible and implausible sentences scaled in units of the canonical form.

(collapsed over word length) are shown in Figure 1. The main effects of Word Order, $F(4,75) = 3.66, MS = 0.02001$, $p < .01$, and Plausibility, $F(1,75) = 15.48, MS = 0.09474$, $p < .001$, were significant. No other effect or interaction reached significance as all yielded an $F < 1$.

The effect of Semantic Plausibility suggests that, as in the latency analysis, the various orders of implausible sentences were less different from their SVO form than were the plausible sentences, averaging 1.010 and 1.044, respectively. The effect of Word Order again suggests that there were differences among the permutations. To further assess the effect of order with respect to the SO/OS contrast, a second analysis was conducted on the ratios, this time limited to the semantically plausible sentences. This analysis looked at four sentence types: VSO, VOS, SOV, and OSV. The first two contrast SO and OS when the verb is the first word of the sentence, and the second two set up the contrast when the verb is the last word. The SO/OS contrast with V in the middle could not be included since SVO, as the standard, would enter only ones into the analysis. A 2 (Verb Location) × 2 (SO/OS) × 2 (Word Length) analysis of variance revealed one significant effect: SO/OS, $F(1,60) = 4.18, MS = 0.00839$, $p < .04$. That is, SO constructions differ from OS constructions. All other F's were less than 1 except Verb Location, $F(1,60) = 1.14, MS = 0.00839$, $p > .25$.

All analyses showed that word order influences the speed with which speakers/readers of Serbo-Croat can semantically evaluate a three-word sentence. This was true when the canonical SVO form was contrasted with all others as well as in the general contrast between SO and OS forms. Other standard influences on verification—typicality, relatedness, ambiguity—do not differ from one word order to the next. Typicality might be relevant if it is construed as structural typicality, but it should be noted that the SO/OS contrast overrode the frequency differences among the various word orders. That is to say, even though OSV is second to SVO in frequency of occurrence, it was evaluated slowly relative to SO sentences. Similarly, VSO (as a V-initial order) is infrequent but it was evaluated quickly, relative to OS orders. These differences were consistent with the prediction motivated by observations of linguistic universals of word order.

Experiment 2

While ordinary naming tasks seem to be immune to influences of syntactic and message levels of processing, assembling an utterance so that it can be spoken as a unit may be sensitive to such influences insofar as they affect something like prearticulation planning (e.g., Cooper & Ehrlich, 1981) or articulatory routines for stress groups (Sternberg, Monsell, Knoll, & Wright, 1978). For example, one of the uses of word order variants in Serbo-Croat is to emphasize points of interest about the sentence ("the dog (not the child) ate the slipper," "the dog ate (rather than buried) the slipper," "the dog ate the slipper (not its dinner)"). To the extent that different stress patterns require different articulatory planning or work, word order ought to matter. These differences may themselves be linked to the canonical/noncanonical distinction with the dominant form being easier or faster.

A number of investigations of sentence production have hinted at an influence from structural variables. But, as observed by Fodor, Bever, and Garrett (1974), "What is needed is a paradigm in which structural variables are manipulated independent of content variables in a production task" (p. 404). They suggest that the data of Tannenbaum and Williams (1968), who found faster response times for active voice than passive voice descriptions of a line drawing, might be interpreted this way. Their measure was not latency to initiate a response but, rather, complete response time so the finding is inconclusive (e.g., passivation introduces extra syllables). Johnson (1966) reported longer initiation latencies for sentences whose beginning segments had higher depth (cf. Yngve, 1960) than shallower sentences matched for lexical content. These responses were to a paired associate prompt, though; as the deeper sentences were more awkward, they may have been learned less well. More recently, Cooper and Ehrlich (1981) found no effect of early versus late clause boundaries (e.g., "I jog with the pitcher, and the umpire and Pete work out in the gym" vs. "I jog with the pitcher and the umpire, and Pete works out in the gym"). But, in addition to the reliance on memory, these sentences introduced a difference in the meaning of the contrasted sentences.

Experiment 2 does not demand that subjects learn or remember anything. Nor does the structural manipulation change the number of words to be uttered or the voice or the meaning. So, to reiterate, word order change in Serbo-Croat is a clean manipulation of structural properties and its influence is expected to be felt in the globally oriented reading task.

Method

Subjects. Forty-two students from the Faculty of Philosophy at the University of Belgrade participated in the experiment in partial fulfillment of a course requirement. All had participated previously in reaction time experiments. There were 42 subjects in each of three experimental groups. Each experimental group was further subdivided into three counterbalancing groups with 14 subjects each.

Materials. Target sentences were variants of a basic set of 27 three-word sentences, largely the same as the six-letter word plausible sentences of Experiment 1. Again, all words were singular in number, verbs were transitive third person present tense, subjects were nouns in the nominative case, and objects were nouns in the accusative case, gender and animacy were not restricted. Each of the 27 sentences was typed in uppercase Roman in each of the 6 word orders, to prepare a total of 162 black on white slides, and each was centered on a slide. All words were checked in a naming study to ensure that they were not associated with differential pronunciation latencies. Subjects (788 ms), Verbs (787 ms), and Objects (794 ms) did not differ, $F < 1$.

Design. In order to limit the burden on a given subject, the six word order comparisons were distributed over three experimental groups, with SVO always compared with two other orders. Three experimental lists were composed for each counterbalancing group (A, B, and C) in each experimental group (1, 2, and 3). Group 1A saw nine of the sentences as SVO, nine as VSO, and nine as OSV. Groups 1B and 1C saw the same 27 sentences but with the word orders applied to different sets of 9. The same was true of groups 2 (SVO, SOV, and VOS) and 3 (SVO, OVS, and SOV, which was repeated so that the groups would be equated with respect to how many sentences a subject saw). In sum, each subject saw the same sentences as every other subject but not necessarily in the same word order; no subject ever experienced the same sentence more than once (not even in a different word order); and each sentence appeared in every word order.

Procedure. On each trial, a single sentence was exposed for 1500 ms (with a 5000 ms intertrial interval) in one channel of a three-channel Scientific Prototype Model GB tachistoscope, illuminated at 10.3 cd/m^2. The subject's task was to name this sentence aloud as rapidly as possible but necessarily "in one breath." That is, once pronunciation had commenced, it had to proceed without any stammering or stuttering. Naming latency was measured from the onset of the slide by a voice-operated trigger relay constructed by Dr. M. Gurjanov of the Faculty of Electrical Engineering at the University of Belgrade. The experimental procedure was preceded by a practice sequence of 27 different sentences, which were the same for all experimental groups. Considerable feedback was provided on these practice trials to ensure that the utterances were produced smoothly.

Results and Discussion

Each group of comparisons was analyzed separately. Because the means and variances of the treatment levels were related, a square root transformation was performed on the latency data. For Group 1, the SVO/VSO/OSV comparison was significant, $F(2,82) = 11.73, MSerr = 0.597$, $p < .001$. The Group 2 SVO/SOV/VOS comparison was also significant, $F(2,82) = 4.69, Mserr = 0.920$, $p < .01$. The Group 3 comparison of SVO/SOV/OVS was not significant, $F(2,82) = 2.55, Mserr = .638$, $p < .08$. Protected t-tests (Cohen & Cohen, 1975) performed on Groups 1 and 2 revealed that the SVO (897 ms)/VSO (899 ms) contrast was not significant, $t(42) = 0.02$, $p > .50$, but SVO (897 ms)/OSV (941 ms), $t(42) = 4.05$, $p < .001$, SVO (922)/SOV (947 ms), $t(42) = 2.14$, $p < .05$, and SVO (922 ms)/VOS (961 ms), $t(42) = 3.33$, $p < .$[illegible], were significant. For ease of comparison with the verification data presented in Figure 1, each noncanonical order was scaled by the appropriate SVO latency and these ratios are shown in Figure 2. Again, it is apparent that the SO constructions differ less from the canonical form than do the OS constructions.

Word-order influences on initiating an utterance were generally consistent with those from sentence verification. SVO was always the fastest order, significantly different from one SO and two OS orders. Group 3 differences, though not significant, were in the same direction, with SVO

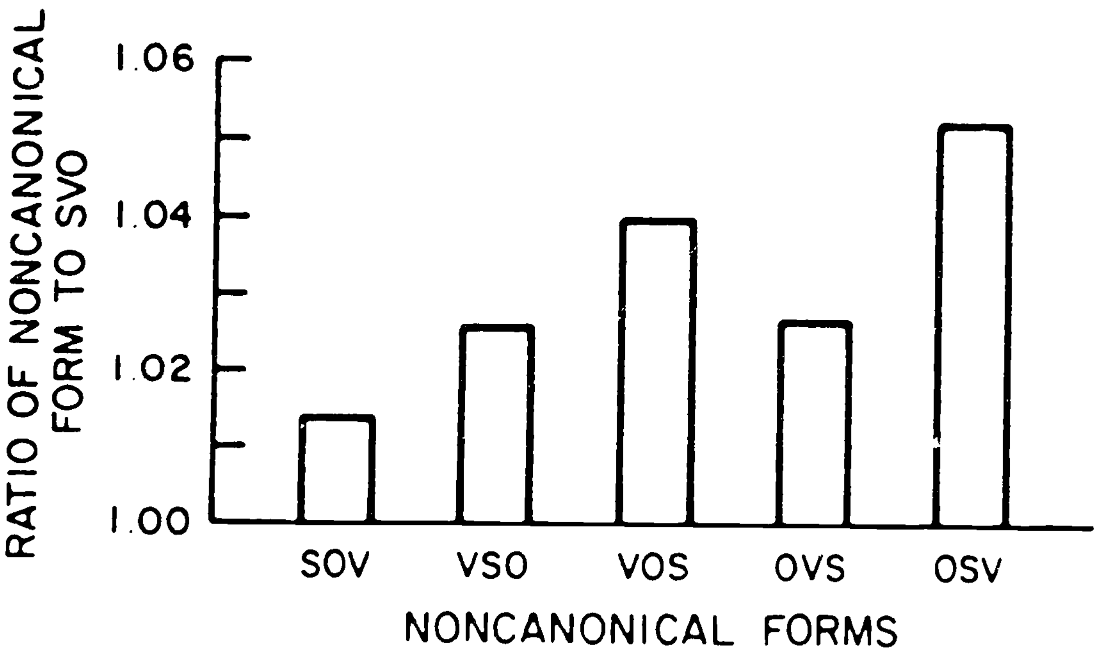


Figure 2. Naming latencies scaled in units of the canonical form.

being 4 ms faster than the SO and 21 ms faster than the OS in that group. Once again, the SO/OS contrast, and not the frequency of occurrence of a given type of order, governed the pattern of results. The reading task is really a hybrid of naming and sentence production tasks. As with naming, it requires no task-specific learning or memory but, as with sentence production, more than one word is uttered on a given trial. A number of manipulations that influence initiation latencies in these two types of tasks, however, do not seem to be a factor here, again because word order changes in Serbo-Croat leave so many properties unchanged. Even articulatory differences do not seem to be straightforward. Although Sternberg and his colleagues (Sternberg et al., 1978; Sternberg, Wright, Knoll, & Monsell, 1980) find that response latency is affected by the number of words with primary stress, this number does not vary with word order. The stress *pattern* changes but, given the finding that latency is not affected by the insertion of words with little or no stress (Sternberg et al., 1978), this should not matter. Moreover, simply changing from a standard or more frequent word order does not, in and of itself, influence response latency. No latency differences were found among utterances of days of the week ordered normally, randomly, or limited to the repetition of one day (Sternberg et al., 1978). Rather, our word order differences seem to reflect something about the structure of sentences.

Experiment 3

We mentioned earlier that lexical decision is hastened by a variety of contexts including associative (relative to neutral or unrelated), syntactic (relative to agrammatic), and semantic (relative to anomalous). An advantage is found for words in sentence contexts over words in isolation (i.e., preceded by a neutral context, Schuberth & Eimas, 1977; West & Stanovich, 1982). It would seem, then, that lexical decision is a particularly sensitive task. Nonetheless, we have already remarked that, in contrast to sentence verification and sentence reading, it is a locally focused task. While this fact has not precluded grammatic and semantic influences on the decision-making device, there are already hints that word order processing does not enjoy the same status. An investigation of the grammatical congruency effect in noun-verb pairs and verb-noun pairs found no difference between the two word orders (Kostić, 1983). Those one word contexts obviously did not explore the salience of SO constructions or the canonical form that provide the focus of Experiments 3 and 4. Experiment 3 was limited to verb targets in SO or OS contexts (which produced semantically plausible sentences) or in isolation (preceded by a row of asterisks). An advantage of meaningful contexts over isolation was expected in SV, VS, or asterisk contexts.

Method

Subjects. Seventy-five high school seniors from the Fifth Belgrade Gymnasium served voluntarily as subjects. None had previous experience with visual processing experiments.

Materials. Target words were 30 third person singular verbs in present tense. The length of target words varied from four to seven letters. Thirty pseudoverbs were generated from the real verbs by changing one letter but maintaining the verb affix. Vowels were substituted for vowels and consonants were substituted for consonants. Twenty SO/OS contexts were constructed as before, with singular subjects in the nominative case and singular objects in the accusative case. Again, subjects and objects bore no obvious associative or semantic link. All sentences with word targets described semantically plausible situations. Word strings with pseudoverb targets were necessarily meaningless.

Design. Each subject saw ten (SO)V, ten (OS)V, and ten (***)V situations with verbs as targets and the same number with pseudoverbs as targets. Whether a given target was seen with an asterisk, SO, or OS context was counterbalanced over subjects. Each subject saw the same sentences (or, for the asterisk context, same targets) as every other subject but not necessarily in the same word order; no subject ever experienced the same sentence more than once although pseudoword targets were repeated with different contexts; and every target appeared with SO, OS, and asterisk contexts. The type of context was randomized over trials.

Procedure. A subject was seated before the CRT of an Apple IIe computer in a dimly lit room. A fixation cross was centered on the screen. On each trial, the fixation point disappeared and a centered context (SO, OS, or asterisks) appeared for 900 ms, followed by an interstimulus interval of 100 ms before the target was presented, also in the center of the screen, for a maximum of 1500 ms. All letter strings appeared in uppercase Roman. Subjects were informed that some contexts would consist of words and others of asterisks. They were instructed to read the contexts where appropriate and, for all cases, decide as rapidly as possible whether or not the target was a word (periodic checks were made to ensure that subjects were reading the contexts). Decisions were again indicated by depressing a telegraph key as in Experiment 1, and latencies were measured from the onset of the target. In the event of an error, a message appeared on the screen and that trial was repeated (but its decision time was discounted). The experimental sequence was preceded by a practice session in which subjects had to achieve an error rate less than 10% over 40 trials.

Results

Minimum and maximum acceptable latencies were set at 400 ms and 1500 ms, respectively. Average lexical decision latencies to word and nonword verb targets in SO, OS, and asterisk contexts are shown in Table 3. Because the nonword targets were filler items and not counterbalanced across groups, the analysis of variance was limited to the word data. The effect of context was significant, $F(2,148) = 16.38, MSerr = 2361.13$, $p < .001$. Significant protected t's were found for the SO/asterisk contrast, $t(150) = 6.18$, $p < .001$, and the OS/asterisk contrast, $t(150) = 7.62$, $p < .001$. Importantly, there was no significant difference between SO and OS, $t(150) = 1.44$, $p > .10$.

Table 3

Average Lexical Decision Latencies (in ms) to Verb and Pseudoverb Targets as a Function of Context

Lexicality	Context		
	(SO)V	(OS)V	(**)V
Verbs	737	728	771
Pseudoverbs	827	845	830

Experiment 4

Subjects. Seventy-five high school seniors from the Fifth Belgrade Gymnasium served voluntarily as subjects. None had previous experience with visual processing experiments.

Materials. Target words were 30 feminine singular nouns in the accusative case. The length of the target words varied from four to eight letters. Thirty pseudonouns were generated from real nouns by changing one or two letters but maintaining the accusative inflection. Twenty SV/VS contexts were constructed as before with singular subjects in the nominative case and singular verbs in the present tense.

Design and procedure. The remainder of the method was the same as Experiment 3 with the appropriate exchanges of O and V.

Results and Discussion

Minimum and maximum acceptable latencies were set at 400 ms and 1500 ms, respectively. Average lexical decision latencies to word and nonword object targets in SV, VS, and asterisk contexts are shown in Table 4. Because the nonword targets were filler items and not counterbalanced across groups, the analysis of variance was limited to the word data. The results were exactly the same as Experiment 3: The effect of context was significant, $F(2,148) = 10.62, MSerr = 1607.36$, $p < .001$, with significant protected ts for the SV/asterisk contrast,

$t(150) = 4.70$, $p < .001$, and the VS/asterisk contrast, $t(150) = 6.26$, $p < .001$, but not the SV/VS contrast, $t(150) = 1.56$, $p > .10$.

Table 4

Average Lexical Decision latencies (in ms) to Noun and Pseudonoun Targets as a Function of Context

Lexicality	Context		
	(SV)O	(SV)O	(**)O
Nouns	682	675	704
Pseudonouns	800	783	791

Both Experiments 3 and 4 replicate the common finding that a sentence context hastens lexical decision relative to a neutral context. This was true for the four word orders that were investigated. Neither experiment found the sentence contexts to differ, however, not even SO and OS. These findings contradict the expectations from the perspective of word order universals but are consistent with the results of Kostić (1983) on grammatical congruency of noun-verb pairs.

General Discussion

Three classes of experiments have been conducted whose focus is the contribution of word order to aspects of linguistic processing of Serbo-Croatian sentences of the subject-verb-object sort. Experiment 1 addressed the comprehension of the events to which the sentences referred: Were they semantically plausible? The experiment sought to determine whether reversals of the universally preferred subject-object ordering (Greenberg, 1966) in general, and departures from the canonical form (SVO) in particular, affected the latency with which evaluations of semantic plausibility could be made. Both the particular and the general perturbations of the canonical form retarded such decisions relative to the canonical SVO form and the SO ordering. The second experiment looked at printed three-word sentences and the time elapsing between their presentation and the initiation of their reading. Similar to the first experiment, it sought to determine the extent to which perturbations of the canonical word order made a difference. The outcome closely parallels that of the first experiment. Orders other than SVO were associated with generally longer latencies and OS orderings were generally slower than SO orderings. The goal of Experiments 3 and 4 was to see whether or not the time needed to decide about a printed word's lexical status was sensitive to the ordering of the subject, verb, and object. Experiment 3 posed the question by investigating lexical decisions on verb targets following either an OS or SO ordering; in Experiment 4 the investigation focused on object targets following either a VS or SV ordering. The outcome of both experiments was the same: With meaning and grammar held constant, a change in word order did not differentially affect lexical decision times although lexical decision times benefited markedly from the sentential contexts.

The weight of the evidence favors the conclusion that word order does have psychological consequences for adult speakers of Serbo-Croat. The conclusion, however, must be qualified by noting that word order effects were not manifest in all three classes of experiments. This "inconsistency" is to be expected. Effects of linguistic variables on linguistic processing depend on the kind of processing under inquiry. This reduces, by and large, to the kind of experimental task used to embody the linguistic process of interest. Before addressing the differences among the three experimental tasks, however, it will serve us well to conjecture about how word order influences arise in the linguistic subsystems proposed in the construal of the language processor due to Forster (1979) but found in closely similar forms elsewhere (e.g., Fodor, 1983).

We remarked in the introduction that, although the various orderings of subject, verb, and object have the same referential meanings, in actual usage the variants may have specifiable contextual meanings related, for example, to intended points of emphasis in the sentence. This suggests the possibility that certain word orders might be perceptually more complex than others. In those cases where SO and OS expressions are, in fact, equivalent, the former may be evaluated faster than the latter because the latter allows or signals the possibility that normal precedence relations may not hold. So, although the resulting representation is the same, the paths taken to reach that result may not be. The syntactic parser may employ two different routines for processing OS and SO sentences: One routine is used for those sentences in which the first noun phrase is marked as subject and another is used for those in which the first noun phrase is marked as object.

Does the failure of the lexical decision task to yield word order effects invalidate this argument? Perhaps not. As we have pointed out, the lexical decision task—unlike the reading and sentence verification tasks—is explicitly aimed at single word rather than sentence-size units. Response times for lexical decision are measured from the onset of the last word, whereas in the other tasks, time to process the entire sentence is included. This allows the possibility that word order controls sentential processing time, but this is not being measured by the lexical decision task. But, then, why were clear sentence context effects found with the lexical decision task? The growing consensus is that this task involves more than the mere checking of a presented letter string against representations in the internal lexicon. The decision part of the task is influenced by other sources of evidential support (e.g., in the form of grammatic and semantic coherency checks [e.g., Gurjanov et al., 1985; Kostić, 1983; West & Stanovich, 1981]). If lexical decision effects are really only decision effects, then responses will be delayed by any property of the target word that does not fit the context and will be hastened by any property that does fit the context. Implausible or ungrammatical combinations, therefore, both increase decision time. Variations in SO order do not affect decision time because, in all cases, the target word fits the context perfectly well. Sentence contexts facilitate lexical decision relative to asterisk contexts because, in the latter, contextual appropriateness is irrelevant.

If a modified lexical decision task requiring the consideration of whole sentences (e.g., the scanning technique advocated by Sanocki et al., 1985) was similarly insensitive to word order influences, then the foregoing line of argument would be supported. If, instead, the whole-sentence technique revealed word order effects, then the presentation technique would be implicated. This last possibility suggests caution in drawing conclusions from research conducted at some remove from the communicative function of language. That is to say, lexical decision is a rather vacuous task relative to normal language comprehension. Nonetheless, providing a context (however meager) exerts an influence on decision time, presumably due to the functional integrity of the language processor. If enriching the lexical decision task even a little seems to reveal organizational properties of the language processor as it normally functions, then we can only speculate that enriching that and other tasks more in the direction of normal language use will reveal a great deal about the functional role of a variety of linguistic structures. Although we have argued that word order effects in Serbo-Croat appear to reflect some seemingly fundamental linguistic universals, this does not preclude the possibility that "preferences" for the canonical SVO form and SO orderings might be overridden in a contextual niche that favors OS orderings. In the present series of experiments, with all things being equal, some word orders were more equal than others. But in normal language use, all things need not be equal and the special fit of a particular linguistic form to a communicative function may, in fact, overcome the SO advantage. It must be remembered, however, that in the experiments reported here performance reflected the SO/OS contrast indifferent to the frequency of use of the various word orders in normal speech. This suggests that even dominance in linguistic usage cannot deny the constraining role of linguistic universals.

References

Ashcroft, M. H. (1976). Priming and property dominance effects in semantic memory. *Memory & Cognition, 4*, 490-500.

Barclay, J. R., Bransford, J. P., Franks, J. J., McCarrell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior, 13*, 471-481.

Belić, A. (1933). *Srpski Jenik Predavanja* (Serbian language lectures). Belgrade.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlational analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cooper, W. E., & Ehrlich, S. F. (1981). Planning speech: Studies in choice reaction time. In T. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 537-552). Hillsdale, NJ: Erlbaum.

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

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language*. New York: McGraw-Hill.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. Walker (Eds.), *Sentence processing* (pp. 27-85). Hillsdale, NJ: Erlbaum.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous processing. *Quarterly Journal of Experimental Psychology, 33A*, 465-495.

Forster, K. I. (1985). Special purpose computation: All is not one. *The Behavioral and Brain Sciences, 8*, 1-42.

Greenberg, J. H. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language* (pp. 73-113). Cambridge, MA: MIT Press.

Gurjanov, M., Lukatela, G., Lukatela, K., Savić, M., and Turvey, M. T. (1985). Grammatical priming of inflected nouns by the gender of possessive adjectives. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

Hayes-Roth, B., & Thorndyke, P. W. (1979). Integration of knowledge from text. *Journal of Verbal Learning and Verbal Behavior, 18*, 91-108.

Javarek, V., & Sudjić, M. (1963). *Serbo-Croat*. London: English University Press.

Johnson, N. F. (1966). On the relationship between sentence structure and latency in generating the sentence. *Journal of Verbal Learning and Verbal Behavior, 5*, 375-380.

Kostić, A. (1983). *Verb valence and lexical decision*. Unpublished doctoral dissertation, University of Connecticut.

Lupker, S. J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behavior, 23*, 709-733.

Meyer, D. E. (1970). On the representation and retrieval of stored semantic information. *Cognitive Psychology, 1*, 242-300.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word-recognition. In P. M. A. Rabbitt (Ed.), *Attention and performance V* (pp. 98-118). London: Academic Press.

Oden, G. C., & Spira, J. L. (1983). Influence of context on the activation and selection of ambiguous word senses. *Quarterly Journal of Experimental Psychology, 35*, 51-64.

Ripps, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12*, 1-20.

Sanocki, T., Goldman, K., Waltz, J., Cook, C., Epstein, W., & Oden, G. C., (1985). Interaction of stimulus and contextual information during reading: Identifying words within sentences. *Memory & Cognition, 13*, 145-157.

Schuberth, R. F., & Eimas, P. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 27-36.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489-537.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Memory & Cognition, 12*, 315-328.

Singer, M. (1981). Verifying the assertions and implications of language. *Journal of Verbal Learning and Verbal Behavior, 20*, 46-60.

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition, 12*, 229-265.

Smith, E. E. (1967). Effects of familiarity on stimulus recognition and categorization. *Journal of Experimental Psychology, 74*, 324-332.

Smith, E. E., Shoben, E. J., & Ripps, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decision. *Psychological Review, 81*, 214-241.

Starchuk, E., & Chanal, G. (1963). *Essentials of scientific Russian*. London: Wesley Publishing Company.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117-152). New York: Academic Press.

Sternberg, S., Wright, C. E., Knoll, R. L., & Monsell, S. (1980). Motor programs in rapid speech: Additional evidence. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 507-534). Hillsdale, NJ: Erlbaum.

Tannenbaum, P. H., & Williams, F. (1968). Generation of active and passive sentences as a function of subject and object focus. *Journal of Verbal Learning and Verbal Behavior, 7*, 246-250.

Theios, J., & Muise, G. (1977). The word identification process in reading. In N. J. Castellan, Jr., D. B. Pisoni, & G. R. Potts (Eds), *Cognitive Theory* (Vol. 2). Hillsdale, NJ: Erlbaum.

West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 5*, 385-399.

Wilkins, A. (1971). Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior, 10*, 382-385.

Yngve, V. M. A. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society, 104*, 444-466.

SHORT-TERM MEMORY, PHONOLOGICAL PROCESSING AND READING ABILITY*

Susan Brady†

Abstract. *Verbal short-term memory deficits are a common characteristic of children with reading problems, and may markedly increase the difficulty of learning to read. Previous work suggests that the basis of the short-term memory deficit may involve limitations in phonetic coding. In the present paper, a series of experiments are reviewed that examined the role of phonological processes in short-term memory. First, a developmental study is described in which a significant relationship was found between phonetic processes and verbal memory span, but not between phonetic processes and nonverbal memory. Second, additional studies are reviewed that collectively found that children with reading problems are less accurate at phonetic encoding than are good readers, and that performance on phonetic processing corresponds with verbal memory span. No reading group differences were obtained on nonverbal perception or memory tasks. These findings suggest that both developmental and individual differences in verbal memory span are related to the efficiency of phonological processes. Practical implications of current cognitive research are discussed.*

In the last 15 years exciting progress has been made in understanding the factors in learning to read. A strong case can now be given for the importance of linguistic skills in reading ability (e.g., Liberman, 1983; Mann, 1984; Perfetti, 1985; Stanovich, 1985; Vellutino, 1979). One of the most robust findings in this area is that there is a relationship between level of reading ability and performance on short-term memory (STM) tasks, with better reading skill being associated with superior STM. This has been observed in adults (Baddeley, Thomson, & Buchanan, 1975; Daneman & Carpenter, 1980), and more extensively in children (see Jorm, 1983, and Stanovich, 1982, for reviews). Deficits in STM for individuals with reading problems have been demonstrated with digit span measures, letter strings, sentence tasks, and recall for pictures of familiar objects. Interestingly, the memory problem for poor readers generally has been found to be specific to linguistic material or to other material that is easy to represent linguistically. When instead the memory task consists of nonsense figures, photographs of strangers, or symbols from an unfamiliar writing system, recall has not been found to be closely related to reading ability (Katz,

---

* *Annals of Dyslexia, 1986, 36*, 138-153.

† Also, University of Rhode Island

Aknowledgment. The research was supported by NIH Grant HD-01994 to Haskins Laboratories. The author gratefully acknowledges the comments of Donald Shankweiler and Anne Fowler on an earlier draft.

Shankweiler, & Liberman, 1981; Liberman, Mann, Shankweiler, & Werfelman, 1973; Vellutino, Pruzek, Steger, & Meshoulam, 1973).

The repeated finding that short-term memory deficits for language co-occur in children with reading problems raises an important point. Rather than poor readers having difficulties restricted to identifying printed words, the STM deficits suggest a broader problem in processing language. In keeping with this, STM limitations have been suggested to contribute to other language impairments of poor readers such as both spoken and written language comprehension (Perfetti & Lesgold, 1977; Shankweiler & Crain, 1986). The presence of language deficits extending beyond failures in word decoding indicates that it is necessary to look for a more general processing problem. With this in mind, our research has not pursued other legitimate questions such as how readers process text (Liberman, 1983) or how the STM deficits might be related to the decoding problem (Perfetti, 1985), but has focused on investigating the origin of the poor readers' underlying deficit in verbal short-term memory.

We have supposed that the basis of the short-term memory deficit in poor readers may involve limitations in phonetic coding. In this paper, a series of experiments will be described in which we examined the phonetic skills of good and poor readers, and investigated the role of phonetic coding in short-term memory. However, first a brief review of the prior evidence linking the poor readers' verbal STM deficit to faulty coding processes is in order.

Background Evidence

To account for the verbal memory deficit, it has been hypothesized that less efficient coding processes in short-term memory may be the basis for some of the language deficiencies characteristic of poor readers. In tests of this hypothesis, the effect of phonological confusability on recall was evaluated by giving good and poor readers strings of rhyming and nonrhyming items. It was reasoned (cf. Baddeley, 1966; Conrad, 1972) that if good readers are better able to form phonetic representations, they should recall more than poor readers on phonetically distinct (nonrhyming) stimuli. However, when the test is saturated with confusing (rhyming) items, the good readers' superior phonetic skills should actually work against them, resulting in more similar performance for the reading groups. This prediction was confirmed in several studies (Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977; Mann, Liberman, & Shankweiler, 1980; Mark, Shankweiler, Liberman, & Fowler, 1977; Olson, Davidson, Kliegl, & Davies, 1984).[1]

One of the most noteworthy experiments, from our present perspective, demonstrated that the reading group differences in performance level and in sensivity to phonetic similarity were virtually identical for visual presentation and for auditory presentation (Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979). Thus it has been suggested that poor readers have a general problem with the use of a phonetic code, independent of how material is presented, not a difficulty that is restricted to the reading process itself.

---

[1] While the hypothesis concerning the phonetic coding ability of good and poor readers has been supported by several lines of evidence (to be discussed in this paper), it is now apparent that the "rhyme effect" per se is subject to age effects (Olson et al., 1984) and task factors (Hall, Wilson, Humphreys, Tinzmann, & Bowyer, 1983).

Analysis of the errors good and poor readers make on short-term memory tasks provides additional evidence that poor readers have less efficient phonetic coding (Brady, Mann, & Schmidt, 1985; Brady, Shankweiler, & Mann, 1983). This interpretation is supported by the observation that both reading groups make errors that indicate use of a phonetic coding strategy. For example, like good readers, poor readers are apt to make transposition errors, recombining the phonetic information from adjacent items in the list, especially when items have phonetic features in common. While the nature of errors suggests that both reading groups are employing a phonetic coding strategy, poor readers' low efficiency in phonetic coding is revealed by a greater frequency of errors.

In sum, past research has shown that poor readers have a specific difficulty with verbal memory tasks; they show reduced sensitivity to rhyme; verbal STM is deficient regardless of whether verbal or auditory presentation is used; and poor readers produce a greater frequency of phonetic errors of transposition. This evidence converges to suggest that poor readers have a language problem in the phonological domain that is not restricted to reading tasks.

Processing in Short-Term Memory

In recent years my colleagues and I have been interested in trying to find out more about the nature of the phonetic coding problem of poor readers. Our aim, as stated earlier, has been to understand better the basis of the short-term memory deficits common for these individuals. If we consider the requirements of an STM task, we can, at least conceptually, consider different processes that are involved, any one of which might be the source of a deficit. For example, in order for something to be remembered, it must first have been perceived. The stimulus, whether it is seen or heard, has to be perceived and represented phonetically. So a deficit might arise from problems in initial phonetic encoding. (In making this assumption, it is perhaps necessary to underscore the evidence that humans have a ubiquitous tendency to use phonetic codes whenever possible. Even when alternative strategies might have seemed likely, such as in reading by congenitally deaf individuals or for readers of a logographic script such as Chinese, evidence of phonetic coding has been obtained [Hanson, Liberman, & Shankweiler, 1984; Tzeng, Hung, & Wang, 1977]). Or, once a phonetic representation is created, it has to be retained, and a deficit might stem from a limitation in rehearsing or retrieving the information in memory.

We have viewed short-term memory, or working memory, as a limited capacity system (Baddeley & Hitch, 1974) that briefly stores information in a phonetic code. Since the amount of information it can handle is limited, the functional capacity of STM can be expected to affect performance on a number of cognitive tasks such as listening comprehension, decoding written language, and reading comprehension. (The role of memory in reading has been discussed in detail elsewhere [e.g., see Liberman et al., 1977; or Perfetti & Lesgold, 1977]). Further, within STM the resources or attention required by aspects of processing (such as encoding, storage, and retrieval) are assumed to become more automatic with experience and to require less resource allocation (LaBerge & Samuels, 1974; Perfetti, 1985). These characteristics, widely supported in the cognitive literature, are taken to reflect normal constraints on human information processing.

Within this framework, we have explored the question of whether the poor reader's phonetic coding deficit in memory might arise from less efficient encoding processes. Part of the impetus for this approach came from experimental evidence with adults that perceptual difficulty can determine memory performance. In a study by Rabbitt (1968), memory span was measured for lists of digits that were presented clearly or degraded by the addition of noise. In addition, the ability of subjects to identify the individual digits was determined in a condition in which the subjects repeated the individual items. Noise levels that caused no effect on perception of the single items significantly impaired recall of strings of the same stimuli. Therefore, adding noise, and increasing the perceptual difficulty, reduced memory span. Rabbitt argued that in a limited capacity system, if perception (or encoding) is difficult, fewer resources will be available for memory.

Research on individual differences in adults has further supported a connection between phonetic processing and memory span. It is known that adults' memory span can be predicted on the basis of the number of words that a subject can read in approximately two seconds (Baddeley et al., 1975). In the same vein, Hoosian (1982) reported a negative correlation between naming speed for digits and digit span. In both studies, phonetic processes involved in encoding and articulating a response were statistically related to memory performance. These results are compatible with the view that perception and memory processes in STM share limited resources.

Developmental Factors in Short-Term Memory

Similar arguments have been made to explain the commonly observed developmental differences in memory span. It is well known that measures of memory span show considerable improvement from early childhood to adulthood. It is estimated that span approximately triples from an average of slightly more than two items at age two to an average of nearly seven items at adulthood (Dempster, 1981). Related to the approach that Rabbitt's work suggests, one hypothesis about the cause of this increase is that, with experience, the component processes in STM come to operate more efficiently, reducing the resource demands and resulting in a developmental increase in functional storage space. [2]

Pertinent to our interests, encoding processes are presumed to increase in efficiency, contributing to a concomitant improvement in memory capacity. These developmental changes in STM are assumed to occur at an automatic, nonstrategic level. We were interested in pursuing this explanation (as will be described shortly) because looking at the basis of developmental improvements offers a way to examine the role of phonetic processes in STM.

Two recent experiments have produced compelling developmental evidence that perceptual processes may be importantly related to memory capacity. Testing children from three to six

---

[2] A second category of explanation of developmental increases in memory focuses on strategic or mnemonic variables that are under conscious direction. These include rehearsal, chunking, organization, and retrieval strategies. According to this view, as children get older they become increasingly aware that particular strategies will aid recall, and increasingly able to employ the appropriate procedure. It is clearly the case that children do improve in the use of these metacognitive skills (Chi, 1976, reviews this evidence). However, it doesn't look as though this is the sole basis for the age differences in span. For example, Samuel (1978) equated children's use of the organization control process by temporally grouping items in recall lists. This procedure improved recall for all the age groups tested (ranging in age from 6 to 29), and thereby preserved the age differences. Similar results have been obtained by Huttenlocher and Burke (1976) and by Engle and Marshall (1983). So strategic variables, while showing increments with age, do not seem to account completely for developmental differences in span.

years of age, all of whom are unlikely to use mnemonic strategies, Case, Kurland, and Goldberg (1982) observed a strong correlation between how fast the children could repeat words and the size of their memory span: the older children, who could repeat faster, remembered more words. The authors argued that the basic encoding and retrieval operations in STM show a developmental increase in "operational efficiency," resulting in more functional storage space. In a convincing test of this, Case and his colleagues equated six-year-olds and adults on speed of counting by manipulating word familiarity (adults were forced to count in an unfamiliar language). Correspondingly, memory span with these stimuli was also demonstrated to be no different for the two age groups. In a second study, it has likewise been reported that although younger children recall less, the same linear function relates speaking rate to short-term memory for subjects ranging in age from four years old to adulthood (Hulme, Thomson, Muir, & Lawrence, 1984). The results of these studies indicate that the buffer component of working memory is phonologically organized, and that developmental increases in memory span are related to the efficiency of phonological processes.

Current Research

Reiterating the main points of the paper thus far, there are three interesting conclusions to be drawn. First, poor readers have a short-term memory deficit that is specific to tasks requiring phonetic coding. Second, adults' performance in verbal short-term memory is related to their perceptual abilities. Third, there is some evidence that developmental increases in STM may be the by-product of increases in the efficiency of phonetic encoding. Influenced by these findings, our experimental work has proceeded in two directions: 1) we focused on the developmental link between phonetic processes and verbal STM to determine more precisely how closely they are connected with one another, and 2) we examined the role of perceptual processes in the short-term memory deficits of poor readers. In the remainder of the paper, the results of these experiments will be summarized.

A. *Development of Phonetic Skills and Memory Span*

The developmental relationship between phonetic processes and memory span was examined in a study with $4\frac{1}{2}$-year-olds, $6\frac{1}{2}$-year-olds, and $8\frac{1}{2}$-year-olds (Merlo & Brady, in preparation). A number of tasks were used to evaluate the correspondence of different aspects of phonetic and memory processes. The phonetic tasks included word repetition with monosyllabic and multisyllabic words, and repeated production of disyllabic tongue twisters (e.g., /sishi/). The responses were scored for reaction time and for the number of correct productions. In this way both speed and accuracy of phonetic processing were measured, allowing separate analyses for these two elements of phonetic efficiency. In the word repetition tasks, response time would reflect both encoding and articulation. We were interested in isolating these components of word repetition speed. To that end, a control task was included in which the children would hear a nonspeech tone to which they were to respond with a prespecified word (either monosyllabic /cat/ or multisyllabic /banana/). In this task, no speech signal has to be encoded; hence it provides an indication of the output or articulatory process.

The goal in this experiment was to examine the developmental link between the phonetic processes of encoding and articulation, and verbal short-term memory capacity. To assess verbal STM, word strings of familiar concrete nouns were given. A nonverbal STM task, the Corsi block test (Corsi, 1972), was included for comparison to help ascertain whether increases in phonetic processes were related to increases in verbal memory. In this test nine identical black blocks are mounted in a scattered arrangement on a wooden surface. The experimenter points, one by one, to a series of blocks. The subject must then repeat this sequential pattern.

All of the measurements on the phonetic and memory tasks showed improvement with age. The linear correlations between verbal STM (VSTM) and each of the measures of phonetic processing skill were all found to be significant: as VSTM increased, verbal reaction and response time decreased, as did the number of errors. It was necessary to check that the variation in age, or some cognitive factor that improves with age, was not a spurious basis for the reported significant relationships among VSTM and the phonetic processing measures. To assess this, a series of correlational analyses were conducted with the effects of age controlled statistically. The resulting partial correlation coefficients showed that when the effects of age were removed, a statistically significant linear relationship continued to exist between VSTM and phonetic processing. The highest correlations were obtained on the accuracy measures for the more demanding tasks: multisyllabic words and tongue twisters. On the speed of processing measures there was an interesting split. The reaction time scores were significantly correlated with verbal memory span after age was partialled out, with two exceptions. These were the control tasks in which the subjects said a word as soon as they heard a tone. This lack of relationship demonstrates that the process tapped by these measures (speaking rate) was not closely related to memory. In any case from the overall results, it seems reasonable to conclude that the ability to process phonetic information efficiently plays at least a partial role in verbal short-term memory functioning.

The importance of this connection between phonetic processing and VSTM is shown by contrasting these results with the correlations with nonverbal memory (the Corsi task). When age was partialled out, verbal and nonverbal memory spans were not significantly correlated, nor was nonverbal memory performance closely related to the phonetic processing measures. Only one variable, monosyllabic reaction time, produced a significant correlation with nonverbal memory (with this number of comparisons that could be a chance result). Certainly the general picture for the relationship of phonetic skills to nonverbal memory was very different.

Demonstration of the close ties between performance on the phonetic tasks and verbal STM span strengthens the evidence that phonetic encoding processes in memory contribute to the basis of individual differences in memory capacity. These results both corroborate and extend previous research findings (Case et al., 1982; Hulme et al., 1984) that showed evidence for developmental improvements in word repetition speed and in verbal STM. In the Merlo and Brady study (in preparation), these findings were confirmed with a variety of phonetic tasks, and there was also evidence for age-related increases in accuracy. Thus older children had longer memory spans and were able to repeat verbal stimuli more accurately and rapidly than younger children. The importance of the relationship between the phonetic variables and memory was underscored by demonstrating that relationship even when age was removed as a factor, and showing no such link with nonverbal memory. Thus when we examine short-term memory developmentally, there is strong evidence that the efficiency of phonetic processes is centrally related to verbal short-term memory span.

B. Phonetic Processes, Memory Span, and Reading Ability

As noted in the section on background evidence, the results of several studies indicate that the problems poor readers demonstrate in verbal STM are related to their use of phonetic coding.

In light of the evidence of the developmental relationship between phonetic processes and verbal memory, we can ask more analytic questions about the phonetic basis of the poor readers' memory deficit. One such question is whether reading group differences in memory span arise from the perceptual requirements of STM: do good and poor readers differ in the efficiency of phonetic encoding? [3]

Effects of noise on phonetic encoding. A few years ago we began to investigate the phonetic skills of children varying in reading performance. In Brady et al. (1983), third-grade children varying in reading ability were tested on a word repetition task consisting of monosyllabic nouns chosen to control for phonetic and syllabic construction and for word frequency. Since the perceptual abilities of poor readers are within normal limits, we anticipated that detecting a phonological deficit in poor readers with a perceptual task might require an especially demanding task. Therefore, the words were recorded in two noise conditions to vary the perceptual difficulty. In one condition, the words were presented clearly; in the second, amplitude-matched noise was added to each stimulus. On the no-noise task, both the good and the poor readers had nearly perfect performance. The addition of noise made it harder for all the subjects to identify the words, but reduced the poor readers' accuracy significantly more than that of the good readers.

The study showed that the perception skills of poor readers are less effective, but that this difference is only observable when the task is difficult. To determine whether poor readers have a general problem with auditory perception, these same subjects were tested on perception of environmental sounds (e.g., human activities [knocking on a door]; animal noises [frogs croaking]; mechanical sounds [a car starting up and driving off]). Once again the stimuli were presented against a background of silence and in (white) noise. As was found with the speech stimuli, performance on environmental sounds was high in the noise-free condition and deteriorated with noise. However, the reading groups did not differ significantly from each other on either condition. If anything, the poor readers identified the sounds-in-noise somewhat more accurately than their better-reading peers. This pattern of results suggests that the difficulty poor readers manifest in perceiving speech-in-noise is not the consequence of generally deficient auditory perceptual ability, but is related specifically to the processing requirements for speech. Thus, paralleling the memory results, poor readers have a problem specific to language that entails phonetic coding.

Effects of phonological difficulty on phonetic encoding. We hypothesized that poor readers may have a problem forming a phonetic representation in short-term memory, whether they are establishing a speech label for a visual stimulus (a letter or a word or a nameable picture), or whether they are listening to a spoken word. Given that however language is perceived it is necessary to encode the material, we conjectured that the deficiency of poor readers in processing phonological structures is a constant feature. Further, this problem may limit the resources in memory and impair recall on memory tasks. We speculated that the poor readers' encoding

---

[3] As in the developmental literature, researchers have also asked whether differences in reading skill might be related to differences in the use of cognitive strategies (i.e., conscious control processes such as rehearsal.) As Stanovich (1985) has noted in a review (see also Jorm, 1983), there is evidence that poor readers are less likely to employ cognitive strategies deliberately that enhance memory performance, and this must be acknowledged as a factor in poor readers' inferior recall. However, there are also studies demonstrating that when good and poor readers are equated on the use of particular strategies, differences in performance level are still present (Cermak, Goldberg, Cermak & Drake, 1980; Cohen & Netley, 1981; Torgesen & Houck, 1980).

problem for speech was present for all language tasks, but was discernible only if the task were sufficiently difficult. To test this line of reasoning, it was necessary to determine whether differences in perception can be demonstrated even under clear listening conditions. In the previous study, with both groups at near ceiling performance levels on the noise-free condition, the task may not have been sensitive enough.

If reading group differences in perceptual processing efficiency are present for clear listening, let's consider how such differences might occur: poor readers might be slower at encoding a phonetic item and/or the quality of the phonetic representation might be less fully accurate. With relatively easy phonetic material and no time limitations, poor readers could possibly perform adequately with either or both of these processing limitations.

With these questions in mind, Brady, Poggie, and Merlo, 1986, examined third-grade good and poor readers on a speech perception task designed to discover whether reading group differences are evident, under some circumstances, when stimuli are presented clearly. Three kinds of stimuli were presented: monosyllabic words, multisyllabic words, and pseudowords. In this way the phonological demands of the task were varied in case monosyllabic words (previously tested) were not sufficiently difficult to process to reveal potential group differences. Therefore the length of the stimuli was increased in the multisyllabic condition, and the familiarity was decreased by using novel constructions (that could occur in English) in the pseudoword condition. Both of these are known to increase processing demands. Reaction time measures were collected to assess processing speed and the responses were scored for accuracy.

When the perceptual skills of good and poor readers were examined, we found that the poor readers were significantly less accurate on the more difficult multisyllabic and pseudoword stimuli. While there were group differences in accuracy, no difference between good and poor readers was obtained for speed of responding. These findings suggest that the critical differences in phonological processing between good and poor readers are related to the accuracy of formulating phonetic representations, not with the rate of processing them. This result complements and clarifies the earlier finding of inferior speech perception by poor readers with speech-in-noise. In sum, it seems when the speech items are somewhat difficult to perceive for any reason, either because of the signal-to-noise ratio, or the phonological complexity or novelty, children with reading problems do significantly less well. Thus when the task is demanding, the lesser ability of poor readers to extract and process the phonetic information from the stimulus items becomes apparent.

Phonetic encoding and STM. Our motive for looking at perceptual skills was to address whether the poor readers have encoding deficits that may contribute to their characteristic limitation in verbal STM. Having demonstrated phonological deficits in poor readers on perceptual tasks, the next step was to see how closely perception and memory performance correspond for children who vary in reading ability. We must return at this point to the developmental study by Merlo and Brady to which we referred earlier. As part of that study, two additional groups of $8\frac{1}{2}$-year-olds were tested: good readers and poor readers. Recall that in that study children were tested on both phonetic and memory tasks. The phonetic tasks were word repetition for monosyllabic and multisyllabic words, repetition of tongue twisters, and words spoken when a nonspeech tone was heard. Speed and accuracy measures were collected. The memory tasks were word strings and the nonverbal Corsi test.

As expected, the reading groups differed significantly on verbal memory, but did not differ on nonverbal STM. On the phonetic tasks, the results replicated the pattern in the study described just prior to this (Brady et al., 1986). Poor readers were significantly less accurate in phonetic processes, and this was again apparent on the more demanding tasks. Once again the reading groups did not differ on speed of processing. Therefore we have a consistent result that poor readers, though not slower, have less accurate phonetic coding, and verbal short-term memory deficits. It is noteworthy that significant correlations were obtained for the reading groups between the accuracy scores on the more demanding measures of phonetic skill and the verbal STM performance: multisyllabic errors and VSTM, r= -.52; tongue twister errors and VSTM, r= - .40. These correlations imply that the reading group differences in verbal short-term memory are associated with differences in phonetic encoding processes common to perception and memory tasks.

Concluding Remarks

To summarize, in this paper we referred to extensive evidence that poor readers have verbal short-term memory deficits. In trying to investigate the basis of the memory deficit, we have adopted a cognitive framework for viewing short-term memory as a limited capacity system. Within that system, if initial encoding is less efficient, the ultimate resources for memory and comprehension will be reduced. In the research presented here, this approach was supported in a developmental study showing a close link between the efficiency of phonetic skills and the available storage capacity in short-term memory. Turning to children with reading problems, we then demonstrated that poor readers are relatively less efficient at phonetic processes (as seen in the accuracy measures), and that this performance corresponds with their reduced memory space (as seen in the correlations between accuracy and memory span).

The identification of less accurate phonetic encoding by chidren who are poor readers provides a parsimonious account of their difficulties on a number of language tasks involving memory and/or perception, including sentence comprehension (see Shankweiler & Crain, 1986, for elaboration of this idea). Further, the present research adds to the burgeoning evidence that poor readers have difficulty in the phonological domain. The problems in phonetic processing, here observed in verbal short-term memory and in speech perception, have also been noted in other language tasks: poor readers are generally lacking in metalinguistic awareness of phonological structure, produce more errors when naming objects, and are less accurate at comprehension of spoken sentences. As Liberman and Shankweiler (1985) discuss in a review of these findings, all of the difficulties appear to stem from deficiencies involving the phonological component of the language. To build on the current knowledge about the origin of reading disability, it would be valuable: 1) to explore the interrelationship of these diverse abilities in normal development and in children encountering difficulty in reading; and 2) to understand better the underlying phonetic and memory processes.

Practical Implications

Practitioners working to alleviate reading problems in children will reasonably be asking if the current research, though short of complete knowledge, has any practical application. We have seen that poor readers are less efficient at phonetic coding and have memory deficits. The developmental literature suggests that with experience encoding operations increase in efficiency and memory span increases. Thus, it may generally be useful to employ more memory games and tasks in school to provide greater practice (cf. Mann & Liberman, 1984). For those individuals with memory deficits, it becomes particularly important in reading acquisition (as Perfetti & Lesgold, 1979, and Perfetti, 1985, have stressed) to overlearn decoding skills to increase encoding speed. In this way the individual can compensate for a further limitation in STM, which is already, by its nature, brief and limited in capacity.

While these recommendations are perhaps not curative, understanding the underlying problems for children with reading difficulty can serve as an important aid in avoiding inappropriate diagnoses and inappropriate treatment. There are some practical implications in this direction from the cognitive research. In closing, these will be briefly described:

1. Evidence is steadily mounting that the problem poor readers have is specific to language. Many prereading programs incorporate practice in labeling environmental sounds. Our research indicates this is a separate, unrelated skill: that children with reading problems are not having a difficulty with nonspeech sound categorization, thus suggesting a teacher would be better off spending time on relevant language tasks such as auditory training on phonological awareness (see Liberman, Shankweiler, Blachman, Camp, & Werfelman, 1980, for examples.)

2. If the errors poor readers made on STM tasks provided evidence that these children were using some other strategy, such as visual or semantic coding, one might conclude they should be taught by a different method. But evidence that poor readers use deviant coding strategies has not been obtained in the research surveyed here. Error analyses confirm that good and poor readers use the same strategies, but poor readers are less accurate.

3. Sometimes the well-intended motive to teach "to the child's strength" goes off-track. Obviously, a practioner needs to deal pragmatically and sensitively with the child's self-esteem and self-confidence, but what has been observed is that the poor reader needs help in phonological processes. The memory research, and that on linguistic awareness, points directly to this. Avoidance of the problem will not improve the essential phonological skills.

4. A similar comment can be made about the recommendation to "teach to the right hemisphere." From what is known about language processing, it is evident that when humans are presented with a letter or word, whether by eye or by ear, a phonetic representation is formed in short-term memory. Neurological studies confirm that this involves regions of the left cerebral hemisphere (Milner, 1974). Going to the whole word method of instruction in no way bypasses linguistic processing: the child is simply left with the more difficult task of figuring out phonic relationships and decoding skills without helpful instruction, but the left-hemispheric linguistic processes involved have not been eliminated.

5. Another lesson from the research on phonological deficits pertains to the issue of "visual learners vs. auditory learners," "visual and auditory dyslexics," etc. On the visual and auditory short-term memory tasks for verbal material, poor readers were found to have a *general* phonetic coding deficit. The modality of presentation did not matter.

Why might the misunderstanding about visual and auditory learning have come about? If the verbal material has been presented visually (either letters, or words, or, maybe especially, pictures), the person testing the subject may focus on the visual nature of the task and interpret a performance deficit as visual in origin. But again, from the studies that have been reviewed, it is known that the child will encode the presented material phonetically. So linguistic processing is central to this visual task. It is important to recognize the language requirements of visual tasks, so as not to draw the wrong conclusion and possibly recommend an inappropriate activity for remediation.

6. In the perception studies, the necessity of having sufficiently sensitive tasks for identifying deficits has been demonstrated. In three different studies, poor readers have been observed to have perceptual problems, but these were only discernible on the more demanding tasks. Many of the standard tests for screening children have rather easy tasks that may demarcate large developmental differences, but not be sensitive to important individual differences. Therefore, it is appropriate to consider whether a task has been sufficiently difficult before concluding that a child has adequate skills in a given area.

References

Baddeley, A. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362-365.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8). New York: Academic Press.

Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589.

Brady, S., Mann, V., & Schmidt, R. (in press). Errors in short term memory for good and poor readers. *Memory & Cognition.*

Brady, S., Poggie, E., & Merlo, M. (1986). An investigation of speech perception abilities in children who differ in reading ability. *Haskins Laboratories Status Report on Speech Research, SR-85*, 23-27.

Brady, S., Shankweiler, D., & Mann. V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35*, 345-367.

Case, R., Kurland, D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386-404.

Cermak, L., Goldberg, J., Cermak, S., & Drake, C. (1980). The short-term memory ability of children with learning disabilities. *Journal of Learning Disabilities, 13*, 25-29.

Chi, M. T. H. (1976). Short-term memory limitations in children: Capacity or processing deficits? *Memory & Cognition, 4*, 559-572.

Cohen, R., & Netley, C. (1981). Short-term memory deficits in reading disabled children in the absence of opportunity for rehearsal strategies. *Intelligence, 5*, 69-76.

Conrad, R. (1972). Speech and reading. In J. F. Kavanaugh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading*. Cambridge, MA: MIT Press.

Corsi, R. M. (1972). *Human memory and the medial temporal region of the brain*. Unpublished doctoral dissertation, McGill University.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*, 63-100.

Engle, R., & Marshall, K. (1983). Do developmental changes in digit span result from acquisition strategies? *Journal of Experimental Child Psychology, 36* 429-436.

Hall, J., Wilson, K., Humphreys, M., Tinzmann, M., & Bowyer, P. (1983). Phonemic-similarity effects in good vs. poor readers. *Memory & Cognition, 11*, 520-527.

Hanson, V. L., Liberman, I. Y., & Shankweiler, D. (1984). Linguistic coding by deaf children in relation to beginning reading success. *Journal of Experimental Child Psychology, 37*, 378-393.

Hoosian, R. (1982). Correlation between pronunciation speed and digit span size. *Perceptual and Motor Skills, 55*, 1128.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38*, 241-253.

Huttenlocher, J., & Burke, D. (1976). Why does memory span increase with age? *Cognitive Psychology, 8*, 1-31.

Jorm, A. (1983). Specific reading retardation and working memory: A review. *British Journal of Psychology, 74*, 311-342.

Katz, R., Shankweiler, D., & Liberman, I. (1981). Memory for item order and phonetic recoding in the beginning reader. *Journal of Experimental Child Psychology, 32*, 474-484.

LaBerge, D., & Samuels, S. (1974). Towards a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Liberman, I. (1983). A language-oriented view of reading and its disabilities. In H. Myklebust (Ed.), *Progress in learning disabilities* (Vol. V). New York: Grune & Stratton.

Liberman, I., Mann, V. A., Shankweiler, D., & Werfelman, M. (1982). Children's memory for recurring linguistic and nonlinguistic material in relation to reading ability. *Cortex, 18*, 367-376.

Liberman, I., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education, 6*, 8-17.

Liberman, I. Y., Shankweiler, D., Blachman, B., Camp, L., & Werfelman, M. (1980). Steps toward literacy. In P. Levinson & C. H. Sloan (Eds.), *Auditory processing and language: Clinical and research perspectives*. New York: Grune & Stratton.

Liberman, I., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. (1977). Phonetic segmentation and recoding in the beginning reader. In A. S. Reber & D. Scarbrough (Eds.), *Toward a psychology of reading: The proceedings of the CUNY conference*. Hillsdale, NJ: Erlbaum.

Mann, V. (1984). Reading skills and language skill. *Developmental Review, 4*, 1-15.

Mann, V. A., & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory: Can they presage early reading problems? *Journal of Learning Disabilities, 17*, 592-599.

Mann, V., Liberman, I., & Shankweiler, D. (1980). Children's memory for sentences and word strings in relation to reading ability. *Memory & Cognition, 8*, 319-335.

Mark, L. S., Shankweiler, D., Liberman, I. Y., & Fowler, C. A. (1977). Phonetic recoding and reading difficulty in beginning readers. *Memory & Cognition, 5*, 623-629.

Merlo, M., & Brady, A. (in preparation). *Reading and short-term memory: The role of phonetic processing.*

Milner, B. (1974). Hemispheric specialization: Scope and limits. In F. O. Schmitt & F. G. Worden (Eds.), *The neurosciences: Third study program*. Cambridge, MA: MIT Press.

Olson, R., Davidson, R., Kliegl, R., & Davies, S. (1984). Development of phonetic memory in disabled and normal readers. *Journal of Experimental Child Psychology, 37*, 187-206.

Perfetti, C. (1985). *Reading ability*. New York: Oxford University Press.

Perfetti, C., & Lesgold, A. (1977). Discourse, comprehension, and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum.

Perfetti, C., & Lesgold, A. (1979). Coding and comprehension in skilled reading and implications for reading instruction. In L. B. Resnick & P. Weaver (Eds.), *Theory and practice in early reading*. Hillsdale, NJ:Erlbaum Associates.

Rabbitt, P. M. A. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology, 20*, 241-248.

Samuel, A. G. (1978). Organizational vs. retrieval factors in the development of digit span. *Journal of Experimental Child Psychology, 26*, 308-319.

Shankweiler, D., & Crain, S. (1986). Language mechanisms and reading disorders: A modular approach. *Cognition, 24*, 139-168.

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 531-545.

Stanovich, K. (1982). Individual differences in the cognitive processes of reading: I. Word decoding. *Journal of Learning Disabilities, 15*, 485-493.

Stanovich, K. (1985). Explaining the variance in reading ability in terms of psychological processes: What have we learned? *Annals of Dyslexia, 35*, 67-96.

Torgesen, J., & Houck, D. (1980). Processing deficiencies of learning-disabled children who perform poorly on the digit span test. *Journal of Educational Psychology, 72*, 141-160.

Tzeng, O. J. L., Hung, D. L., & Wang, W. S. Y. (1977). Speech recoding in reading. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 621-630.

Vellutino, F. (1979). *Dyslexia: Theory and research*. Cambridge, MA: MIT Press.

Vellutino, F., Pruzek, R., Steger, J., & Meshoulam, V. (1973). Immediate visual recall in poor and normal readers as a function of orthographic-linguistic familiarity. *Cortex, 9*, 368-384.

PHASE-LOCKED MODES, PHASE TRANSITIONS, AND COMPONENT OSCILLATORS IN BIOLOGICAL MOTION*

J. A. S. Kelso,** G. Schöner† J. P. Scholz,†† and H. Haken†††

Abstract. *We review the results of joint experimental and theoretical work on coordinated biological motion demonstrating the close alliance between our observations and other nonequilibrium phase transitions in nature (e.g., the presence of critical fluctuations, critical slowing down). Order parameters are empirically determined and their (low-dimensional) dynamics used in order to explain specific pattern formation in movement, including stability and loss of stability leading to behavioral change, phase-locked modes and entrainment. The system's components and their dynamics are identified and it is shown how these may be coupled to produce observed cooperative states. This "phenomenological synergetics" approach is minimalist and operational in strategy, and may be used to understand other systems (e.g., speech), other levels (e.g., neural) and the linkage among levels. It also promotes the search for additional forms of order in multi-component, multi-stable systems.*

1. Introduction

One may well ask: What role might the theoretical and experimental study of coordinated movement play in a conference on the Physics of Structure and Complexity? There are at least two reasons for its inclusion. One is that the movements of animals and people are *ordered spatiotemporal structures* that arise in a system composed of very many neural, muscular and metabolic components that operate on different time scales. The order is such that we are often

---

* Paper presented at Conference on *The Physics of Structure and Complexity*, International Center for Theoretical Physics, Trieste, Italy, September 1-5, 1986. Published in *Physica Scripta* (1987), *35*, 79-87.

** Also Center for Complex Systems, Florida Atlantic University, Boca Raton, FL

† Center for Complex Systems, Florida Atlantic University, Boca Raton, FL

†† Georgia State University

††† Institut für Theoretische Physik, Universität Stuttgart, FRG

Acknowledgment. The work reported here was supported, in large part, by Contract No. N00014-83-K0083 from the U. S. Office of Naval Research (H. Hawkins, Scientific Officer) and by NINCDS Grants NS-13617 and BRS Grant RR-05596. G. Schöner was supported by a Forschungsstipendium of the Deutsche Forschungsgemeinschaft, Bonn, FRG.

able to classify it, like the gaits of a horse, for example, or the limited number of basic sounds (the so-called phonemes), that are common to all languages. Structure, then, emerges from complexity in a fashion reminiscent of the spontaneous formation of structure in open, nonequilibrium systems (e.g., Haken, 1983a; 1983b).

A second reason for allying movement to a physics of complexity is that biological systems are *behaviorally complex.* They are multifunctional in the sense that they are capable of producing a wide variety of behaviors often using the same set of anatomical components (e.g., speaking and chewing). In certain cases, this behavioral complexity may, nevertheless, have a common basis. For example, common to many creatures–vertebrate and invertebrate–is the ability to generate *rhythmical* acts such as walking, flying, and feeding. Since rhythmic behaviors are supported by such a diversity of neural processes, they may be a good starting place to look for laws underlying behavioral complexity.

How then is order in biological coordination to be characterized? Ideally, one would like to have a model system that affords the analysis of pattern and change in pattern, both in terms of experimental data and theoretical tools. Here we describe an ongoing program of research in which theory and experiment have gone (literally), hand in hand, and whose main aims are to understand: The formation of ordered, cooperative states in biological motion; the stability of these observed states; and the conditions that give rise to switching cooperative states. Although our experimental paradigms are concerned with movement control, we believe they might also offer a window into stability and change in general, in a biological system whose real-time behavior can be monitored continuously. We start with some basic facts obtained by ourselves and others. Then we map these observations onto an explicit model that in turn predicts additional aspects that are also studied.

2. Phase Transitions in Biological Movement

Our experiments deal with rhythmical finger (or hand) movements in human subjects. We monitor the kinematic characteristics of these movements using infrared light-emitting diodes attached to the moving parts. The output of these diodes is detected by a *Selspot* optoelectronics camera system. On occasion, we also record from relevant muscles using surface or fine-wire platinum electrodes as the occasion demands (see e.g., Kelso & Scholz, 1985). Thus the behavioral phenomena can be examined at both kinematic and neuromuscular levels. All data are recorded on a 14-channel FM recorder for later off-line digitization at 200 samples/sec., and consequent computer analysis.

Following a paradigm introduced by Kelso (1981, 1984), subjects oscillate their index fingers bilaterally in the transverse plane (i.e., abduction-adduction) in one of two patterns (in-phase or anti-phase). In the former pattern, homologous muscles contract simultaneously; in the latter, the muscles contract in an alternating fashion. Using a pacing metronome, the frequency of oscillation is systematically increased from 1.25 Hz to 3.50 Hz in 0.25 Hz steps every 4 sec (see Kelso & Scholz, 1985; Kelso, Scholz, & Schöner, 1986).

Data from the last 3 sec of each frequency plateau (600 data samples) are used for the calculation of averages to secure stationarity. Figure 1 shows a time series when the system is prepared initially in the anti-phase mode. Obviously, at a certain critical frequency the subject switches spontaneously into the in-phase mode. No such switching occurs when the subject starts

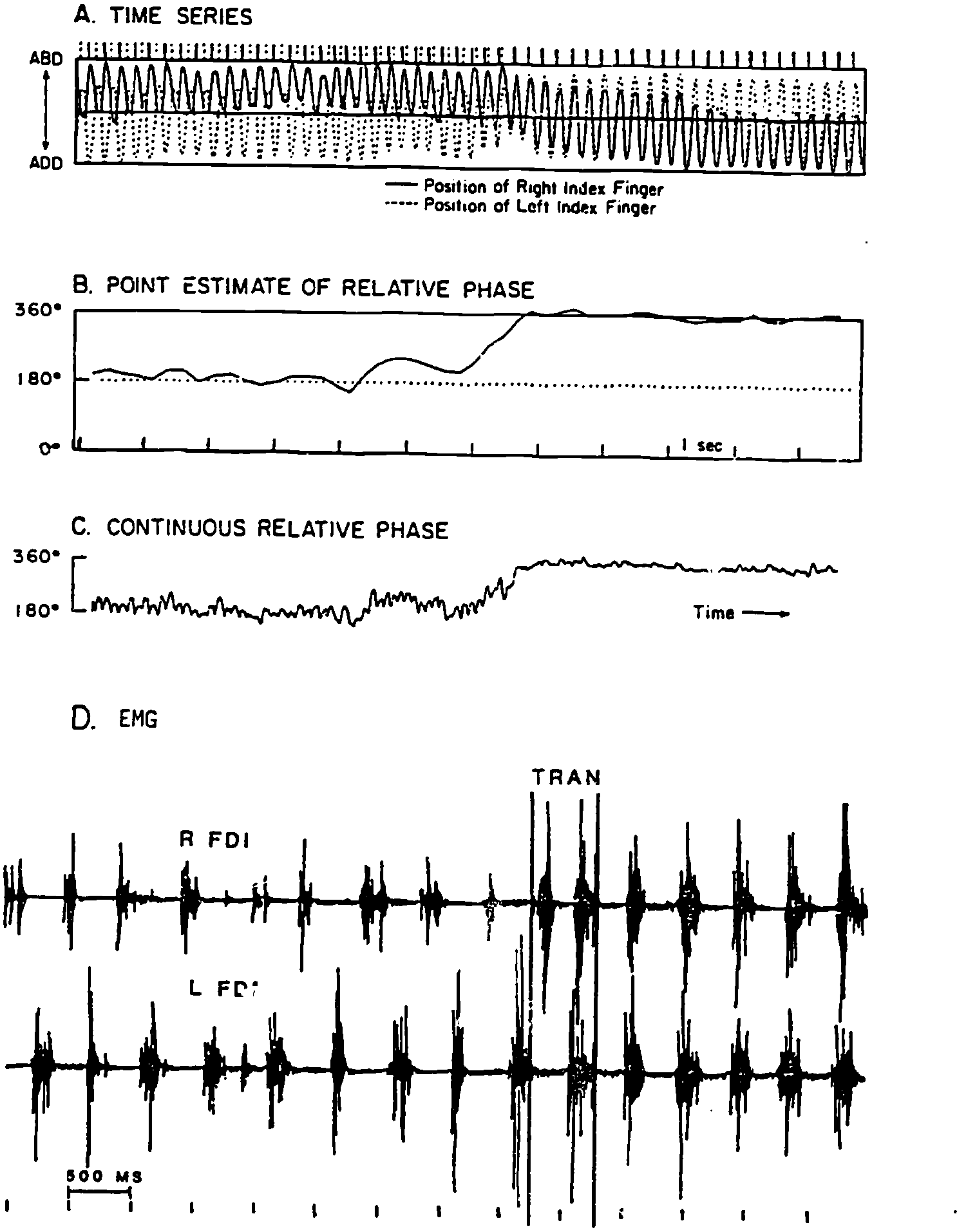


Figure 1. **A.** Time series of left and right finger position; **B.** Point estimate of the relative phase; **C.** Continuous relative phase; **D.** The EMG record of FDI from right and left index finger movements (see text for details).

in the in-phase mode. Thus, while there are two stable patterns for low frequency values, only one pattern remains stable as frequency is scaled beyond a critical region. This transition behavior can be monitored by calculating the relative phase between the two fingers. A *point estimate* of relative phase is the latency of one finger with respect to the other finger's cycle, as determined from peak-to-peak displacement. A *continuous estimate* of relative phase (i.e., at the sampling rate of 200 Hz) can be obtained from the phase plane trajectories of both fingers. (The velocities are obtained by a central difference numerical differentiation procedure). Normalizing the finger oscillations to the unit circle, the phases of the individual fingers can be obtained simply from the arctan ($\dot{x}/x$) if x is normalized finger position (see Kelso, Saltzman, & Tuller, 1986). Relative phase is then just the difference between these individual phases. In Figure 1 the relative phase fluctuates before the transition and stabilizes thereafter (cf. Kelso, 1984; Kelso & Scholz, 1985; Kelso et al., 1986).

3. Dynamical Modeling

In order to understand temporal order and the observed change of such order in terms of dynamics, we need to address the following questions: First, what are the essential variables (order parameters) and how can their dynamics be characterized? Second, what are the 'control' parameters that move the system through its collective states? Third, given a model, what new observations does the model predict? In a first step, relative phase, ϕ may be considered a suitable collective variable that can serve as order parameter.

The reasons are as follows: 1) relative phase, ϕ, characterizes the observed, coordinative modes; 2) ϕ changes abruptly at the transition and is only weakly dependent on parameters outside the transition; 3) ϕ has very simple dynamics in which the ordered phase-locked states are characterized by fixed point attractors. Since the prescribed frequency of oscillation, manipulated during the experiment, is followed very closely, frequency does not appear to be system dependent and can be considered the control parameter.

It is thus possible to determine the dynamics of ϕ from a few basic postulates: 1) The observed stationary states of ϕ at 0 deg and ±180 deg are modeled as point attractors. The dynamics are then assumed to be purely relaxational. This is a minimality strategy in which only the observed attractor type (point attractor) appears in the model. 2) The model must reproduce the observed bifurcation diagram (i.e., bistable in a certain parameter régime, monostable in another parameter régime). 3) Pursuing again a minimality strategy, only the point attractors of the bifurcation diagram should appear. 4) Due to the angular character of ϕ the dynamics have to be 2π-periodic. 5) From the left-right symmetry found in the data, the model is required be symmetric under the transformation $\phi \rightarrow -\phi$. The most general model obeying 1) to 5) is:

$$\dot{\phi} = \frac{-dV(\phi)}{d\phi} \tag{1}$$

where

$$V(\phi) = -a\ cos(\phi) - b\ cos(2\phi) \tag{2}$$

(cf. Haken, Kelso, & Bunz, 1985). This is an explicit model of the dynamics of the relative phase with two parameters, a and b. Haken et al. (1985, Figure 5) show that equations (1) and (2) indeed capture the bifurcation diagram, which has minima at $\phi = 0$ and $\phi = \pm 180$ deg for $a/b < 4$ with the latter minimum turning into a maximum for $a/b > 4$. The order parameter dynamics (1), (2) for ϕ can be derived from nonlinear oscillator equations for the two hands with a nonlinear coupling between them (Haken et al., 1985). We shall postpone discussion of the individual component's dynamics to Section 7.

A chief strategy of the foregoing dynamical analysis is to map the reproducibly observed states of the system onto attractors of a corresponding dynamical model. Thus, *stability* is a central concept, not only as a characterization of the two attractor states, but also because it is *loss of stability* that plays a chief role in effecting the transition. Stability can be measured in several ways. (1) If a small perturbation applied to a system drives it away from its stationary state, the time for the system to return to its stationary state is independent of the size of the perturbation (as long as the latter is sufficiently small). The "local relaxation time", τ_{rel}, (i.e., local with respect to the attractor) is therefore an observable system property that measures the stability of the attractor state. The smaller τ_{rel} is, the more stable is the attractor. The case $\tau_{rel} \rightarrow \infty$ corresponds to a loss of stability.

(2) A second measure of stability is related to noise sources. Any real system described by low dimensional dynamics will be composed of, and be coupled to, many subsystems. These act to a certain degree as *stochastic forces* on the collective variables (cf. Haken, 1983a, Sect. 6.2 and references therein). The presence of stochastic forces and hence of *fluctuations* of the macroscopic variables, is not merely a technical issue, but of both fundamental and practical importance (cf. Haken, 1983a, Sec. 7.3). In the present context, the stochastic forces act as continuously applied perturbations and therefore produce deviations from the attractor state. The size of these fluctuations as measured, for example, by the variance or SD of ϕ around the attractor state, is a metric for the stability of this state. The more stable the attractor, the smaller the mean deviation from the attractor state for a given strength of stochastic force. Let us see how these stability measures behave experimentally.

4.0 Pattern Change: Loss of Stability

Even cursory examination of typical experimental trajectories (see Fig. 1) reveals the presence of fluctuations. One may even suspect a non-trivial dependence on the control parameter frequency in that the relative phase plots look increasingly noisy near the transition (critical fluctuations?) as well as slower in returning from deviations (critical slowing down?). The task, then, is to test these features rigorously.

4.1 Critical Fluctuations: The mean relative phase and its standard deviation were calculated for each frequency plateau (and averaged over ten runs). Figure 2 shows the results for a typical subject.

The mean relative phase stays roughly constant if parameter scaling begins in the in-phase mode, but exhibits a transition if scaling starts in the anti-phase mode. A striking feature is the enhancement of fluctuations (as measured by SD) in the anti-phase mode before and during the transition. In contrast, a roughly constant level of fluctuations is observed in the in-phase mode. We draw the reader's attention especially to the enhanced SD on the pre-transitional plateau, where the system is still stationary (cf. Kelso & Scholz, 1985; Kelso et al., 1986). This provides experimental evidence for the presence of critical fluctuations and a quantitatively consistent fit between data and theory (within the limits of precision in such an approach, see Schöner, Haken, & Kelso, 1986, and Section 5).

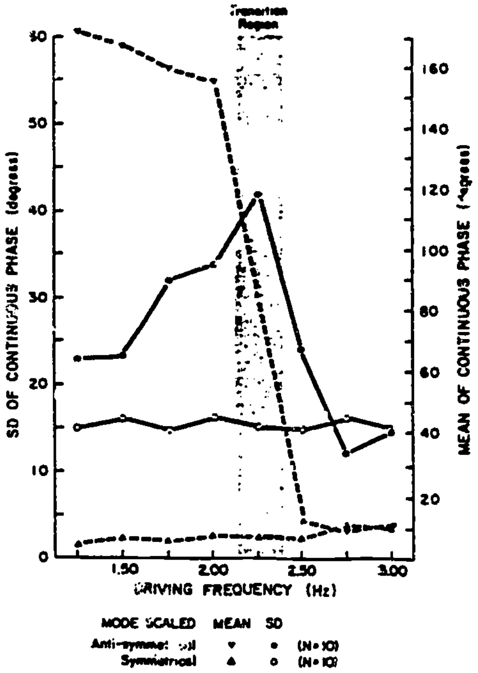


Figure 2. The average mean relative phase modulus from the in-phase (open triangle) and anti-phase (closed triangle) modes of coordination and the average SD (in phase = open circles, anti-phase = closed circles) as a function of driving frequency (in Hz) for a set of 10 experimental runs. On a given run, the mean and SD were calculated for the last 3 sec. (600 samples) at a given frequency (from Kelso, Scholz, Schöner, 1986).

4.2 Critical Slowing Down: Direct evidence of critical slowing down has been obtained from perturbation experiments carried out recently by Scholz (1986) as part of his Ph.D. dissertation (Scholz, Kelso, & Schöner, submitted). The basic experiment involved perturbing one of the index fingers with a torque pulse (50 ms duration) as subjects performed in the two basic coordinative modes (cf. Kelso et al., 1981). The apparatus consisted of a freely rotating support for each index finger that allowed flexion and extension about the metacarpophalangeal joint in the horizontal plane (see Kelso & Holt, 1980, for description.) Electronics provided for the direct transduction of position, velocity, and acceleration, and for the application of a predetermined magnitude of torque to either of the index fingers. In Scholz's experiment, only the right index finger was perturbed at random times during a trial and the torque onset was electronically timed to the peak flexion velocity of that finger. Torque magnitude was set individually for each subject in order to produce readily observable displacement of the finger into extension. Subjects (N=5) were asked to move their fingers rhythmically in one of two modes of coordination: in-phase (relative phase ≈ 0 deg), and anti-phase (relative phase ≈ 180 deg). Scaling trials consisted of increasing the frequency of oscillation in nine 0.2 Hz steps every ten sec starting at 1.0 Hz. When scaling began in the anti-phase mode of coordination, a transition invariably occurred to the in-phase mode.

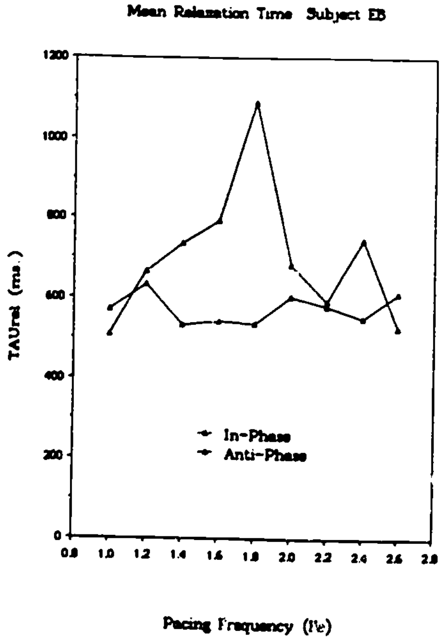


Figure 3. The mean relaxation time (in ms.) as a function of pacing frequency for the two coordinative modes, in-phase (closed triangles) and anti-phase (open triangles). Most of the transitions occur for this subject at 1.8 Hz. Each triangle contains at least 10 observations.

Relaxation time was operationally defined as the time from the offset of the perturbation until the continuous relative phase—calculated from the difference between each hand's phase plane trajectory—returned to its previous steady-state value. Except for the lowest movement frequencies (e.g., 1.0 to 1.2 Hz), relaxation time of the anti-phase mode was significantly longer than that of the in-phase mode for all five subjects. Furthermore, as the critical frequency for mode transition was approached, the relaxation time increased linearly. The result is revealed by a significant positive correlation between the pacing frequency and relaxation time up to the transition ($p < .001$). No such increase occurred when scaling was carried out over the same frequency range beginning in the in-phase mode. Here, either no relationship existed between pacing frequency and relaxation time (N=3/5), or relaxation time actually decreased with increasing frequency of movement ($N = 2/5; p < .001$). The results of one of the subjects are presented in Figure 3.

Additional, though more preliminary evidence for critical slowing down has also been found (see Kelso et al., in press). The method for determining relaxation time in this case uses the power spectrum of the continuous relative phase calculated for the stationary portion of each frequency plateau. Due to the Weiner-Khinchin theorem, this function is just the Fourier transform of the relative phase autocorrelation function. It has a low frequency peak that reflects the relaxational dynamics of relative phase. Stochastic theory (see Section 5) tells us that the line width of this low frequency peak is a measure of relaxation rate (the reciprocal of relaxation time). Figure

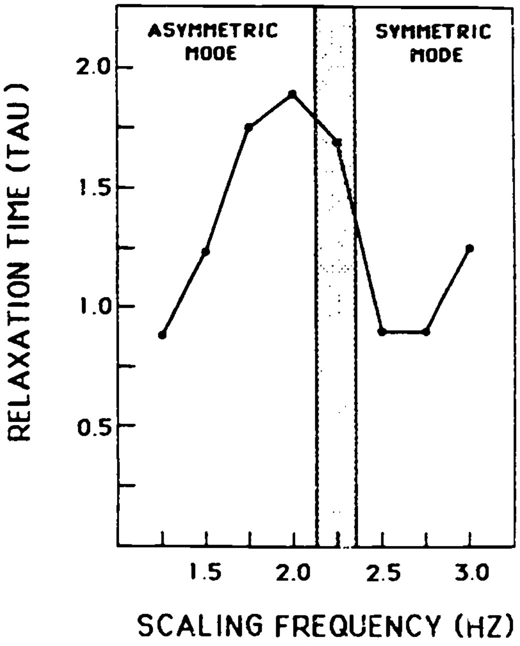


Figure 4. Plot of mean relaxation time, calculated as the inverse of the line width of the power spectra of relative phase at half-power, versus the scaling frequency. One subject's data averaged across 10 trials. Power spectra calculated for 3 sec. of data (600 points) on each trial at each value of the scaling frequency.

4 is a plot of relaxation time, τ_{rel}, as determined from the line width of the power spectrum of relative phase for one subject (the same data as shown in Figure 2).

The strong enhancement of τ_{rel} before the transition can be clearly seen. Nevertheless, a lot of work remains to be done on this topic. An algorithm for determining the line shape of the relative phase spectral density function has to be developed. Detailed modeling is necessary to compare theory and experiment with respect to relaxation time in a quantitative fashion.

5. Stochastic Model of the Phase Transition

With respect to the previously discussed features of the phase transition the model (1), (2) is as yet incomplete because it lacks a representation of fluctuations. Including fluctuations in the model allows us to: 1) test the consistency of the dynamical model experimentally (via time scales relations); 2) determine model parameters quantitatively; and 3) make non-trivial predictions that can be further tested experimentally. The latter point aims at finding lawful relations on the level of dynamics rather than a mere redescription of observations.

To account for fluctuations, and again guided by several reasonable assumptions, we add a stochastic force to equation (1). We assume: 1) the source of the noise consists of many, weakly interacting degrees of freedom (e.g., on the neuromuscular level); 2) the noise sources are correlated over short times compared to the observed macroscopic dynamics (otherwise they are included in the deterministic part of equation (1)); and 3) the noise sources are regular during the transition. These assumptions imply that stochastic forces can be modeled as additive, gaussian, white noise ξ_t:

$$\dot{\phi} = \frac{-dV(\phi)}{d\phi} + \sqrt{Q}\ \xi_t \tag{3}$$

with

$$< \xi_t >= 0; < \xi_t \xi_t{}' >= \delta(t - t') \tag{4}$$

where the parameter Q measures the noise strength (Schöner et al., 1986). This stochastic model is still incomplete, however. In order to solve equ. (3) and compare its solution to experimental data, we have to furnish initial conditions. Finding the appropriate initial conditions requires a discussion of several relevant *time scales* of the present system. Because this is an important point for modeling biological dynamics in many situations, we will be somewhat more general here.

To characterize the state of a biological system within a stochastic dynamical description three types of time scales are relevant. The first is the typical time scale on which the system is *observed*, τ_{obs} (i.e., how long the experimenter observes the system in a given preparation). The second is the previously discussed local relaxation time τ_{rel} (cf. Sect. 3), that is specific to a given attractor. The third one is the so-called *equilibration time* (or global relaxation time), τ_{equ} which is defined as the time it takes the system to achieve the stationary probability distribution from a typical initial distribution. In a bistable situation like ours, below the transition τ_{equ} is determined mostly by the typical time it takes to cross the potential hill (see e.g. Gardiner, 1983).

If these time scales fulfill the following relation

$$\tau_{rel} << \tau_{obs} << \tau_{equ} \tag{5}$$

then the interpretation of observed states as attractor states is consistent. That is, the system has relaxed to a stationary state on the observed time scale, but is not yet distributed over all coexisting attractors according to the stationary probability distribution. When stationary states in an experiment are referred to, what is meant is that the time scales relation (5) is obeyed.

It is important to realize that much of the work in dynamical modeling of biological systems uses deterministic models only and thus implicitly makes the assumption that (5) holds (see, for instance, the contributions to the 1982 Conference on Nonlinearities in Brain Function (Garfinkel & Walter, 1983) for typical examples). To neglect fluctuations and assume (5) throughout is dangerous, however, because (a) the relation (5) breaks down at critical points; (b) fluctuations are an important feature of bifurcation phenomena, and (c) fluctuations are essential in bringing about transitions. Let us examine these three points in more detail.

In our system (3), as the transition is approached (i.e., the anti-phase mode loses its stability), the local relaxation time (with respect to the anti-phase mode) increases, while the global relaxation time decreases (because the potential between 0 and 180 deg vanishes). At the critical point, however, both are of the same order as the observed time and one can see the transition. Thus, at the transition point, the time scales relation (5) is violated and an additional time scale assumes importance, namely the *time scale of parameter change*, τ_p. This reflects the fact that in our system (as often in biological systems) the control parameter that brings about the instability is itself changed in time. The relation of the time scale of parameter change to the other system times plays a decisive role in predicting the nature of the phase transition. If, for example,

$$\tau_{rel} << \tau_{\rho} << \tau_{equ} \tag{6}$$

then the system changes state only as the old state actually becomes unstable. This transition behavior is sometimes referred to as a *second order phase transition* (because of an analogy with equilibrium phase transitions, see Haken, 1983a, Sect. 6.7). In that case features of critical phenomena (such as critical fluctuations or critical slowing down, see below) are predicted. If, on the other hand

$$\tau_{rel} << \tau_{rel} << \tau_{\rho} \tag{7}$$

then the system, with overwhelming probability, always seeks out the lowest potential minimum. It therefore switches state before the old state actually becomes unstable. Jumps and hysteresis (among other features) are generally predicted. This behavior is also called a *first order transition*, again in reference to equilibrium phase transitions [Note: In catastrophe theory, these two different transition behaviors are sometimes referred to as conventions, although they can, of course, be derived from the experimentally accessible relations (6) and (7). It is the failure to treat fluctuations that renders catastrophe theory incomplete in this respect.]

Following these more general remarks, let us return to the concrete stochastic model (3). Because τ_{ρ} and τ_{obs} are of the same order in the experiment, we expect time scales relation (6) to hold up to the transition. At the transition all time scales may then be of the same order. This requires us to differentiate two parameter régimes: (a) The noncritical régime, where the system is stationary in the sense of (5); and (b) The critical régime, where the system exhibits transient behavior. In these régimes one can now solve the stochastic equation (3) via the corresponding Fokker-Planck equation (Schöner et al., 1986). For the noncritical régime stationary probability distributions of local models (that have only one stationary state at either 0 or 180 deg) can be determined. From these the standard deviation (SD) as a measure of the width of the distribution can be calculated. As the transition is approached, the SD of the local model of the anti-phase mode increases, reflecting the enhancement of fluctuations. Using the experimental information on the local relaxation time and the SD in the noncritical régime, one can determine all model parameters a, b, and Q (Schöner et al., 1986). In the critical régime the full Fokker-Planck equation can now be solved numerically, using an appropriate distribution from the pre-transitional régime as an initial condition. Without further adjustable parameters the model accounts nicely for the transient behavior (Schöner et al., 1986).

The stochastic model contains another feature that can be compared experimentally. This is the duration of the transient from the anti-phase state to the in-phase state—which we call the *switching time*. The basic idea is that during the transition the probability density of relative phase—initially concentrated at $\phi \approx \pm 180$ deg flows to $\phi \approx 0$ deg and accumulates there until the "new" peak at $\phi \approx 0$ deg is dominant and stationary. The model predicts the duration of this process both in terms of its distribution and its mean (Schöner et al., 1986). These switching times have been extracted from the experimental data described in Section 4.2 above. In most cases they were easy to calculate as the time between the relative phase value immediately before the transition and the value assumed immediately following the transition. The distribution of switching times for all five subjects is shown in Figure 5. The match between theoretical prediction (cf. Schöner et al., 1986, Fig. 11) and empirical data is impressive, to say the least, even to the

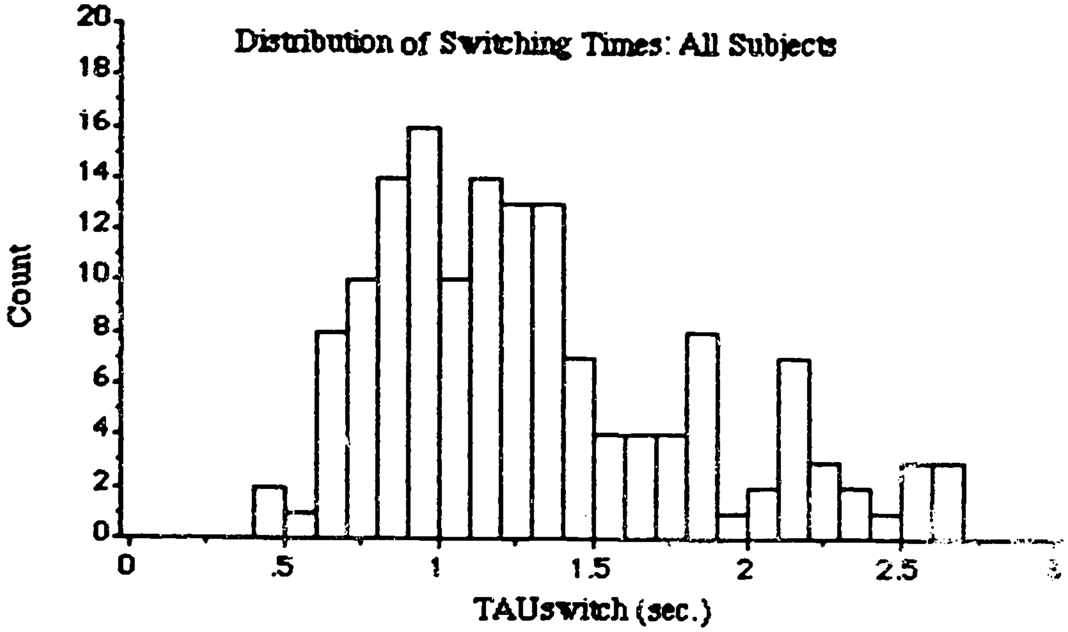


Figure 5. The distribution of switching times for all subjects in Scholz's phase transition experiment (see Text for details).

shapes of the switching time distributions. This new aspect is particularly interesting, because it shows that the switching process itself is quite closely captured by the stochastic dynamics of (3). Using the language of phase transitions may thus be adequate to understand the present phenomenon – even beyond the more superficial level of analogy.

6. Phase-Locked Modes: The "Sea Gull Effect"[1]

In our discussion so far we have always assumed that the system has only two phase-locked patterns—in-phase and anti-phase—at its disposal. Is this assumption really valid? The experiments described in Sections 2 and 4 probe only these two states and their local environment (local in relative phase). We shall now discuss experiments (Tuller & Kelso, 1985; Yamanishi et al., 1980), that allow us to establish the stabilities of all relative phase values, thus affording a view of the whole "potential landscape."

In Tuller and Kelso (1985), the subject's task was simply to tap with the left index finger every time a light for the left hand flashed, and to tap with the right index finger every time a light for the right hand flashed. The set of conditions involved different lag times between onsets of the two lights varying in 100 ms steps from synchrony to a 500 ms lag (or .5 out-of-phase) and back to synchrony. The cycle time for the lights was constant at 1 sec. The phases of the lights did not change within a trial. Four 24-sec trials of each of ten phase conditions were presented randomly.

---

[1] The reason we dub this the "sea gull effect" is obvious from the shape of the function shown in Figure 6 (bottom). Perhaps our proximity to the Atlantic Ocean played a role in naming it the way we did (for another example, see Yamanishi et al., 1980, Figure 2).

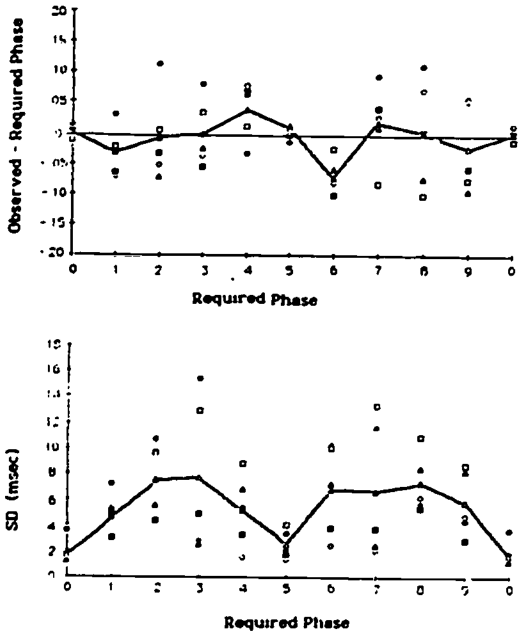


Figure 6. Top: The phase of the pacing lights specifying the required phase plotted against the mean difference between the phase required and the phase actually produced. A negative number means that the required phase was underestimated. Each symbol represents an individual subject's mean computed over $N > 80$ movements. **Bottom:** Standard deviation of the phase produced plotted against the required phase. Symbols same as above.

The top portion of Figure 6 shows the mean deviation from required relative phase as a function of the required phase difference. The bottom portion of Figure 6 shows the standard deviation (SD) of the observed relative phase between the hands as a function of the required relative phase. The different symbols refer to different subjects and the open triangles connected by straight lines are the means across subjects. Obviously in-phase (at zero) and anti-phase (at 0.5) movements are the most stably produced. [This is the case in both musicians and non-musicians as well as in split-brain patients (Tuller & Kelso, 1985)]. Moreover, the top portion of Figure 6 shows how these two states attract neighboring states—the difference between observed and required phase passes through 0.0 and 0.5 with a negative slope. These findings are highly consistent with our basic modeling assumption, namely, that in-phase and anti-phase are the two basic stable phase-locked patterns. To make the implications of this experiment more stringent, however, we have to generalize our model to include the externally imposed required phase.

A simple way to include the external pacing in the equation (3) for relative phase is to alter the potential. The potential (2) represents the system's intrinsic cooperativity. We assume that this remains valid under the paced conditions. Thus, we represent the required phase by adding a term to the potential that attracts the 'intrinsic' relative phase toward the required phase. The simplest function that does this (while conforming to certain periodicity requirements, see Schöner, Kelso, and Tuller, forthcoming, for details) is $\cos\left(\frac{\phi-\psi}{2}\right)$, where ψ is the required phase.

The new potential thus reads:

$$V\psi(\phi) = -a\cos\phi - b\cos(2\phi) - c\cos\left(\frac{\phi-\psi}{2}\right) \qquad (8)$$

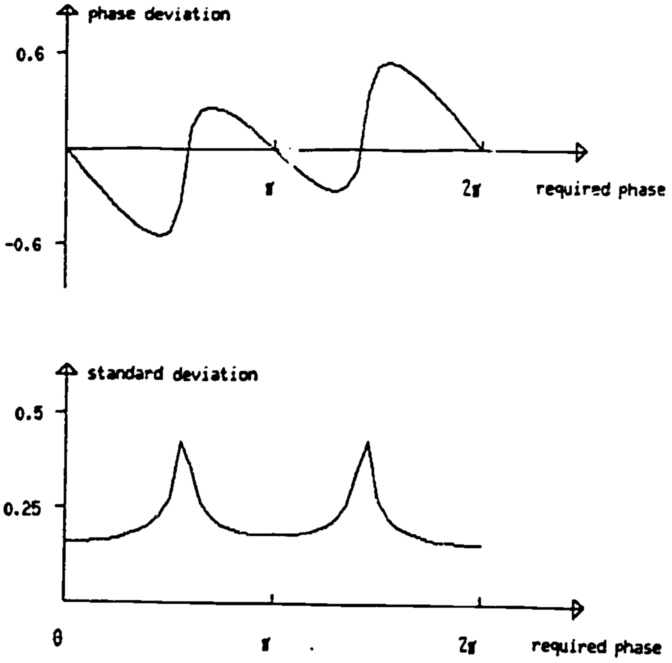


Figure 7. **Top:** The deviation of relative phase from required phase as a function of required phase as calculated from the minima of V_{ψ} (equ.(8)). The parameters were $a = 1.Hz, b = 1.Hz, c = 20Hz$. **Bottom:** The standard deviation as determined from a local model around the minima of V_{ψ}. The parameters a, b, c were as above and $Q = 0.5Hz$.

(Note that the new term breaks the $\phi \to -\phi$ symmetry as the pacing does in the experiment). The zero of this potential are the new stationary states. [Unfortunately the corresponding transcendental equation cannot be solved analytically. We determined the stationary states numerically, using for parameters a and b values similar to those previously established to account for experimental data (Schöner et al., 1986), and choosing c sufficiently large to see an effect of the new term. No systematic attempt to optimize parameters has yet been made, however.] In the top portion of Figure 7 the deviation of the stationary solution for relative phase from the required phase ($\phi_{stat} - \psi$) is plotted as a function of required phase, ψ. Obviously our model captures the attractivity of the two basic modes. To determine the stochastic properties we proceed as in the previously discussed case of noncritical properties for the model (3) (cf. section 5; Schöner et al., 1986): Expanding the potential about its stationary solution and determining the stationary probability distribution of the resulting local model allows us to calculate the SD in this approximation. The bottom portion of Figure 7 shows the SD as a function of required phase ψ (where a, b, and c are as in the top portion and Q is chosen as in the model for the phase transition (Schöner et al., 1986)). Obviously this function captures the qualitative features of the experimental data (although it is a somewhat edgy "sea gull"!). Our modeling here acquires additional credibility through the fact that the potential (8) can again be derived from a model of the oscillatory components of the system, in which the external pacing has been incorporated. This is briefly discussed in the following section. A more detailed analysis of these models will be published elsewhere (Schöner, Kelso, & Tuller, forthcoming). In summary, the global stability measurement performed in this experiment together with a theoretical model shows the consistency of our conceptual approach. We consider this continued close match of theory and experiment to be quite remarkable.

7. Modeling the Subsystems

A key feature of the approach thus far has been to characterize coordinated states entirely in terms of the dynamics of macroscopic, collective variables (in this case relative phase is an order parameter). Here we address the nature of the subsystems themselves and how these can be coupled so as to produce coordinated states. We start at the next level down, as it were, which is the individual hands themselves. Experimentally, the behavior of the individual hands is observed as finger position, x, and velocity, $\dot{x}$. The stable and reproducible oscillatory performance of each hand is modeled as an attractor in the phase plane $(x, \dot{x})$, in this case a limit cycle. Several experimental features constrain the modeling. Kinematic relationships, such as those between amplitude, frequency, and peak velocity, have been measured (Kay et al., 1987). Figure 8 (from Kay et al., 1987) shows the amplitude-frequency relation for oscillatory movements of only one hand. The observed monotonic decrease of amplitude with frequency can be modeled by a combination of the well-known van der Pol and Rayleigh oscillators (Haken et al., 1985; Kay et al., 1987):

$$\ddot{x} + f(x, \dot{x}) = 0 \tag{9}$$

with

$$f(x, \dot{x}) = \alpha\dot{x} + \beta\dot{x}^3 + \gamma\dot{x}x^2 + \omega^2 x \tag{10}$$

In mapping the observed oscillatory state onto a limit cycle the notion of stability is, once again, a key feature of our theory. This can again be tested by measuring the relaxation time after a perturbation of the hand in a fashion similar to that described in Section 4. Such experiments have been recently performed by Kay (1986) under the direction of the first author. Along with the observed kinematic relations, relaxation time measures allow one to determine all parameters in (9), (10).

Another assumption implicit in (9), (10) is that the oscillation is essentially autonomous. This can be experimentally tested in the perturbation paradigm by phase resetting techniques (see e.g., Winfree, 1980). The basic idea is that the phase of an autonomous oscillator is marginally stable, unlike that of a driven oscillator, which is locked to the driving function. This assumption has also been checked in Kay's experiments. Preliminary results show that the oscillation is indeed autonomous (in the absence of external pacing).

How can the components with their dynamics (9) and (10) give rise to the phase-locked coordinative modes? Obviously their dynamics have to be coupled. Haken et al. (1985) have determined coupling structures that can account for the observed phase-lockings. The simplest model that achieves this is a van-der-Pol-like coupling of the form:

$$\ddot{x} + f(x_1, \dot{x}_1) = (\dot{x}_1 - \dot{x}_2)\{A + B(x_1 - x_2)^2\} \tag{11}$$

$$\ddot{x}_2 + f(x_2, \dot{x}_2) = (\dot{x}_2 - \dot{x}_1)\{A + B(x_2 - x_1)^2\} \tag{12}$$

where f is the oscillator function (10) and A and B are coupling constants. The experimental observation, that the kinematic relations (e.g., amplitude-frequency relation) are not significantly different between the coordinative modes and the single hand movements, shows that the coupling constants A, B are small compared to the corresponding coefficients α, γ of the oscillator function (10) [Kay et al., 1987]. In spite of this the coupling structure (11), (12) gives rise to the two phase-locked states. Indeed Haken et al. (1985) were able to derive the equation for relative phase (1), (2) from (11), (12) using the slowly varying amplitude and rotating wave approximations. These

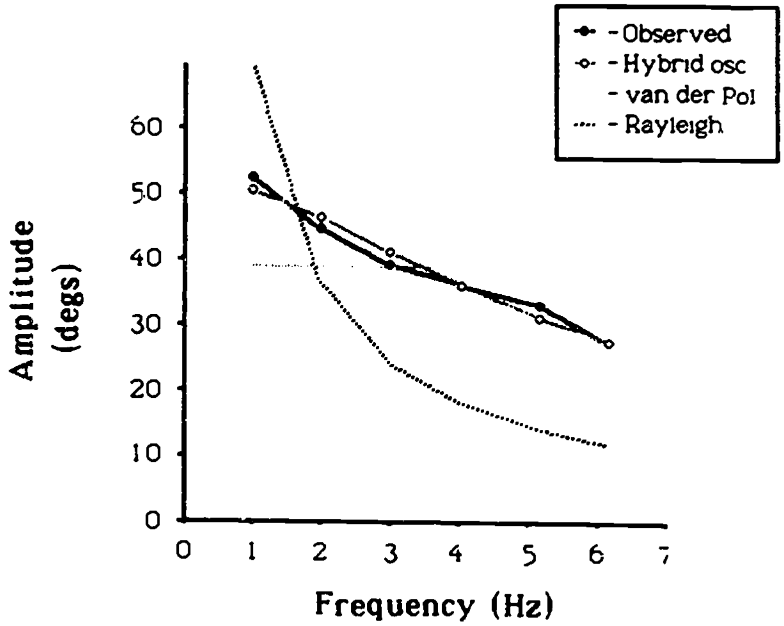


Figure 8. Amplitude versus frequency for rhythmical movements of a single hand. The experimental points (full circles) are means over subjects, experimental sessions and trials. The (hybrid) oscillator of equ. (9) (10) was fitted to the data (least squares). For illustration purposes fits of the van der Pol and the Rayleigh oscillators alone are also shown.

results not only provide further support for the dynamical model on the collective variable level, but also establish in a rigorous fashion the relation of the two levels of description.

Finally, we indicate briefly how the pacing of both hands in the "sea-gull effect" (Section 6) may be incorporated into the model at the component level. The basic idea is similar to that used to determine the potential (8): We assume that the system's intrinsic dynamics are still intact and the pacing acts as an additional external force. For the oscillator equations this can be done by adding a periodic driving force to their oscillator function, for example,

$$\ddot{x}_1 + f(x_1,\dot{x}_1) = (\dot{x}_1 - \dot{x}_2)\{A + B(x_1 - x_2)^2\} + F cos(\omega t) \tag{13}$$

$$\ddot{x} + f(x_2,\dot{x}_2) = (\dot{x}_2 - \dot{x}_1)\{A + B(x_2 - x_1)^2\} + F cos(\omega t + \psi) \tag{14}$$

Here F is the coupling constant of the driving force and ψ the required relative phase. For convenience we have chosen the natural frequency of the oscillators as identical to the driving frequency. In fact, the ability of the subjects to entrain their rhythmic movements without a phase lag to the external pacing lights is an interesting subject in its own right and deserves further theoretical and experimental study. Using again the slowly varying amplitude and the rotating wave approximations we were able to derive the following equations for the relative phase ϕ and the phase sum θ:

$$\dot{\phi} = -\frac{\partial V(\phi,\theta)}{\partial \phi} \tag{15}$$

$$\dot{\theta} = -\frac{\partial V(\phi,\theta)}{\partial \theta} \tag{16}$$

with a potential:

$$V(\phi,\theta) = -a\ cos\phi - b\ cos2\phi - c\ sin\left(\frac{\theta-\psi}{2}\right) cos\left(\frac{\phi-\psi}{2}\right) \tag{17}$$

where

$$a = -A - 2Br_0^2,\ b = \frac{1}{2}Br_0^2$$

and

$$c = F/(\omega r_0)$$

with amplitude

$$r_0 = \sqrt{\frac{\alpha}{3\beta w^2 + \gamma}}$$

If one assumes the phase sum to have relaxed to its stationary value $\theta = -\pi + \psi$ the resulting equation for relative phase is exactly:

$$\dot{\phi} = -\frac{dV\psi(\phi)}{d\phi} \tag{18}$$

with the potential $V\psi$ of equ (8). Thus again we have derived the collective variable dynamics from the component level. The details of these calculations will be published elsewhere (Schöner et al., forthcoming).

8. Conclusions: "Phenomenological Synergetics"

The laws governing the dynamic patterns produced by complex, biological systems that possess very many degrees of freedom (such as the human brain which has $\sim 10^{14}$ neurons and neuronal connections) are, in general, not known. Unlike certain physical systems, the path from the microscopic dynamics to the collective order parameters—as in Haken's *slaving* principle—is not readily accessible to theoretical analysis. Here we suggest that an understanding of biological order may still be possible via an alternative approach, namely one in which the nature and dynamics of the (low-dimensional) order parameters are first empirically determined, particularly near nonequilibrium phase transitions (or bifurcations). Then the relevant subsystems and their dynamics can be identified. This approach, which we may call "phenomenological synergetics," has provided the conceptual framework for the empirical studies of spatiotemporal order in bimanual coordination that we have reported here. Parenthetically, we have preliminary, but exciting evidence that the approach is useful for understanding another system, namely the multiple articulator movements that structure the sounds of speech (Kelso et al., 1986).

"Understanding" is sought, in the present approach, not through some privileged scale of analysis, but within the abstract level of the essential (collective) variables and their dynamics, *regardless of scale or material substrate*. Not only may the language of dynamics be appropriate at the behavioral level (e.g., in the patterns among muscles and kinematic events), but also, we hypothesize, at the more microscopic scale of neurons and neuronal assemblies. Many of the dynamical features we have observed and modeled in our experiments, for example, synchronization, phase-locking, switching, etc., can also be observed in the neuronal behavior of even the lowliest creatures, for example, in the buccal ganglion of the snail, *Helisoma* (as in Kater and colleagues' work, cf. Kater, 1985, for review) or in the motoneuronal firing patterns of *Pleurobranchaea* during feeding (cf. Mpitsos & Cohan, 1986, for review). Thus the long sought-for link between neuronal activities (microscopic events) and behavior (macroscopic events) may actually reside in the coupling of dynamics on different levels (cf. Section 7 above). Relatedly, the classical dichotomy in biology between structure and function may be one of appearance only. The present theory promotes a unified treatment, with the two processes separated only by the time scales on which they live.

Finally, we want to stress—after Bridgeman (1953) and Haken (1985)—the operational nature of the present approach. That is, dynamics are formulated for observable variables only and predictions are made that can be experimentally tested. As much as possible, a minimality strategy is followed in which all consequences of a theoretical formulation are checked for their empirical validity. Insight is not necessarily gained by increasingly accurate quantitative descriptions of data, or by using increasingly complicated dynamical equations. Rather, we seek to account for a larger number of experimental features with a smaller number of theoretical concepts.

References

Bridgeman, P. W. (1953). *The nature of physical theory*. New York: Dover Publishers.

Gardiner, C. W. (1983). *Handbook of stochastic methods for physics, chemistry and the natural sciences*. New York: Springer-Verlag.

Garfinkel, A., & Walter, D. O. (Eds.). (1983). Proceedings of "Symposium on Nonlinearity in Brain Function." *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology, 14*, R450.

Haken, H. (1983a). *Synergetics: An introduction* (3rd ed.). Heidelberg: Springer-Verlag.

Haken, H. (1983b). *Advanced synergetics*. Heidelberg: Springer-Verlag.

Haken, H. (1985). Operational approaches to complex systems. An introduction. In H. Haken (Ed.), *Complex systems–Operational approaches in neurobiology, physics* and *computers*. Berlin: Springer-Verlag.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*, 347-356.

Kater, S. B. (1985). Dynamic regulators of neuronal form and connectivity in the adult snail Helisoma. In A. I. Selverston (Ed.), *Model neural networks and behavior*. New York: Plenum.

Kay, B. A. (1986). *Dynamic modeling of rhythmic limb movements: Converging on a description of the component oscillators*. Doctoral dissertation, University of Connecticut, Storrs.

Kay, B. A., Kelso, J. A. S., Saltzman, E., & Schöner, G. (1987). The space-time behavior of single and bimanual rhythmical movements: Data and limit cycle model. *Journal of Experimental Psychology: Human Perception and Performance, 13* (2).

Kelso, J. A. S. (1981). On the oscillatory basis of movement. *Bulletin of the Psychonomic Society, 18*, 63.

Kelso, J. A. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology, 15*, R1000-R1004.

Kelso, J. A. S., Holt, K. G. (1980). Exploring a vibratory systems analysis of human movement production. *Journal of Neurophysiology, 43*, 1183.

Kelso, J. A. S., Holt, P. G., Rubin, P. & Kugler, P. (1981). Patterns of human interlimb coordination emerge from the properties of nonlinear limit cycle oscillatory processes: Theory and data. *Journal of Motor Behavior, 13*, 226-261.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. (1986). The dynamical perspective in speech production: Data and theory, *Journal of Phonetics, 14*, 29-59.

Kelso, J. A. S., & Scholz, J. P. (1985). Cooperative phenomena in biological motion. In H. Haken (Ed.), *Complex systems: Operational approaches in neurobiology, physical systems and computers*. Berlin: Springer-Verlag.

Kelso, J. A. S., Scholz, J. P., Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physica Letters A, 118*, 279-284.

Kelso, J. A. S., Schöner, G., Scholz, J. P., & Haken, J. (in press). Nonequilibrium phase transitions in coordinated movement involving many degrees of freedom. *Perspectives in Biological Dynamics and Theoretical Medicine, Annals of the New York Academy of Sciences.*

Mpitsos, G. J., & Cohan, C. S. (in press). Convergence in a distributed nervous system: Parallel processing and self-organization. *Journal of Neurobiology.*

Scholz, J. P. (1986). *A nonequilibrium phase transition in human bimanual movement: test of a dynamical model*. Doctoral dissertation, University of Connecticut, Storrs.

Scholz, J. P., Kelso, J. A. S., & Schöner, G. (Submitted) Nonequilibrium phase transitions in coordinated biological motion: Critical slowing down and switching time. *Physica Letters.*

Schöner, G., Haken, H., & Kelso, J. A. S. (1986). A stochastic theory of phase transitions in human hand movement. *Biological Cybernetics, 53*, 247-257.

Schöner, G., Kelso, J. A. S., Tuller, B., forthcoming.

Shimizu, H. (1982). In H. Haken (Ed.), *Evolution of order and chaos in physics, chemistry and biology*. Berlin: Springer-Verlag.

Shimizu, H., Haken, H. (1983). Cooperative dynamics in organelles. *Journal of Theoretical Biology, 104*, 261-273.

Tuller, B., & Kelso, J. A. S. (1985). Coordination in normal and split-brain patients. Paper presented at *Psychonomic Society*, Boston, MA.

Winfree, A. T. (1980). *The geometry of biological time*. New York: Springer-Verlag.

Yamanishi, J., Kawato, M., & Suzuki, R. (1980). Two coupled oscillators as a model for the coordinated finger tapping by both hands. *Biological Cybernetics*, 37, 219.

BEYOND ANATOMICAL SPECIFICITY*

M. T. Turvey†

I wish to address Berkinblit, Fel'dman, and Fukson's treatment of movement plasticity with respect to two classical assumptions that figure prominently in their analysis. These assumptions are (1) the existence of elemental behaviors specific to anatomical structures and (2) the formation of complexity from simplicity. The first assumption can be introduced through the Doctrine of Specific Nerve Qualities advanced by Muller and Helmholtz. According to this doctrine a given sensation (e.g., of brightness, hue, loudness, etc.) is specific to the activity in a given nerve. What causes the nerve to be active is immaterial; *that* the nerve is active is all that matters. Patently, any perception and any action must be specific to something. Generalizing the Doctrine of Specific Nerve Qualities leads to the traditional claim that the something in question is a part of the body (more exactly, the states of some of its neurons). This is the notion of anatomical specificity that has enjoyed widespread acceptance since DesCartes framed his Doctrine of Corporeal Ideas (Reed, 1982a). The study of the control and coordination of movement adverts to anatomical specificity in several ways. A host of responses, conventionally termed "reflexes," are said to be specific to afferent-efferent linkages in the spinal cord. Oscillatory movements are said to be specific to particular single neurons (pacemakers) or to circumscribed ensembles of neurons. Whole sequences of responses are said to be specific to command neurons.

Anatomical specificity implies invariable context-free movement units. For example, in the ordinary understanding of reflexes if the anatomy can be effectively isolated, then when the appropriate afferent path is stimulated a fixed efferent outflow and skeletomuscular patterning should result. Movement plasticity gets defined, therefore, as the problem of how anatomically specific, context-free movement units are adjusted and combined. One can think of the problem in these terms: Successful activity in real environments requires that movements be variable and specific to tasks (the animal's goals or intentions with respect to the surrounding layout of surfaces, including its own); the movements in the basic repertoire, however, are assumed to be invariable and specific to anatomical structures. Attempts to resolve this predicament ordinarily invoke specialized devices mediating the inputs and outputs of the anatomical structures. In that such devices embody a degree of unexplained intelligence, their inclusion in the account of spinal capabilities has been, historically, a source of controversy (Fearing, 1930/1970).

Now personally I doubt that there are, in fact, invariable anatomically specific movements. I think they owe their "existence" largely to a methodology that lifts isolated fragments out of context. As Bernstein (1967) underscored, instances of coordinated movement always involve two sets of forces: those provided muscularly and those provided nonmuscularly (reactive, coriolis, frictional). An animal generates a set of forces that, in conjunction with the nonmuscular forces,

---

* Comment on Berkinblit, M. B., Fel'dman, A. G., & Fukson, O. I. Adaptability of innate motor patterns and motor control mechanisms. *The Behavioral and Brain Sciences*, 1986, *9*, 624-625.

† Also University of Connecticut

Aknowledgment. The writing of this commentary was supported in part by an Office of Naval Research Contract N00014-83-C-0083 awarded to the Haskins Laboratories.

realizes its aim. In more picturesque terms, an animal "bends" the (variable) force structure that is given to produce the force structure that is required (Fowler & Turvey, 1978). Insofar as this would always have been true for the activity of any creature, from the smallest to the largest, it seems most unlikely that evolutionary pressure would have put a premium on "hard molded" context-free movements. What seems to be needed, in evolutionary perspective, are systems that can assemble task- or function-specific movements. Where the focus of anatomical specificity is the issue of how invariable units yield variable products, task- or functional-specificity focuses attention on reconciling invariance of a collective end state with variance in the detail of the component courses of events leading up to it. Outside of animal movements nature furnishes very many examples of multiple constellations of microevents yielding macroproducts of essentially the same standard type (Stevens, 1974; Weiss, 1967). The implication is that the phenomenon reflects self-organizing strategies of great generality. It is notable that morphogenetic processes (e.g., embryogenesis) persistently resist interpretation in terms of anatomical specificity. (Like Berkenblit et al., I believe that there are useful parallels, even convergences, to be drawn between coordinated movements and morphogenesis.) A mixture of isolated single cells, drawn randomly from an already functioning embryonic kidney and then scrambled, lumped and suitably nourished, will assemble itself into a functionally adept miniature kidney. Such examples abound and are well known (Weiss, 1967). One important lesson they provide is that a cell's role is largely determined by the conditions prevailing at the cell's location. From locale to locale conditions vary dependent on the configurational dynamics of the ensemble. Embryogenesis implicates a field perspective. Another important lesson of embryogenesis is that for functionally invariant macroproducts to arise through variable microevents, sources of stability and reproducibility (of function) must be realized in the evolving dynamics of the cell ensemble. Peter Kugler and I, working with a morphogenesis-like problem, have taken a leaf from Gibson (1966, 1979) and sought these sources in alternative descriptions (qualitative, macroscopic properties) of the dynamics (Kugler & Turvey, in press). These properties are close relatives to the "order parameters" of cooperative physical phenomena (e.g., Careri, 1984; Haken, 1978), and the equilibrium points of Berkinblit et al.'s λ-theory may well be interpreted in similar fashion.

If the notion of anatomically specific movements is suspect, as I am suggesting (see also Reed, 1982b), then the idea that there are classes of movements (termed "complex") built from simple stereotyped movement units must also be suspect. Thus, although the formation of complexity from simplicity is easily imagined, I am inclined to think that it is an improper image. Activities such as the wiping response of focal concern to Berkenblit et al. are more properly construed, perhaps, as types of simplifications from complexity (Pattee, 1972). Wiping in the frog is a simple dynamical regime as are the frog's forms of locomotion. Each regime is distilled out of an aggregation of many degrees of freedom at the muscle/joint level and very many more at the cellular level. The formation of these simple regimes from an extremely complex, functionally rich interior is to be understood as a collective or global activity contrasted with the locally specific organizational style of combining specific structures. Detailed accounts of the emergence of particular simplicities are hard to come by (cf. Kugler & Turvey, in press) but illustrations of the general process are widely documented (e.g., Haken, 1978, 1981; Prigogine, 1980).

By way of conclusion, I am taking the data provided by Berkinblit et al. on the variability and adaptability of the wiping response as evidence for a design principle of functional specificity (Fowler & Turvey, 1978; Gibson, 1966; Reed, 1982b) rather than as evidence for the modifiability of anatomically specific units. And I am suggesting that functional specificity reflects, in part, an overarching set of physical biological strategies by which state spaces of high dimensionality are systematically compressed into state spaces of low dimensionality.

References

Bernstein, N. A. (1967). *The coordination and regulation of movements*. London: Pergamon Press.

Careri, G. (1984). *Order and disorder in matter*. Menlo Park, CA: Benjamin/Cummings Publishing Company.

Fearing, F. (1930/1970). *Reflex action: A study in the history of physiological psychology*. Cambridge, MA: MIT Press.

Fowler, C. A., & Turvey, M. T. (1978). Skill acquisition: An event approach with special reference to searching for the optimum of a function of several variables. In G. Stelmach (Ed.), *Information processing in motor learning and control*. New York: Academic Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton-Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.

Haken, H. (1978). *Synergetics: An introduction*. Berlin: Springer-Verlag.

Haken, H. (1981). *Chaos and order in nature*. Berlin: Springer-Verlag.

Kugler, P. N., & Turvey, M. T. (in press). *Information, natural law and the self-assembly of rhythmic movement*. Hillsdale, NJ: Erlbaum.

Pattee, H. H. (1972). Laws and constraints, symbols and language. In C. H. Waddington (Ed.), *Towards a theoretical biology*. Chicago: Atherton-Aldine.

Prigogine, I. (1980). *From being to becoming: Time and complexity in the physical sciences*. San Francisco: W. H. Freeman & Co.

Reed, E. S. (1982a). The corporeal idea hypothesis and the origin of experimental psychology. *Review of Metaphysics, 35*, 731-752.

Reed, E. S. (1982b). An outline of a theory of action systems. *Journal of Motor Behavior, 14*, 98-134.

Stevens, P. S. (1974). *Patterns in nature*. Boston: Atlantic Monthly Press.

Weiss, P. (1967). $1 + 1 \neq 2$ (when one plus one does not equal two). In G. C. Quarton, T. Melnechuk, & F. O. Schmitt (Eds.), *The neurosciences: A study program*. New York: Rockefeller University Press.

PERCEPTUAL NORMALIZATION OF VOWELS PRODUCED BY SINUSOIDAL VOICES*

Robert E. Remez, Philip E. Rubin, Lynne C. Nygaard,** and William A. Howell†**

Abstract. *When listeners hear a sinusoidal replica of a sentence, they perceive linguistic properties despite the absence of short-time acoustic components typical of vocal signals. Is this accomplished by a postperceptual strategy that accommodates the anomalous acoustic pattern ad hoc, or is a sinusoidal sentence understood by the ordinary means of speech perception? If listeners treat sinusoidal signals as speech signals however unlike speech they may be, then perception should exhibit the commonplace sensitivity to the dimensions of the originating vocal tract. The present study, employing sinusoidal signals, raised this issue by testing the identification of target /bVt/ syllables occurring in sentences that differed in the range of frequency variation of their component tones. Vowel quality of target syllables was influenced by this acoustic correlate of vocal-tract scale, implying that the perception of these non-vocal signals includes a process of vocal-tract normalization. Converging evidence suggests that the perception of sinusoidal vowels depends on the relations among component tones and not on the phonetic likeness of each tone in isolation. The findings support the general claim that sinusoidal replicas of natural speech signals are perceptible phonetically because they preserve time-varying information present in natural signals.*

Does speech perception depend on the occurrence of specific acoustic elements within the signal spectrum? The customary conceptualizations of phonetic perception have assumed so, and have sought to describe the manner in which elementary acoustic cues bring about the perception of phoneme sequences. For example, isolatable spectral elements, such as brief formant frequency

---

* *Journal of Experimental Psychology: Human Perception and Performance*, 1987, *13*, 40-61.

** Department of Psychology, Barnard College

† Columbia College

Acknowledgment. The authors wish to thank Randy Diehl, Louis Goldstein, Jan Rabinowitz, Jim Sawusch and Michael Studdert-Kennedy for help and criticism. For sponsorship, we thank the National Institute of Neurological and Communicative Disorders and Stroke (Grant NS-22096, to R.E.R.) and the National Institute of Child Health and Human Development (Grant HD-01994 to Haskins Laboratories).

transitions, momentary aperiodicities, and low frequency murmurs, are said to be correlates of the perception of consonantal attributes. These particular three are responsible for the perception of articulatory place, voicing, and nasal manner, respectively (Delattre, Liberman, & Cooper, 1955; Zue & Schwartz, 1980). Other descriptions of the transformation of acoustic structure to phonetic properties have relied on momentary spectral shapes (Stevens & Blumstein, 1981), or spectral sequences (Klatt, 1979), or cue-weighting techniques (Massaro, 1972) or network arrangements (Eimas & Miller, 1978; Elman & McClelland, 1985). Though these mechanisms have been entertained as rivals, among others, for explaining speech perception, they all begin in the same way: They identify or filter particulate elements of the acoustic-auditory pattern, assuming them to be the ingredients essential for the perceptual process.

A few proposals, including one of our own, have departed from this point of view (Bailey & Summerfield, 1980; Jenkins, Strange & Edman, 1983; Liberman, 1970; Remez, Rubin, Pisoni & Carrell, 1981). In our work, we have employed tonal analogs of speech composed of three or four time-varying sinusoids, each one reproducing the frequency and amplitude variation of a vocal resonance in a natural utterance. A sinusoidal sentence pattern retains the overall configuration of spectral variation of the natural utterance on which it is modeled. But, it is devoid of aperiodicities, fine-grained regular (glottal) pulses, harmonic series, and broadband formant structure that compose the natural speech signal, and therefore lacks the rich assortment of acoustic particles that typically receive the emphasis in explanations of speech perception. Perceptual tests employing sinusoidal sentences offer an opportunity to test the effects of structured acoustic patterns independent of the assortment of acoustic elements found in speech signals.

Listeners judge sinusoidal signals to be unlike speech, as might be suspected from considering the short-time acoustic properties of tonal signals. Acoustically and perceptually, sinusoidal signals are grossly unnatural, and naive test subjects who are told only to identify "sounds" tend to perceive sinusoidal sentences merely as several covarying tones (Remez et al., 1981). This outcome is predicted by conventional explanations of speech perception based on acoustic elements. However, when instructed to listen for a linguistic message in the tonal patterns, such test subjects often succeed, though only when the tonal configuration abstractly represents a complex resonance pattern; a single tone representing the second formant, for example, is perfectly untranscribable.

Despite the fact that many listeners have little difficulty in comprehending sinusoidal patterns that replicate several formants, as many as a third of the listeners may fail utterly to do so. They report instead that the tone-complexes do not cohere, which perhaps encourages the view that different and mutually exclusive perceptual organizations are possible for sinusoidal signals (Bailey, Summerfield, & Dorman, 1977; Best, Morrongiello, & Robson, 1981; Cutting, 1974; Remez, in press; Remez et al., 1981; Williams, Verbrugge, & Studdert-Kennedy, 1983). However, we have yet to identify a common independent factor differentiating the majority of listeners who are able to transcribe the patterns from the minority who are unable. Our general hypothesis is that sinusoidal replicas of speech convey linguistic information in patterned variation, for few of the acoustic details typical of natural speech remain to provide the basis for phonetic perception. If our interpretation is plausible, then these phenomena are evidence of a kind of acoustic information in speech perception that is time-varying and relatively independent of elemental constituents. The sufficiency of time-varying information presents a strong challenge to perceptual accounts that describe the information in acoustic signals as discrete cues to be extracted from the speech stream like raisins from pudding. Moreover, phonetic perception of tone complexes resists description by prototype models that favor schematizations of typical acoustic elements or their auditory representations (for example, Massaro, 1972; Samuel, 1982).

Time-varying Phonetic Information

The patterns of variation in sinusoidal sentences are derived from natural utterances, which suggests a clue to understanding the skill shown by listeners transcribing these spectrally unfamiliar tonal patterns. Sinusoidal signals may preserve phonetic information present in the natural signals on which they are modeled. Perhaps such information is available ordinarily in the coherent variation of natural signals. Certainly, listeners require no special training to perform adequately in the transcription test, which implies that sinusoidal signals replicate the information available from coherent signal variation despite the gross short-time differences between sinusoidal and natural spectra. Sinusoidally presented linguistic attributes would therefore be perceptible because this time-varying information is treated in an ordinary way–as occurs with time-varying phonetic information in a natural signal.

As attractive as that conclusion is for us, there is an alternative hypothesis that may fare as well in accounting for the findings. Suppose that sinusoidal imitations of speech signals merely preserve aspects of acoustic structure that are irrelevant to phonetic identification. In the absence of the natural acoustic products of vocalization, the listener may exploit the residual, rough resemblance of sinewave variation to speech signals in order to imagine a linguistic sequence that plausibly fits the tones. If so, then we may expect transcription of sinusoidal replicas of sentences to occur via postperceptual compensation for the failure of phonetic perception, subjecting the nonspeech percept to inference, analogic reasoning, and outright guesswork in concocting a linguistic likeness.

An example from the archives that appears to fit this postperceptual description is contributed by Newton, who wrote:

> Soe ye greatest cavity in ye mouth being first made in ye throate & thence by degrees moved towards ye lipps further from ye larinx causes ye pronunciation of ye vowells in order y i e a o ω u w. The filling of a very deep flaggon with a constant streame of beere or water sounds ye vowells in this order w u ω o a e i y. (Newton, 1665, quoted by Ladefoged, 1967; page 65).

It seems that Newton heard the changing pitch of the flagon resonance as a series of vowels–[u] when the flagon was empty and had low pitched resonance, through [i] when the flagon was almost full and had high pitch. Surely, Newton never thought the flagon spoke vowels to him, yet his transcription performance would presumably have been reliable. This is a credible instance, then, in which an auditory experience is likened to speech by deliberate afterthought, and is distinct from the immediate perception of speech.

Taking Newton's observation into account, our immediate question may be clearly posed: Is perception of sinusoidal signals like speech perception, or rather like listening to nonphonetic sounds and then inventing a sentence to match a nonspeech pattern? The answer determines whether research employing tonal analogs of natural signals contributes a valid approach to studying the perception of speech.

Unfortunately, there is no direct way to establish whether perception of tone complexes is primarily phonetic or nonphonetic. We cannot appeal to a definitive test to discover whether acoustic patterns are perceived to be composed of speech sounds or perceived to be nonphonetic but less unlike some speech sounds than others. We may obtain an indirect answer by observing whether phonetic identification of a sinusoidal signal is affected by the implied dimensions of the vocal tract that seems to produce the tonal sentence. Listeners should demonstrate perceptual normalization of vocal-tract dimensions only if there is information in the sinusoidal pattern to mark the tone patterns as implicitly vocal; and only if the tone variations possess sufficient structure to evoke the low-level perceptual evaluation of the sinusoidal voice relative to the range of potential talkers (see Fant, 1962; Joos, 1948; Nearey, 1978). In other words, the solution to the puzzle can be obtained by seeing whether phonetic transcription of sinusoidal signals exhibits perceptual normalization of the acoustic properties correlated with intertalker variation in vocal-tract dimensions. If so, this would be evidence for a basic and early perceptual function keyed to speech sounds, from which we could infer that sinusoidal transcription is an instance of speech perception.

Experiment 1

Normalization of Vocal-tract Dimensions

Our test makes use of a classic finding by Ladefoged and Broadbent (1957), that the perception of speech entails a process that accommodates the acoustic signal properties attributable solely to variation among individual talkers, independent of linguistic factors. The root of this acoustic variation is the corresponding variation in vocal anatomy and in the control of the articulators (Bricker & Pruzansky, 1976; Joos, 1948). Briefly, when different talkers speak the same word, the specific frequencies traversed by the formants differ from one talker to another. Some vocal tracts are longer, some shorter, producing a kind of scalar variation in formant patterns observed across a range of talkers employing the same linguistic elements (see, also, Fant, 1966; Peterson & Barney, 1952). Differences in vocal gestures and corresponding formant trajectories may also result from differences in articulatory control (Ladefoged, 1980). The consequence is plain; the perceiver cannot identify vowels simply from momentary formant frequency values–vowels must first be implicitly rescaled with reference to the talker's formant range, in other words, with respect to the capability of the vocal tract that produced them. Because the formant pattern of one talker's "bet" can be identical to that of another's "but," this rescaling ensures that the word produced by the talker is perceived as such, all other things being equal.

Ladefoged and Broadbent made perceptual normalization an empirical issue using the technique of speech synthesis. Initially, they produced synthetic imitations of a natural utterance, "Please say what this word is," along with approximations of the words, "bit," "bet," "bat," and "but." The synthetic sentence reflected the vocal-tract dimensions of the talker who produced the natural utterance from which the sentence synthesis was derived. However, transposition of individual formant tracks of the sentence pattern in the frequency domain gave it characteristics of vocal tracts of rather different dimensions. Ladefoged and Broadbent lowered (by 25 percent) or raised (by 30 percent) the first or second formants (or both) to create impressions of different talkers from the original sentence pattern. In the normalization test, listeners labeled the target syllables that differed only in the identity of the vowel. Each of the four targets was presented with the untransposed sentence frame, to establish baseline labeling, and was also presented with one or more transposed sentences as the frame. In the latter cases the distributions of judgments often differed from those that had been observed with the untransposed, natural sentence pattern as the frame.

The experimental outcomes were rationalized by assuming that perceptual rescaling occurred on the basis of the leading sentence frame, and was revealed in the effects of different frames on identification of the same lagging target syllable. The perceiver presumably identified the target vowel by reference to the range of formant variation encountered in the frame. This seemed reasonable because the capability possessed by any talker for producing formant frequency excursions is limited by the size of the vocal cavities, and the formant values associated with any specific vowel posture or gesture are governed therefore by the overall range.

The contingency of vowel identity on the scale of the originating vocal tract has been described as adaptation to the personal dimension of the information in speech signals (Ladefoged & Broadbent, 1957); as perceptual normalization for talker (Gerstman, 1968; Nearey, 1978) based on inferences about the physical characteristics of the sound source (Fourcin, 1968); as calibration or mapping of the talker's vowel space (Dechovitz, 1977a; Verbrugge, Strange, Shankweiler, & Edman, 1976); as vowel normalization (Disner, 1980; Sussman, 1986; Wakita, 1977); and, as vocal-tract normalization (Joos, 1948; Rand, 1971; Shankweiler, Strange, & Verbrugge, 1977; Summerfield & Haggard, 1973). Susceptibility to vocal-tract normalization may serve as an index of the phonetic perceptual effects of sinusoidal signals, because it is unlikely that effects attributable to normalization arise postperceptually. First, normalization in speech perception is rudimentary, and is evident early in development (Kuhl, 1979; Lieberman, 1984). Second, though normalization may be automatic and effortless in speech perception, it is implausible to suppose that listeners are equally adept in deliberate guessing, postperceptually, about articulatory geometry. The history of several centuries of such intuitive efforts antedating the use of modern techniques shows that the dimensions of speech production were commonly described imprecisely or misconstrued (Ladefoged, 1967; chapter 2). Third, it is doubtful that the detailed precategorical acoustic structure presumably required for a postperceptual approximation of normalization survives long enough in memory to be useful in this regard (for example, Darwin & Baddeley, 1974). If listeners show evidence of normalizing sinewave replicas, then the perceptual explanation ordinarily offered for this kind of effect with synthetic speech would therefore apply with equal force to the sinusoidal case.

Therefore, the first experiment that we report here attempted to obtain a verdict about sinusoidal replication: Do subjects perform the transcription of sinusoidal signals by relying on phonetic perception or on a postperceptual elaboration of auditory perception? If we observe effects of normalization, then we may say that perception of tonal analogs of speech occurs in a manner akin to ordinary phonetic perception.

Method

Acoustic Test Materials

Six versions of the sentence, "Please say what this word is," and the four target words *bit*, *bet*, *bat*, and *but* were prepared by the technique of sinewave replication of natural speech signals.

An adult male speaker (R.E.R.) of Northeast American English produced a single utterance of the sentence and of each of the four targets in a sound attenuating chamber (I.A.C.). These were recorded with a condenser microphone (Shure SM-78) on audiotape (Scotch No. 208) using a half-track recorder (Otari MX5050).

Spectra were determined by digitizing the utterances and analyzing the records. After filtering the playback signal by low-passing at 4.5 kHz, it was sampled at 10 kHz with 12-bit resolution and stored on a DEC VAX 11/780-based system. The values of the frequencies and amplitudes of the formants were computed at intervals of 10 ms throughout each natural utterance using the technique of linear prediction (Markel & Grey, 1976). In turn, these values were used to control a sinewave synthesizer (Rubin, 1980), which computes waveforms of signals generated by multiple independent audio-frequency oscillators. Four sinusoids were used to replicate the sentence pattern, one for each of the three lowest oral formants and a fourth sinusoid for the fricative resonance when appropriate. Spectrographic representations and spectral sections are shown in Figures 1 and 2 for the natural sentence that served as the model, and in Figures 3 and 4 for the sinusoidal replica. In the present test this sentence, which replicates the linear prediction estimates of the resonant frequencies of the natural utterance, functions as the untransposed *natural* precursor. Three sinusoids were used to replicate each of the [bVt] targets. Table 1 contains the frequency values for the three tones of each target word determined at the syllable nuclei.

Table 1

Frequency Values of the Vowel Nuclei of the Three-tone Target Syllables

Syllable	Tone 1	Tone 2	Tone 3	Duration
bit	393	1801	2572	110
bet	624	1688	2486	130
bat	805	1489	2376	169
but	624	1408	2477	150

Frequency values in Hz; duration in ms.

Five additional versions of the sentence were prepared to parallel the conditions used by Ladefoged and Broadbent, by transposing the frequency values of the sinusoid that followed the first or the second formant of the natural utterance. They are: (a) with Tone 1 lowered by 25%; (b) with Tone 1 raised by 30%; (c) with Tone 2 lowered by 25%; (d) with Tone 2 raised by 30%; and, (e) with Tone 1 lowered and Tone 2 raised together, respectively, by 25% and 30%.

Test sequences were prepared on the VAX and recorded on half-track audiotape (Scotch No. 208) following digital to analog conversion. The acoustic materials were delivered binaurally to test subjects via playback taperecorder (OTARI MX5050), power amplifier (Crown D-75) and matched headsets (Telephonics TDH-39, 300 ohm/channel), attenuated approximately to 60db (SPL).

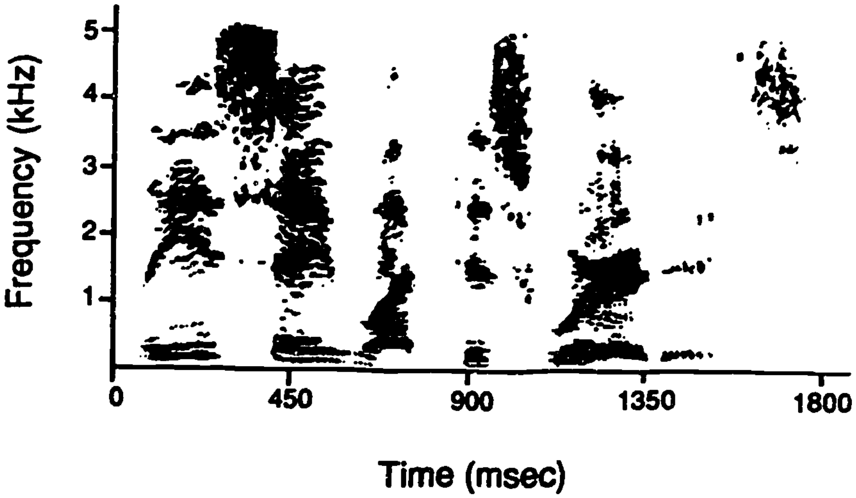


Figure 1. Spectrographic representation of the natural sentence, "Please say what this word is."

Procedure

A testing session included four parts. The first was a warm-up in which eight sinusoidal sentences of varying phonetic composition were presented for transcription. This segment was used simply to accustom the listeners to the unusual timbre of the signals, and was not scored.

The second segment was the normalization test replicating the procedure of Ladefoged and Broadbent, in which target syllables were presented for identification within sentence frames. This test consisted of eleven trials with the same format: a sentence frame (2400 ms), a silent interval (500 ms), and a target syllable (on the average, 140 ms). Successive trials were separated by 10 s of silence. There were eleven different conditions in this test, one trial per condition, each trial separated from the preceding and following trial by a long silence. This procedure was adopted from Ladefoged and Broadbent, and aimed to prevent subjects from developing familiarity with the small set of syllables and identifying them by rote rather than by perceiving the vowels. The target *bit*, was presented with two frames: natural and Tone 1 lowered; *bet* was presented with four frames: natural, Tone 1 lowered, Tone 1 raised, and Tone 1 lowered and Tone 2 raised; *bat* was presented with three frames: natural, Tone 1 raised, and Tone 2 raised; and *but* was presented with two frames: natural and Tone 2 lowered. The eleven trials were presented in random order. Listeners identified the target on each trial by marking a sheet on which the words "BIT BET BAT BUT" appeared. They were instructed to guess if they had no clear phonetic impression of the target.

The third segment of the test session was another test using sinewave signals and is not reported here.

The last segment of the session consisted of a 40-trial identification test in which the target syllables were presented ten times each in random order for labeling in a four-alternative forced-choice paradigm. On each trial subjects reported their impressions on a response form

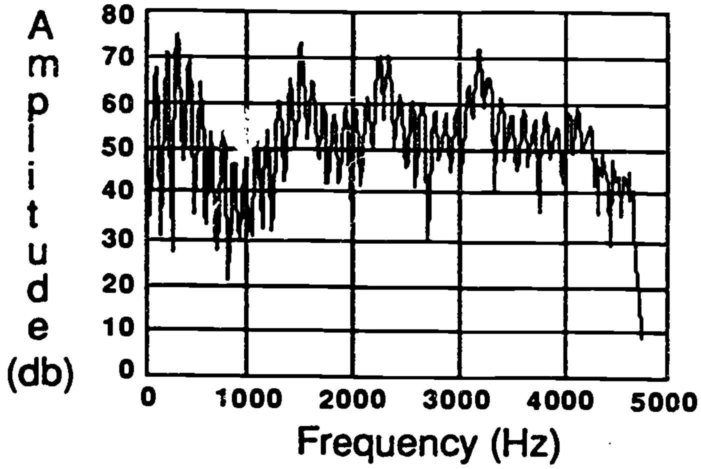


Figure 2. Spectral section of natural sentence.

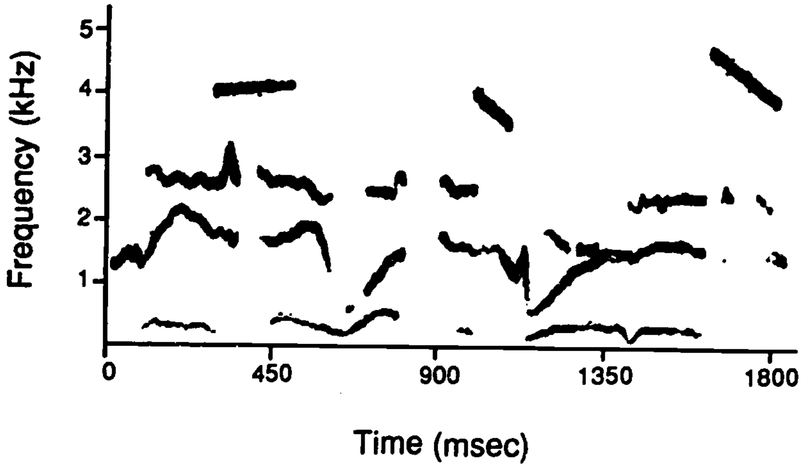


Figure 3. Spectrographic representation of the sinewave pattern derived from the natural sentence, "Please say what this word is," in which natural formant frequencies and amplitude values were used to control digital audiofrequency oscillators. (Three tones follow the oral formants; a fourth tone replaces the fricative formant where appropriate.)

containing the words "BIT BET BAT BUT." This test measured the distinctiveness of the target syllables independent of the influence of the framing sentences, and served as a method for determining which subjects were unable to identify the sinusoidal targets consistently even under highly favorable conditions.

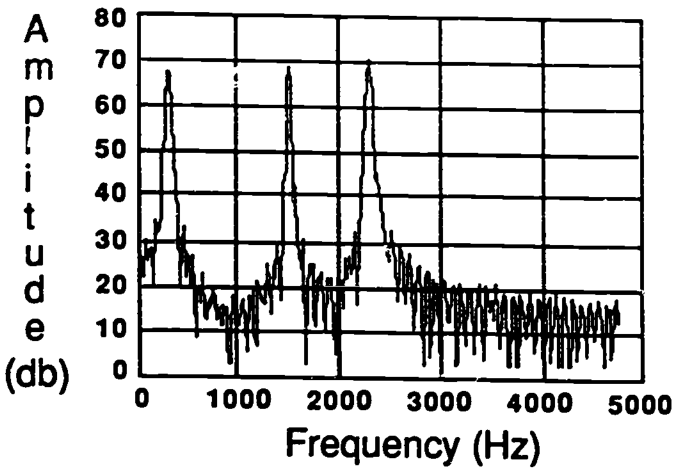


Figure 4. Spectral section of sinewave signal.

Subjects

Ninety-five undergraduate volunteers took the listening tests. None reported a history of speech or hearing disorder. Most were paid; others received course credit in Introductory Psychology for participating. All of the subjects were naive with respect to sinusoidal replicas of speech signals. They were tested in groups of two to six at a time, in visually isolated listening stations.

Results

Identification Test

In this test, the targets were presented in isolation for labeling, thereby determining the stability of the vowel categories given the sinusoidal presentation, while also providing an index of the ability of each subject to treat sinusoidal signals phonetically. To expose both aspects of the data, the results of this test are presented in Figures 5, 6, and 7.

Figure 5 shows the identification performance for the entire group of ninety-five subjects. For each of the four targets, the distribution of labeling judgments is shown, the height of each bar within the foursomes representing the proportion of the responses assigned by the listeners to that category. Each target was presented ten times, hence each group of four bars represents 950 trials. The majority of responses agreed with the intended identity of the target, showing that brief [bVt] syllables are identifiable despite the sinusoidal realization of the resonance pattern. However, there is also evidence of inconsistent identification observable in the responses distributed across the unintended and less preferred alternatives. To resolve the outcome of this test more clearly, we have separately plotted the results contributed by those subjects whose performance is less differentiated overall from those who appear more able to perform the identification task. Figures 6 and 7 portray the identification results for the listeners grouped in two sets: in Figure 6 those

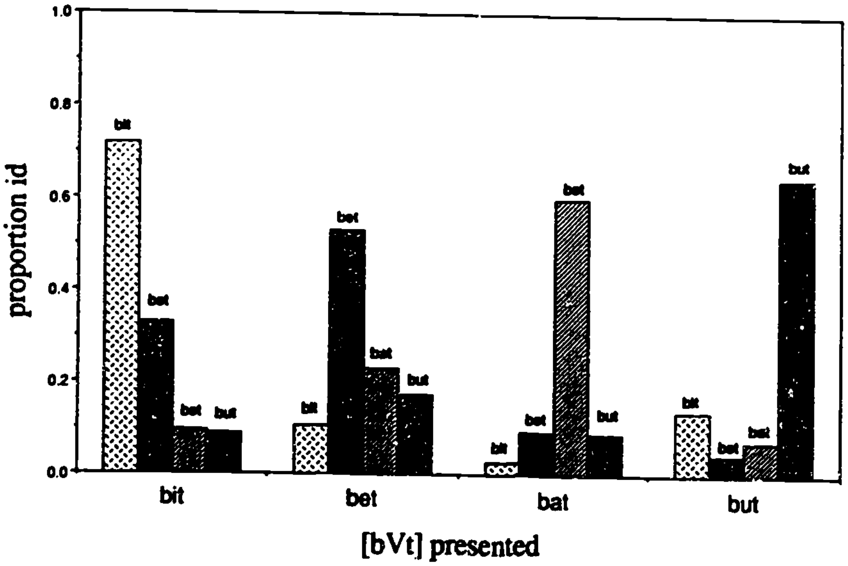


Figure 5. Group identification performance, [bVt] test, for all 95 subjects. (Proportion id = proportion of identification judgments.)

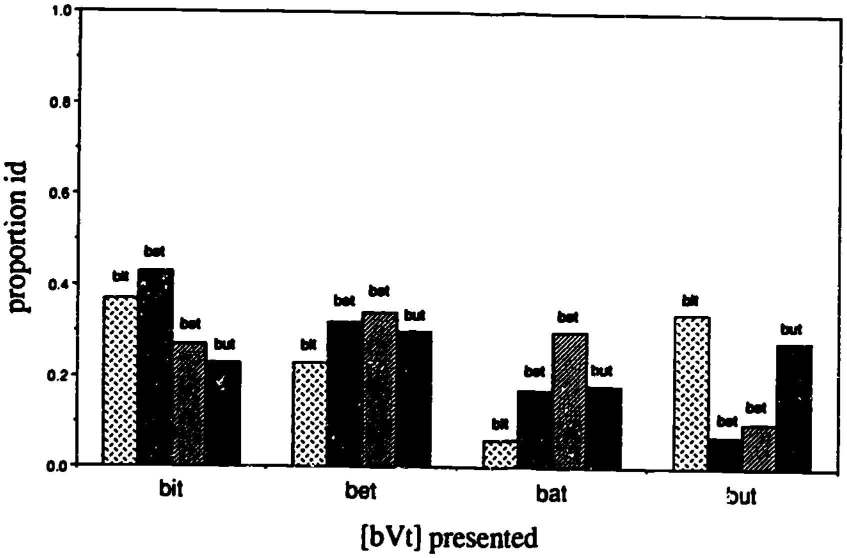


Figure 6. Group identification performance, [bVt] test, for 33 subjects whose performance was inconsistent. (Proportion id = proportion of identification judgments.)

subjects who appeared relatively less able to identify sinewave patterns, and in Figure 7 the others who exhibited greater consistency.

The measure that we used to divide the subjects into two groups–those who could attribute phonetic properties to sinusoidal patterns reliably and those who could not–was straightforward. Because the preferred response alternatives of the entire group of 95 subjects matched the intended

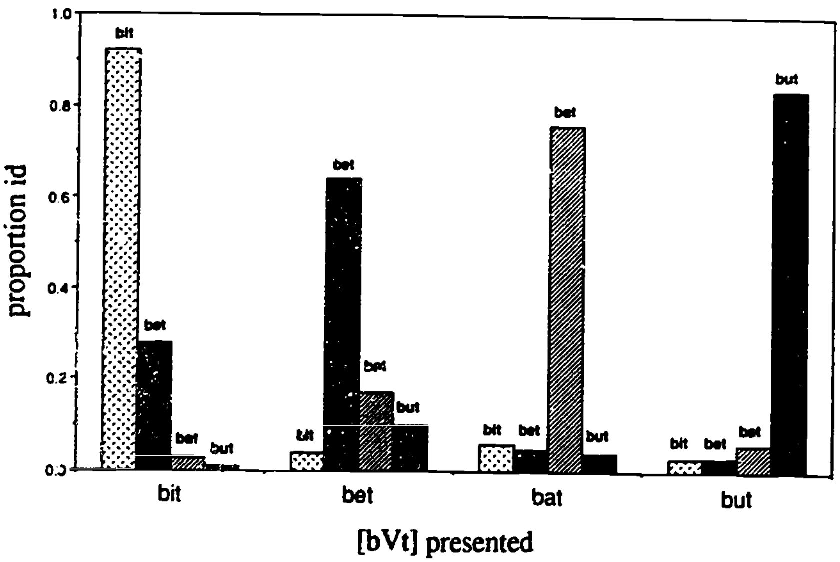


Figure 7. Group identification performance, [bVt] test, for 62 subjects whose performance was consistent. (Proportion id = proportion of identification judgments.)

identity of the sinusoidal syllables, we determined for each subject the percent of responses (out of the 40 trials) that fit this pattern. But, we adopted a rather lenient cutoff, 50% of the available responses, as the threshold for identifying those subjects who simply failed to perform, thereby including with the majority those subjects whose performance may have been phonetically stable only for some of the syllables. In each instance, if fewer than half of a subject's responses were properly assigned to the intended categories across all four target syllables, then we designated that subject as unable to perform the task. Figure 6 shows the response distribution for the group of 33 subjects who selected the less preferred and unintended alternatives more than 50% of the time. Our examination of the individual subject data revealed rather little about this group, which evidently used a mixture of gambits including random labeling and arbitrary response preferences.

Figure 7 presents the identification data obtained from the remaining 62 subjects who assigned their responses to the majority category on more than 50% of the target presentations. This group is clearly able to attach vowel labels to sinusoidal patterns.

Overall, the data reveal that sinusoidal versions of [bVt] syllables contain sufficient variation to permit listeners to categorize the vowels. In addition, as many as one-third of the subjects were found to be incapable of performing the labeling task, even under these conditions of low uncertainty, echoing the findings of Bailey et al. (1977), Best et al. (1981), Cutting (1974), and Williams et al. (1983).

Normalization Test

Because the identification test served as an independent index of each subject's susceptibility to phonetic perception with brief sinusoidal signals, we consider the normalization test results separately for the two groups of listeners whose performance is differentiated by this test. The

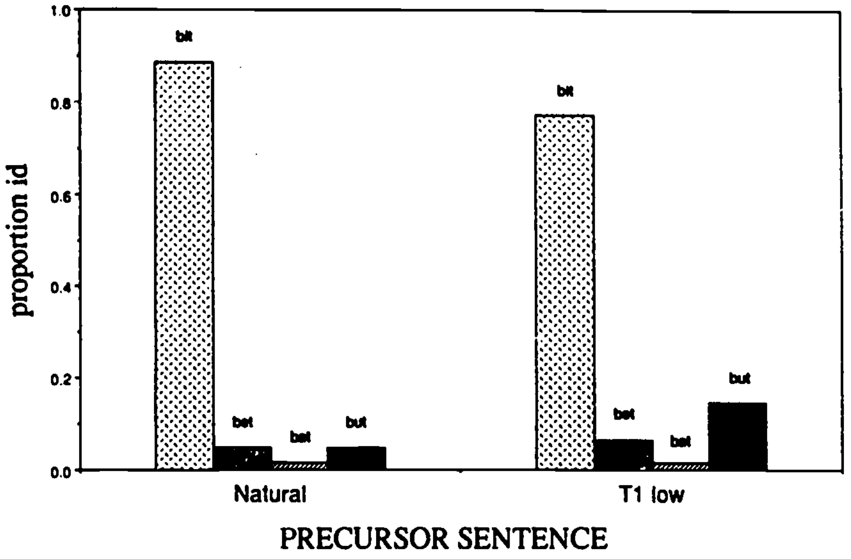


Figure 8. Normalization test performance, phonetic listeners: conditions with *bit* target. (Proportion id = proportion of identification judgments.)

null hypothesis for each of the four target conditions is the same: If normalization did not occur, then identification of the target will be indifferent to the properties of the precursor sentences, and the distribution of the identification judgments will not be altered by different precursor sentences. In fact, to summarize the results, the test showed the influence of the leading sentence in three of the four target conditions with phonetic listeners, and in one of the four conditions with nonphonetic listeners.

Phonetic listeners. The results for this group of 62 listeners are shown in Figures 8-11. When the target syllable was *bit*, shown in Figure 8, no difference was observed in labeling for the two sentence frames: the first natural (the sentence with unmodified frequency values), the second with Tone 1 lowered, $\chi^2(3, N = 62) = 3.62, p > .25$. This differs from the finding reported by Ladefoged and Broadbent, in whose study this modification of the framing sentence caused the target syllable *bit* to appear to be *bet*.

As shown in Figure 9, the target syllable *bet* was presented with four different sentence precursors. Ladefoged and Broadbent had found that one of these, in which the first formant was raised, created the impression that the target was *bit*. Similarly, we found that the analogous transposition of the sinusoidal precursor sentence increased the proportion of *bit* responses to the *bet* target. This effect was significant, $\chi^2(9, N = 62) = 20.72, p < .025$.

Figure 10 shows the outcomes for the three precursor sentences that were used with the target *bat*. Ladefoged and Broadbent had found that raising the first formant in the precursor sentence alone altered the perception of the target, increasing the proportion of trials on which *bat* was identified as *bet*. We also observed the same effect, in which instance raising Tone 1 of the sentence increased the frequency of *bet* labels applied to the *bat* target. In addition, raising Tone 2 made identification difficult, with the single exception that subjects consistently rejected

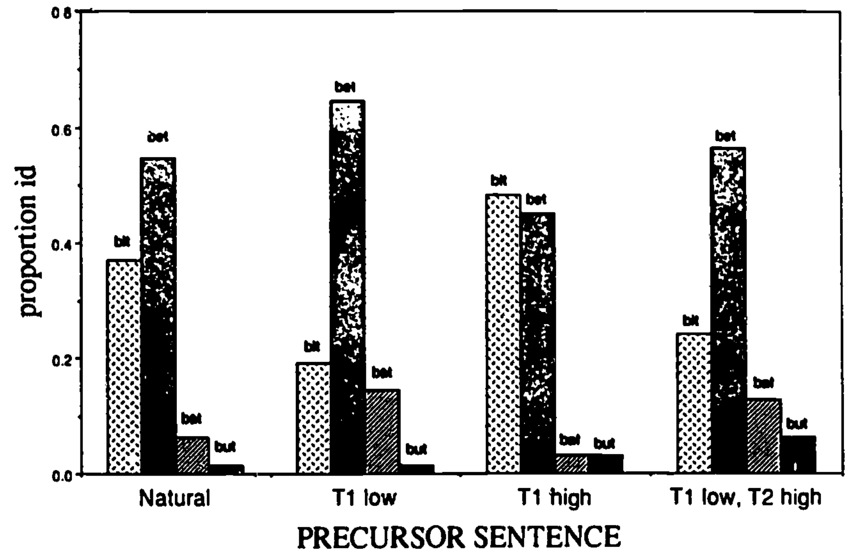


Figure 9. Normalization test performance, phonetic listeners: conditions with *bet* target. (Proportion id = proportion of identification judgments.)

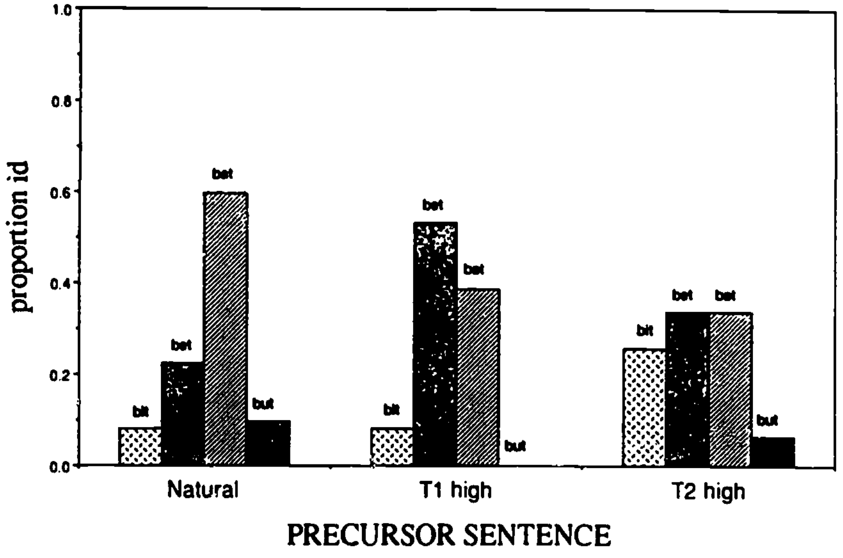


Figure 10. Normalization test performance, phonetic listeners: conditions with *bat* target. (Proportion id = proportion of identification judgments.)

but as the identity of the target. These differences attributable to the precursor sentences are statistically significant, $\chi^2(6, N = 62) = 28.35, p < .001$.

In the last condition, portrayed in Figure 11, the target *but* was presented with two precursors, the natural sentence and a sentence with Tone 2 lowered. The difference in labeling attributable to the precursor is significant, $\chi^2(3, N = 62) = 13.42, p < .005$, though the pattern of outcomes

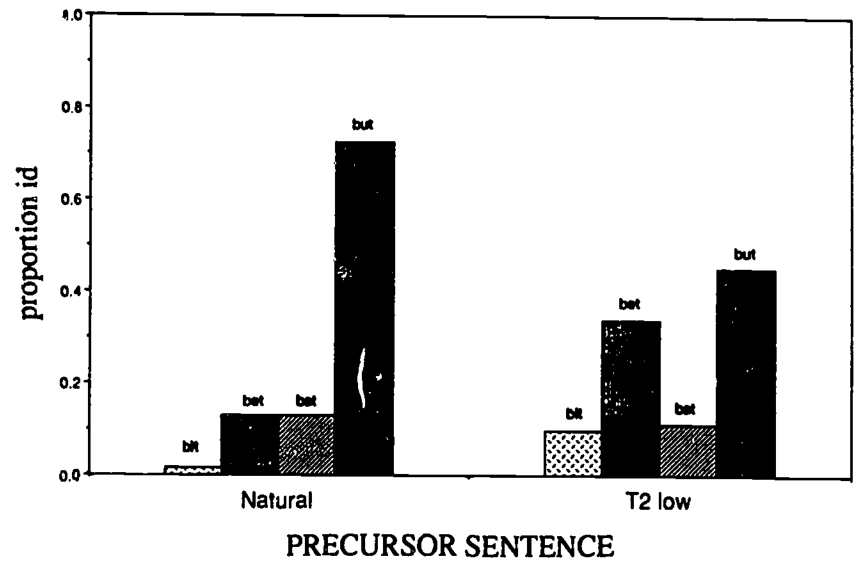


Figure 11. Normalization test performance, phonetic listeners: conditions with *but* target. (Proportion id = proportion of identification judgments.)

is unlike the effects reported by Ladefoged and Broadbent. They observed a shift in responses from *but* to *bat*, while we saw a shift from *but* to *bet*. Nonetheless, in this condition we may again identify the influence of the leading sinusoidal sentence on the vowel of the lagging target syllable.

Nonphonetic listeners. The results of the four target syllable conditions for the 33 nonphonetic listeners designated by the identification test are shown in Figure 12. In three conditions, no statistical differences were observed in the response distributions for different precursor sentences. When *bit* was the target (Figure 12), $\chi^2(3, N = 33) = 2.35, p > .5$; when *bet* was the target (Figure 13), $\chi^2(9, N = 33) = 6.06, p > .5$; when *but* was the target (Figure 15), $\chi^2(3, N = 33) = 1.91, p > .5$.

When the target was *bat*, it was poorly identified with the natural precursor; identified as *bet* when the precursor contained a raised Tone 1; and was identified as *bit* when the precursor contained a raised Tone 2, $\chi^2(6, N = 33) = 23.86, p < .001$. Note that this pattern differs greatly from what we observed for the phonetic listeners (compare Figure 10 and Figure 14), and is also unlike the outcome observed by Ladefoged and Broadbent in the analogous condition.

Discussion

Overall, these data reveal a pattern of vocal-tract scale effects in three-tone vowel identification. This implicit accommodation for variation in vocal-tract dimensions occurs despite the fact that the phonetic and vocal-tract information is carried by an anomalous sinusoidal signal. Although there are a few points of dissimilarity between the present data and those described by Ladefoged and Broadbent, the results are similar. They may therefore be taken here as evidence for talker normalization and, hence, for perceptual evaluation of sinusoidal replicas akin to ordinary perception of speech, and for the likely reliance of this process on the time-varying

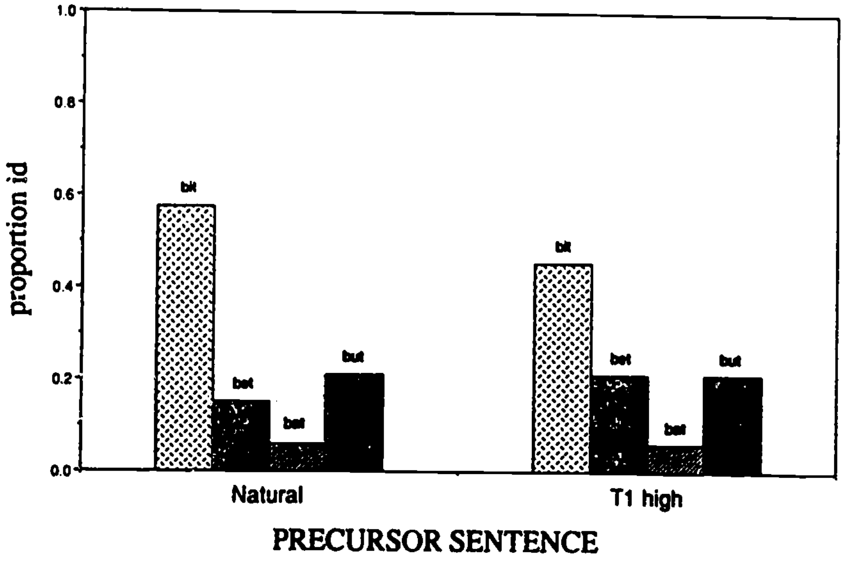


Figure 12. Normalization test performance, nonphonetic listeners: conditions with *bit* target. (Proportion id = proportion of identification judgments.)

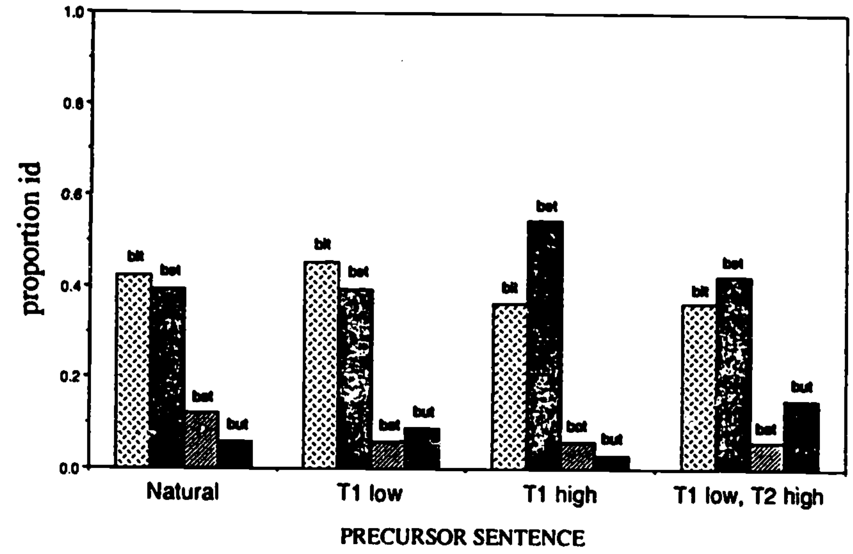


Figure 13. Normalization test performance, nonphonetic listeners: conditions with *bet* target. (Proportion id = proportion of identification judgments.)

information in sinusoids ordinarily present in natural signals. Accordingly, postperceptual guesswork or strategic problem solving is probably not primary in the transcription of tonal patterns, though it cannot be eliminated as a secondary contribution.[1] To offer an account of the outcome,

---

[1] Postperceptual guesswork and strategic problem solving also play a secondary role in ordinary perception of speech.

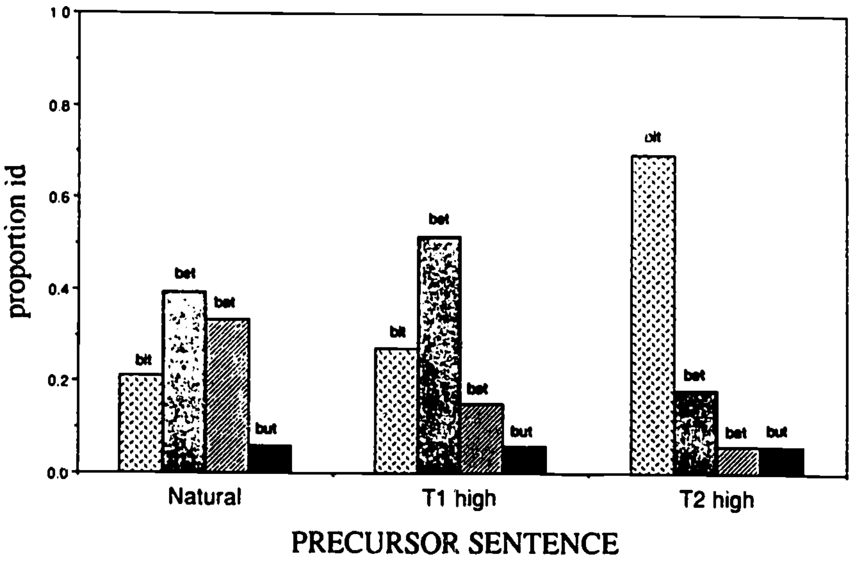


Figure 14. Normalization test performance, nonphonetic listeners: conditions with *bat* target. (Proportion id = proportion of identification judgments.)

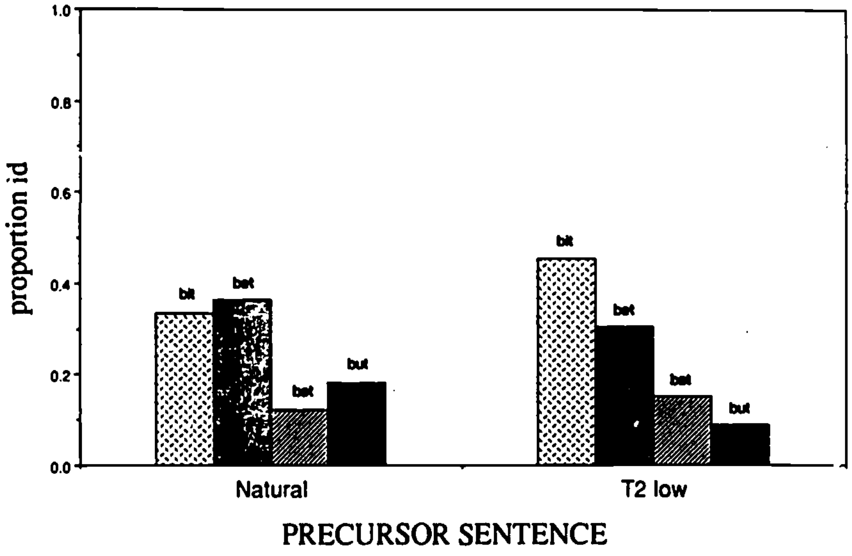


Figure 15. Normalization test performance, nonphonetic listeners: conditions with *but* target. (Proportion id = proportion of identification judgments.)

however, we must address three prominent issues raised by the results: (a) in detail, the results for phonetic listeners in this study differ from the outcomes reported with synthetic speech; (b) despite the apparent inability to identify the targets, the nonphonetic listeners were influenced by the precursor sentences in the instance of the *bat* target; and (c) most generally, our test seems to differentiate phonetic and nonphonetic listeners, though it is unclear what this means.

A Comparison to Ladefoged and Broadbent

The effect of normalization can be likened to recalibration of the formant frequency values of the perceptual standards for vowel categories. Ladefoged and Broadbent took this approach. Presumably, the changes in target identity that we observed in the sinusoidal cases likewise were caused by resetting the standards for vowel perception, all other things being equal (see Ainsworth, 1974, for a discussion of a durational component in normalization). To draw a parallel to Ladefoged and Broadbent's precedent, we adopt their heuristic, defining the perceptual standard for a vowel as a point in two dimensions, frequency of Tone 1 (the analog of the first formant) by frequency of Tone 2 (the analog of the second formant).

Operationally, we assume that the first and second formant frequencies of the syllable nuclei given in Table 1 approximate the vowel standards for the vocal tract that issued the natural sentence and the four target syllables. The effects of normalization can be estimated by transposing the vowel standards in accordance with the various precursor sentences that we used. Then, projecting the untransposed tone frequencies of the targets against recalibrated perceptual standards permits the evaluation of the target within a normalized vowel space.

Four examples of this approach are portrayed in Figures 16-20. We found that when Tone 1 was lowered in the precursor sentence, neither the categorization of *bit* nor *bet* changed. (Following the procedure of Ladefoged & Broadbent, the two other sinusoidal targets were not tested with this precursor.) In Figure 17, we represent the effect of recalibrating the vowel standards by lowering the Tone 1 value by 25%, which places the *bit* and *bet* targets in positions between the putative category centers rather than within new categories. This simplification of normalization agrees with the observations in the sinewave cases. In contrast, Ladefoged and Broadbent had observed *bit* identified as *bet*, and an absence of a normalization effect in the case of the synthetic *bet* target. Our sinewave results appear consistent with the rationale if not the specific finding of the earlier investigation. There is a probable reason that this occurred. The sinewave targets and sentences were linguistically similar to the synthetic speech of the prior study, but were not acoustically identical in the values of spectral peaks, either in the sentence set or in the targets. Our test materials also differed from the original in the dialect of English that was used. In the present test, sinewave synthesis values were derived from new samples of natural speech and therefore replicated natural productions with the approximate linguistic and acoustic attributes, rather than replicating in detail the synthetic acoustic materials of the original test. In consideration of this, a departure from the fine grain, though perhaps not the general finding, of the earlier research is to be expected. The interpretation ultimately hinges on the internal consistency of the present findings and the general correspondence of the sinewave results to the synthetic speech precedent. In this respect, the results of lowering Tone 1 in the sentence frame are best considered along with the outcomes of other conditions.

Raising Tone 1 in the precursor sentence rendered the sinewave *bet* ambiguous, with responses mostly divided between *bet* and *bit*, and made *bat* seem like *bet*. As Figure 18 reveals, the normalized *bet* target in this instance is intermediate between the IH and EH categories, while the *bat* target falls between EH and UH. Because our listeners judged *bat* to be like *bet* on 20% of the Natural context trials, perhaps the intermediacy of the *bat* target between EH and UH should be interpreted as an advantage for EH. A similar argument may apply to the *bet* target in this condition with respect to IH and EH. In both cases, the figure minimally suggests that raising Tone 1 should make the *bet* and *bat* target vowels seem higher along the vowel dimension high-low,

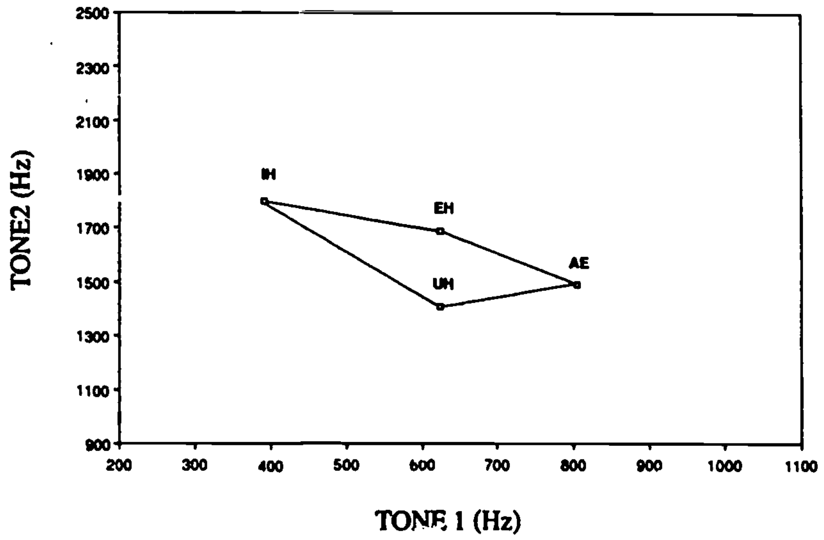


Figure 16. Sinewave nuclei: normalization portrayed as recalibration of perceptual vowel standards. (Hypothetical values for normalized vowel standards are estimated by transposing the first or second formant frequencies observed at the syllable nuclei of the four target vowels, and are labeled IH, EH, AE and UH. Within each condition, the appropriate target syllable is represented as a point in this space, and is labeled in the format bVt. This figure shows the relative positions of untransposed targets).

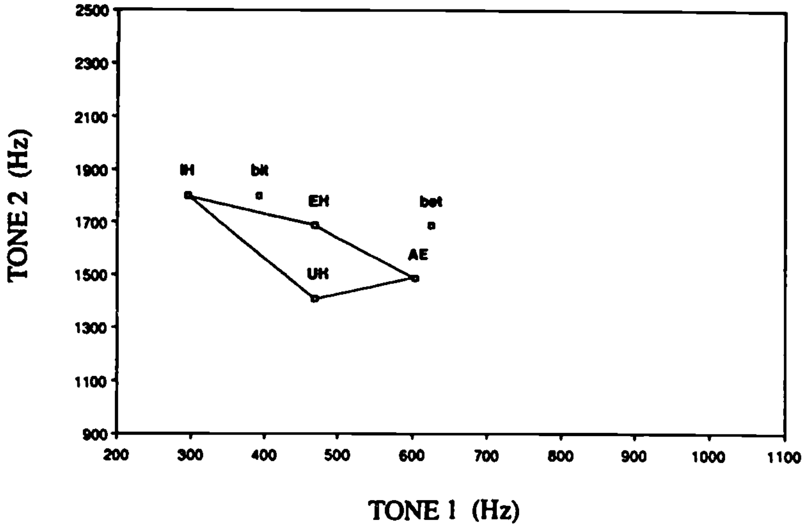


Figure 17. Sinewave nuclei: normalization portrayed as recalibration of perceptual vowel standards–effect of lowering first formant tone by 25%. (Normalized standards=IH EH AE UH targets=bVt.)

and this is a reasonable description of the performance of our listeners. Ladefoged and Broadbent, in conditions that ours paralleled, observed that raising the frequency of the first formant in the

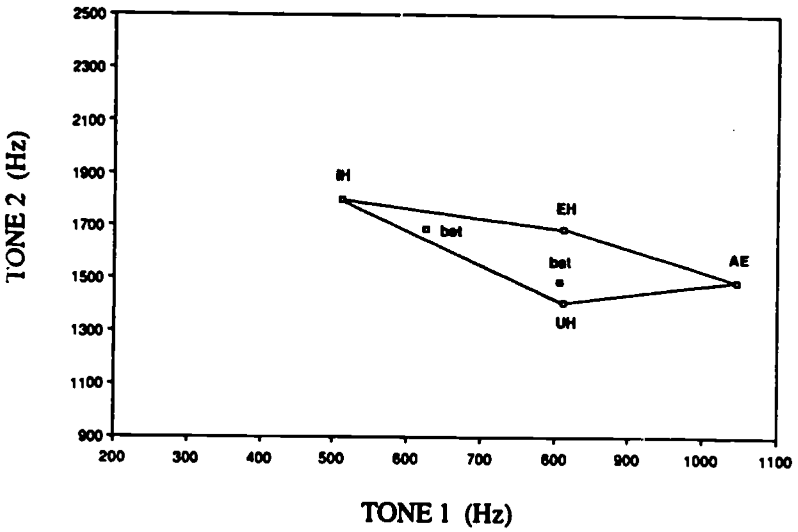


Figure 18. Sinewave nuclei: normalization portrayed as recalibration of perceptual vowel standards-effect of lowering first formant tone by 30%. (Normalized standards=IH EH AE UH targets=bVt.)

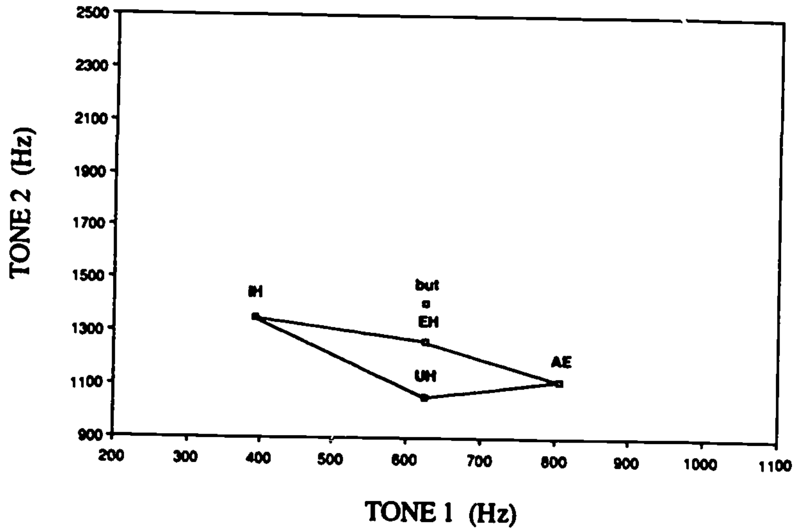


Figure 19. Sinewave nuclei: normalization portrayed as recalibration of perceptual vowel standards-effect of lowering second formant tone by 25%. (Normalized standards=IH EH AE UH targets=bVt.)

precursor sentence caused *bet* to be identified as *bit* and an ambiguous *bat* to be identified as *bet*, both changes in vowel height. In other words, the effects we observed have the same pattern as the synthetic precedent, although they fall short of complete parity because of acoustic differences between the sinewave and synthetic materials, because the variability in the sinewave response

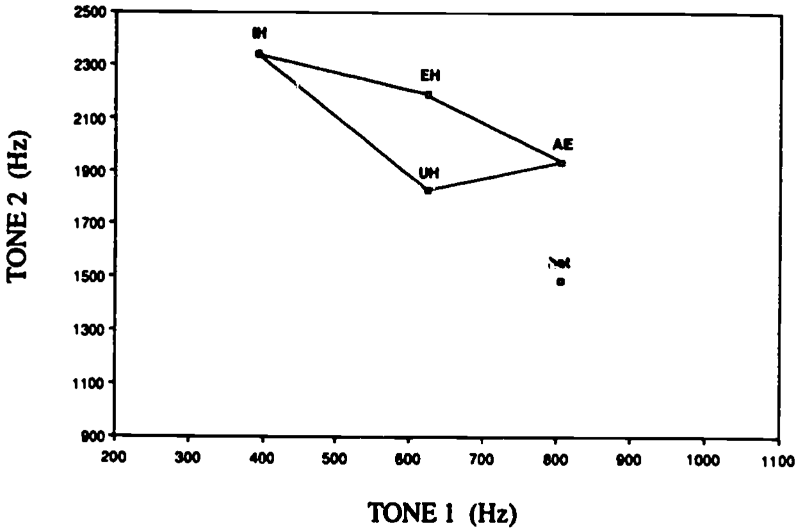


Figure 20. Sinewave nuclei: normalization portrayed as recalibration of perceptual vowel standards–effect of raising second formant tone by 30%. (Normalized standards=IH EH AE UH targets=bVt. The effect of the precursor sentence transposed by simultaneous lowering of Tone 1 and raising Tone 2 is not shown in this set of figures. This condition was effective perceptually neither in the present study nor in the study by Ladefoged & Broadbent, 1957.)

data partly obscures the evidence of normalization, or because the transpositions that we imposed on the precursor sentence were too moderate.

Lowering the frequency of the second tone of the vowel standards brings the *but* target into the EH category, depicted in Figure 19. This approximately matches what our listeners did. In the original synthetic version, however, Ladefoged and Broadbent's listeners had judged *but* as ambiguous with a similar frequency transposition of the sentence frame, half of the judgments for *but* and half for *bat*. From Figure 19, however, it is clear that this sinusoidal condition should project the *but* target into the EH category in the present case.

The precursor with Tone 2 raised was presented with the *bat* target only, in which case we found that listeners apparently could not identify the target, with the exception that they consistently rejected *but* as the target identity. This sentence, then, made the subjects perceive the *bat* target no more precisely than as "not *but*." Figure 20 reveals the normalized standards to be quite remote from the values of the *bat* target, although the exclusion of UH from the possibilities cannot be derived from this representation. Ladefoged and Broadbent found that this transposition of the precursor sentence had no effect on the identification of the *bat* target, although even in their natural context this target was less consistently identified than the other three, and, incidentally, less consistently identified than the sinewave *bat* in our study.

To summarize, the results of the sinewave test appear similar to the original synthetic precedent, and the discrepancies may be accounted for by much the same argument that Ladefoged and Broadbent employed to explain their observations. Essentially, the results obtained with synthetic speech and sinewave replicas agree. In consequence of this, the motivation of the explanation for the synthetic results is no less applicable to the sinusoidal results–that is, for the listeners who perceive sinewaves phonetically. Because this is evidence of normalization specific to the acoustic correlates of vocal dimensions, it is fair to say that sinewave replicas of speech signals and ordinary speech enjoy a common perceptual treatment.

Nonphonetic Influence on the Bat Target

The identification test had designated a group of listeners who exhibited nonphonetic perception of sinusoidal replicas of speech, and these subjects performed the normalization test in a manner largely indifferent to the range of the tone variation of the precursor sentences. However, they did exhibit significantly different distributions of identification judgments across the three precursors used with the *bat* target. The performance on this condition by nonphonetic listeners differed dramatically from that of the phonetic listeners, as Figures 10 and 14 show, which offers a hint at the cause of this outcome. The significant nonphonetic effect may be traced to one detail: The effect of raising Tone 2 in the precursor elicited consistent identification of the target as *bit*. This appears not to be an intrinsically phonetic effect and is completely at odds with the rationale that Ladefoged and Broadbent developed and that we have applied here. Raising Tone 2 places *bat* farthest from the IH standard of the four possible choices, as Figure 20 shows. It is the least likely identification, according to the phonetic rationale. Although the performance of this group attained statistical significance in this target condition, it is unlikely to have arisen due to selective susceptibility to normalization or to phonetic properties of sinusoidal signals when the vowel in the target is AE. This speculation is far from definitive, however, and much remains to be clarified about the performance of subjects who are unable to attain phonetic organization of sinusoidal replicas of speech.

Phonetic and Nonphonetic Listeners

Our prior investigations of the perceptual attributes of sinusoidal signals led us to expect that no less than a third of the subjects we encountered would be incapable of characterizing the signals phonetically. Alternative modes of perceptual organization appear to underlie this circumstance, one a phonetic organization that takes multiple tone variation to be a single fused pattern, and another that takes the tones each to be an independent stream. An account of this difference and the definition of the effective trigger for coherent organization will require additional study. In practice, however, this finding exemplifies an ongoing difficulty with tests of perceptual sensitivity to time- varying information. Because receptivity to sinusoidal messages is not easily predicted, we used the test of identification here to index each subject's capability. This use of a converging identification test separated two groups whose performance on the normalization test differed considerably and therefore appears to be an objective way of responding to the methodological challenge posed by the sinusoidal replication technique.

Other researchers who have conceptualized their sinewave results similarly, as the product of two (or more) perceptually distinct populations, have occasionally—and to good effect—used less formal post hoc practices to identify a subject's membership in one group or the other (Bailey et al., 1977; Best et al., 1981); our a priori approach here is surely not less suited to the problem. A theoretically richer account of individual differences in speech perception, including the issue of perceptual organization that concerns us here, may eventually be pursued with the sinusoidal replication technique, though it is beyond the scope of this study.

Conclusion

Because the outcome of our test using sinusoidal replicas of speech resembles the outcome of the test employing conventional synthetic speech, then one possible interpretation is that the two instances share a common perceptual treatment. We favor this conclusion and the implication that sinusoidal signals preserve information provided by coherent spectral changes in ordinary speech signals. However, an alternative to this perspective would be to view this experiment as a falsification of the classic study after which it was modeled. It could be argued that the replication of rescaling with nonspeech sounds invalidates the account given originally by Ladefoged and Broadbent, which alleged that the effect was based on the personal information available in vocal sound production. In fact, there have been challenges to the account given by Ladefoged and Broadbent pertaining to the role, or the sufficiency, of resonant spectral information in perceptual scaling and vowel perception (e.g., Ainsworth, 1974; Dechovitz, 1979b; Ladefoged, cited in Papçun, 1980; Syrdal & Gopal, 1985; Thompson & Hollien, 1970; Verbrugge et al., 1976). Although these reports proposed technical improvements, variations and extensions of the initial paradigm, warranting the expansion of the proposed mechanism for normalization, the rationale of normalization has proven to be quite durable (see Nearey, 1978).[2] Although the precise perceptual mechanism involved may be obscure, our finding looks like range-appropriate perceptual rescaling, and we therefore appeal to the conventional explanation–vocal-tract normalization.

Looking across the results of Experiment 1, we find that the sinewave outcomes approximated the findings of the synthetic speech version serving as our theoretical and procedural model. The identification of a target word was evidently affected by the variation of tones within the preceding sentence. This was formulated by Ladefoged and Broadbent as a contingency of phonetic perception on personal information about the talker, and we propose to extend this explanation to the present finding. To state the conclusion, the listener perceives sinewave replicas of speech by implicitly recognizing that the tonal complexes originate from a vocal source. Accordingly, sinewave signals appear to be handled by the ordinary means of speech perception, probably because such abstracted spectral patterns nonetheless preserve time- varying phonetic information present in natural signals.

Experiment 2

Control Condition about Vowel Perception

The first experiment of this report tested the assumption that sinusoidal signals are perceived phonetically, using a measure that reflected the contingency of vowel identification on perceptual organization. Evidently, sinusoidal patterns are identified to be vocal in origin, despite the absence of the acoustic products of vocalization in the signal. In consequence, perceptual normalization of vocal-tract dimensions influenced the identification of the vowels in the target syllables, as happens in instances of synthetic and natural speech.

A prominent component of accounts of vowel perception more generally is, simply, that the frequencies of the first and second formants together provide much of the information for vowels in English (reviewed by Ladefoged, 1967; Nearey, 1978). With few exceptions (e.g., Rubin, 1971; Shepard, 1972), these two dimensions of acoustic variation, first and second formant frequencies,

---

[2] Suomi (1984) proposes a model of vowel identification that essentially filters out vocal-tract scale variation, but that also ignores contingent effects of the kind that we re [illegible] here.

have offered a compelling characterization of the perceptually significant acoustic correlates of vowels, especially so because they are held to correspond to articulatory dimensions of sound production. Though we must take the anatomical designations loosely, it is as if the first formant were associated with the height of the tongue, high or low, and the second formant with the advancement of the tongue, front or back.

The sinewave syllables used as targets in Experiment 1 present spectral patterns derived from frequency variation of these acoustic sources of vowel information for the perceiver. Nevertheless, the possibility exists that listeners identified the sinewave vowels in a manner unlike ordinary vowel perception, due to the unspeechlike spectrum that such signals present to the ear. This is encouraged by reports that listeners can attribute speechlike qualities to simple acoustic signals. Of course, it is common to assign a phonetic label to a nonspeech impression in instances of onomatopoeia, though this is presumably distinct from ordinary speech perception, at least in the respect that nonspeech impressions of formant variation do not typically accompany phonetic perception. In the case of sinewave replicas, however, the nonspeech quality of the signal persists even when transcription of linguistic properties occurs (Remez et al., 1981). In order to conclude that the influence of precursor sentences on target identification, observed in Experiment 1, was evidence of vowel normalization, it will be necessary to establish more firmly that the perception of the vowels in the targets was an instance of vowel perception, differing from the attribution of similarity between the nonspeech auditory qualities the listener hears and the vowel features the listener knows.

Vocality

Psychoacoustic investigations of an earlier generation examined the basic sensory attributes of auditory experience, among which are found the familiar pitch, loudness, and timbre. Additionally, simple acoustic presentations were held to cause the experience of *vocality*, a kind of speechlike quality of sound that was irreducible to simpler experience (Boring, 1942; Gatewood, 1920; Köhler, 1910; Modell & Rich, 1915). Operationally, vocality was observed in studies in which the ordinary acoustic dimension of frequency was varied, but the subject responded with a vocal imitation or a phonetic segment name instead of reporting a pitch experience. For example, this sequence of sensations of vocality was held to occur with ascending frequency from 65 Hz to 33.6 kHz:[3] "vvv-mmm-U-O-A-E-I-sss-fff-ch" (Boring, 1942, page 374).

Since these pioneering efforts, the acoustic correlates of phonetic perception have been more realistically defined as complex spectra rather than as an accompaniment of pitch sensation. But for interpreting the present findings, the precedent of vocality studies is uniquely relevant. Specifically, if subjects are able to label nonspeech simple tones with vowel names (Fant, 1959; Modell & Rich, 1915), then the performance of subjects in Experiment 1 may have exploited this capacity for designating likeness rather than for perceiving vowels. Our prior research has shown that Tone 1 is especially prominent perceptually (Remez & Rubin, 1984), which suggests

---

[3] The value given for the upper frequency bound of vocality impressions reflects a compound error (Boring, 1942). Vocality was said to vary across the dimension of frequency, but contemporary estimates of frequency were erroneous. Because of mistaken calibration of the Galton whistles that were used as acoustic sources, the upper frequency bound of audibility had been overestimated at 50 kHz. In consequence, the frequency limits for impressions of vocality were set well within the range of audibility, as it was understood at the time.

that subjects may have analyzed this tone as an independent component and may have based the vowel responses on the phonetic likeness of this tone. If our subjects differentiated the four vowel targets on the basis of the distant phonetic likeness of simple nonspeech spectra, rather than on information given in the multiple spectral properties to which most accounts of vowel perception refer, then a strong claim about speech perception from coherent spectral variation is inappropriate. An experimental investigation of this possibility is clearly warranted.

Experiment 2, then, is a test to distinguish the assignment of vowel names to perceptually prominent tones from the perception of sinusoidal vowels. The objective is to determine whether the identification of the vowel of the three-tone target syllables may be accounted for by vowel labels applied to individual component tones. To the extent that listeners report the same vowel from three-tone and single-tone patterns, we may doubt that vowel perception of sinusoidal syllables involves the ordinary mechanism of speech perception. To the extent that identification of single tone and multiple tone patterns differ, we may conclude that vocality judgments are a different sort of perceptual phenomenon than speech perception from sinusoidal replicas of natural utterances.

Several conditions were used in this attempt to distinguish sinusoidal vowel perception from the attribution of phonetic qualities to nonspeech signals. In the first, we simply replicated the identification test of Experiment 1, in which the four three-tone targets were presented to phonetically disposed listeners. Then, an identification test was presented consisting solely of Tone 1 from each of the four targets. A third test determined the identification of Tone 2 from each target presented in isolation. Last, two conditions investigated the perceptual interaction of spectral peak frequency and gross spectral shape by presenting Tone 1 and Tone 2 synthesized as sawtooth waves rather than as sinewaves. As an ensemble of conditions, these tests define the acoustic correlates of vowel judgments in Experiment 1, revealing whether or not the phonetic likeness of nonspeech impressions plays a role in sinusoidal phenomena.

Method

Acoustic Test Materials

Our design here called for five sets of tonal patterns. The first consisted of the four three-tone [bVt] target syllables of Experiment 1. The second set was composed of the lowest component of each of the four syllables, the Tone 1 set. The third set contained the four isolated Tone 2 sinusoids. The fourth and fifth sets consisted of single tone patterns, realized as time- varying sawtooth waves, in one case following the frequency values of the four Tone 1 patterns, in the other, the patterns of the Tone 2 set. Sawtooth waves presented the same frequency peak in the acoustic spectrum but possessed a shallow roll-off of the high-frequency skirt of the spectrum envelope. Test orders were compiled on the VAX and then output to audiotape, as in the first experiment of this report, and were delivered to listeners binaurally over headsets in the manner of Experiment 1.

Procedure

Following a warm-up sequence of eight sinusoidally synthesized sentences were three brief tests. First, a 40-trial identification sequence was presented in which each of the three-tone syllables occurred 10 times each in a random order. On each trial, subjects marked a response form, circling the word *bit*, *bet*, *bat*, or *but* in a four-alternative forced choice paradigm.

Next, subjects were presented with an 80-trial test presenting the eight single-tone patterns of the Tone 1 and Tone 2 conditions. They occurred intermixed, 10 times each, in random order. Again, a subject chose one of the four alternative responses on each trial by circling the appropriate word in the test booklet.

Finally, a second 80-trial identification test, composed of the eight single sawtooth patterns, was presented. The eight patterns occurred 10 times each in random order. On each trial, a subject identified the word containing the vowel closest to the sound of the tone, marking the response form accordingly. The entire session, including the warm-up and three listening tests, lasted 45 min.

Subjects

Twenty-three volunteer listeners participated in this study. Each had been tested in Experiment 1, and each was a phonetic listener as designated by the identification test. All subjects were paid for their time. They were tested in groups of two to six at a time.

Results and Discussion

Four analyses of variance were performed to assess the effects of the spectrum conditions–whether identification varied across three-tone, single tone, and single sawtooth presentations–for each of the target vowels. The mean proportions of the responses across the five spectrum conditions are shown in Figures 21, 22, 23, and 24. To summarize the outcomes, the identification of the three-tone targets differed from the identification of the single-component tones, and again from the component tones realized as sawtooth waves. In several cases, the identification of a single sinusoid differed from the matched single sawtooth.

The analysis in the case of the *bit* target found that the interaction of the factors Spectrum (three-tone, Tone 1, Tone 2, Sawtooth 1, Sawtooth 2) and Errors (*bat*, *bet*, *but*: the unintended response alternatives) was highly significant, $F(8, 176) = 11.87, p < .0001$. Figure 21 presents this condition. The three-tone target differed significantly from each of the component conditions, determined by post hoc comparisons of the error means (Newman-Keuls). Both Tone 1 and Sawtooth 1 elicited more identifications as UH; Tone 2 drew more identifications as AE, and Sawtooth 2 more identifications as EH. This shows clearly that the identification of the *bit* target is not simply attributable to the vowel likenesses of its component tones, though both versions of the second component seem overall like the vowel IH despite the increase in unintended responses. There were no differences observed between Tone 1 and Sawtooth 1, or between Tone 2 and Sawtooth 2, indicating that the frequency of the amplitude peak in the spectrum, rather than the shape of the spectrum envelope, contributed to the judgment.

The results in the case of the *bet* target are shown in Figure 22. Again, the Spectrum × Errors interaction was significant, $F(8, 176) = 11.64, p < .0001$. The identification of the three-tone pattern differed from each of the single tone versions, with significantly more AE responses to Tone 2, Sawtooth 1, and Sawtooth 2; in the remaining comparison, Tone 1 was identified as UH quite often, also differing from the three-tone pattern. In this case, the imposition of the sawtooth wave on the pattern of Tone 1 proved to be perceptually distinctive, leading to more

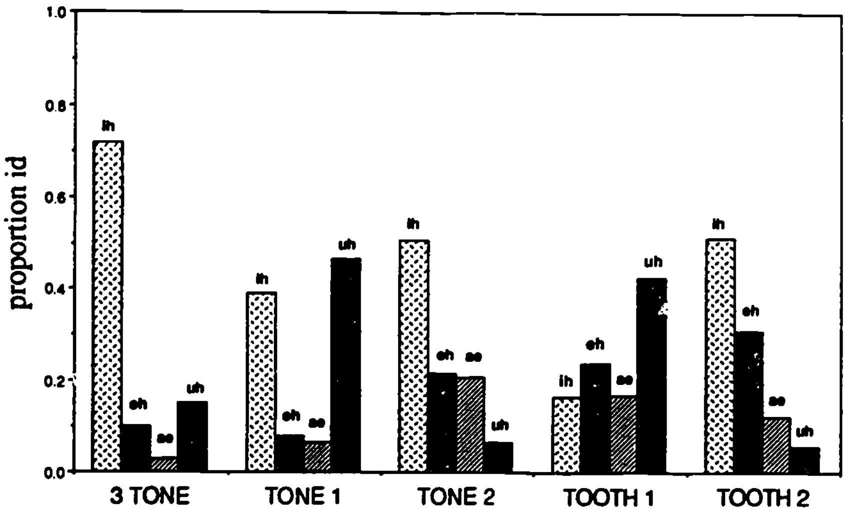


Figure 21. Response distributions for the sinusoidal targets realized as 3-tone patterns, Tone 1 and Tone 2 presented in isolation, and the two lowest tones synthesized as frequency modulated sawtooth waves, Sawtooth 1 and Sawtooth 2: This figure shows the results of the test with the *bit* target. (Proportion id=proportion of identification judgments.)

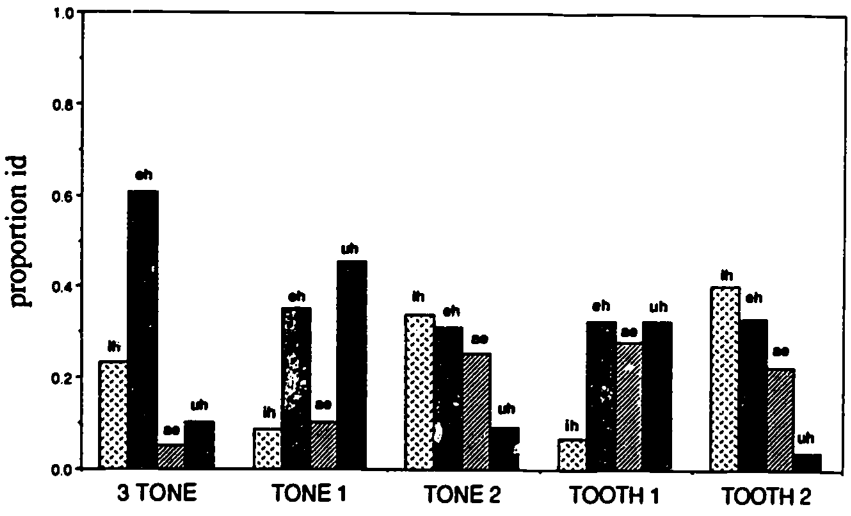


Figure 22. Response distributions for the sinusoidal targets realized as 3-tone patterns, Tone 1 and Tone 2 presented in isolation, and the two lowest tones synthesized as frequency modulated sawtooth waves, Sawtooth 1 and Sawtooth 2: This figure shows the results of the test with the *bet* target. (Proportion id=proportion of identification judgments.)

AE responses and fewer UH responses to Sawtooth 1 than to its mate, Tone 1. By implication,

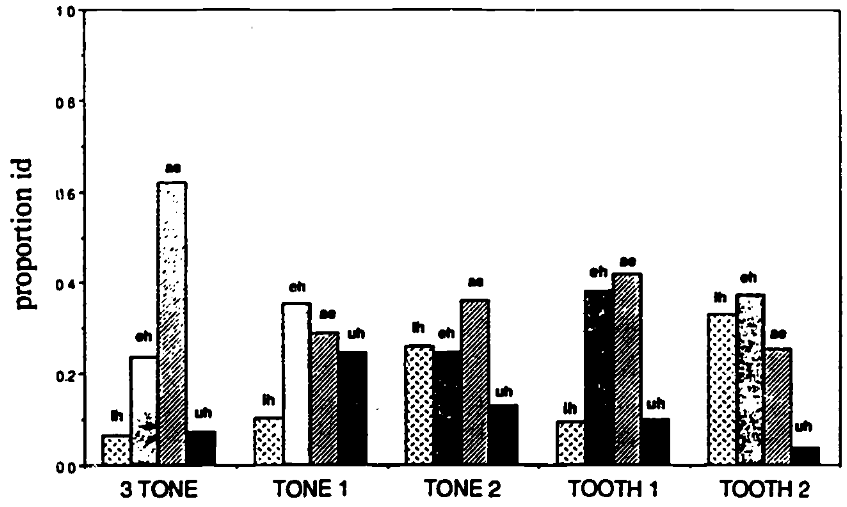


Figure 23. Response distributions for the sinusoidal targets realized as 3-tone patterns, Tone 1 and Tone 2 presented in isolation, and the two lowest tones synthesized as frequency modulated sawtooth waves, Sawtooth 1 and Sawtooth 2: This figure shows the results of the test with the *bat* target. (Proportion id=proportion of identification judgments.)

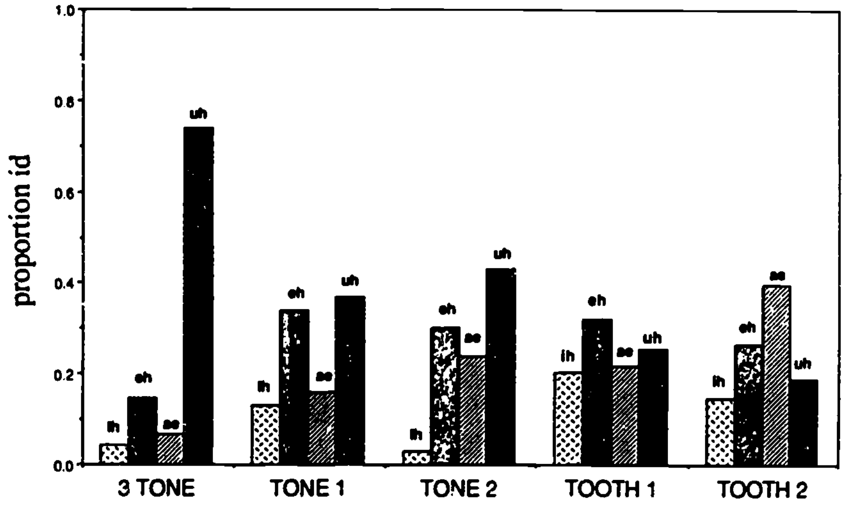


Figure 24. Response distributions for the sinusoidal targets realized as 3-tone patterns, Tone 1 and Tone 2 presented in isolation, and the two lowest tones synthesized as frequency modulated sawtooth waves, Sawtooth 1 and Sawtooth 2: This figure shows the results of the test with the *but* target. (Proportion id=proportion of identification judgments.)

listeners in this case took the spectrum envelope into account in designating a vowel impression for the lowest tone presented in isolation.

The results in the case of the *bat* target are parallel to *bit* and *bet*: the spectrum and errors interaction was significant, $F(8,176) = 4.90, p < .0001$. The three-tone identifications differed from each of the single tone conditions, as is shown in Figure 23. There were significantly more UH responses to Tone 1 and IH responses to Tone 2 and Sawtooth 2; Sawtooth 1 elicited more identifications as EH than did the three-tone pattern. The spectrum envelope also influenced the perception of the lowest tone, inasmuch as Tone 1 was reported as UH more frequently than was Sawtooth 1. These findings again indicate that the perception of the vowel in the three-tone *bat* target is not reducible to the effects of its component tones.

Finally, the *but* target results appear consistent with the others, the three-tone pattern differing from the single tones in each instance. Here, once again, the Spectrum × Errors interaction was significant, $F(8,176) = 2.45, p < .02$. Tone 1 and Sawtooth 1 both produced more EH responses, and Tone 2 and Sawtooth 2 more AE responses. Moreover, Tone 2 and Sawtooth 2 differed in the extent to which they evoked AE labels, which reflects the perceptual influence of spectrum envelope on the perceptual value of the frequency of the amplitude peak.

Conclusion

Overall, this test established that the reports of vowel identity of the three-tone targets do not match the reports for either of the two perceptually important components, Tone 1 and Tone 2, presented in isolation. Though the variation across the target conditions of the responses to Tone 2 approximates the three- tone case, the information that Tone 2 conveys about the vowel was much more effective in the acoustic context of Tone 1. Were this outcome to involve synthetic or edited natural speech signals, it would not be remarkable (Chistovich, Sheikin, & Lublinskaja, 1979; Delattre, Liberman, Cooper, & Gerstman, 1952; but, see Klatt, 1982, 1985). But, in view of the durability of the distinctly tonal qualities when three-tone complexes are perceived phonetically, this result provides a key to understanding the perceptual organization of our present target syllables. The attribution of vowel quality to the sinusoidal signals appears to use information simultaneously available from the two lowest tonal components. This finding distinguishes sinusoidal vowel perception from the attribution of speechlike qualities to spectrally simple nonspeech sounds (for a variant pertaining to perceptual learning, see Grunke & Pisoni, 1982). Although sinusoidal vowel perception appears to rely on acoustic information analogous to that which presumably occurs in ordinary speech signals, the imposition of vowel names on simple tones does not similarly depend on phonetic information. Assigning a vowel name to a simple nonspeech tone may truly require an assertion of remote resemblance between nonspeech and speech impressions, a perceptual circumstance distinct from the apprehension of phonetic information from an acoustic signal.

The comparison of sinusoid and sawtooth versions of the same frequency patterns produced both clear and puzzling results. On the one hand, subjects evidently are influenced by timbre when judging single tones, inasmuch as the frequency at which greatest power in the spectrum occurred was the same in matched sinusoidal and sawtooth instances. This result, noted in two cases involving the second tone and in one involving the first, had no consistent effect. Imposing a sawtooth waveform on the frequency pattern did not create the impression consistently of a lower or a higher vowel; nor did it consistently affect the vowel dimension of advancement; nor did it make the resulting vowel likeness seem consistently more central or less central. In the case of Tone 1 of *bat*, the imposition of the sawtooth waveform was tantamount to decreasing the frequency of the syllable nucleus; in the cases of Tone 1 of *bet* and Tone 2 of *but*, it was equivalent to increasing the frequency. The lack of a clear pattern to the results of this experimental manipulation requires no more than a speculation: Might the effect be specific to the task of judging single tones? Informally, we have observed that imposing sawtooth or triangle waveforms on three-tone patterns affects apparent naturalness, though not the intelligibility of the sentence. Whether the properties of the spectrum envelope create phonetic categorization effects more generally and whether the present outcome requires a phonetic or an auditory motivation remain, therefore, topics for further study.

General Discussion

Which properties of the acoustic speech signal provide information to the listener about the talker's message? Our studies of sinusoidal replicas of natural utterances offer a contrast to the perspective that the perceiver is a meticulous listener whose attention is devoted to the elemental attributes of speech signals. When these elements, which have often been presumed to be essential for perception, are removed from the signal and replaced by sinusoids, phonetic perception persists. Or so it seemed. The present experiments tested this proposition, that the listener transcribed sinusoidal sentences by virtue of ordinary perception, and in essence did not use an explicit strategy or rationalization. Because the perception of sinusoidal vowels in [bVt] syllables resembles vowel perception and because vowel identification of sinewave syllables involves the concurrent properties of the components of multitone patterns, it seems as though interconsonantal vowel perception with three-tone sinewave syllables is phonetically motivated. Moreover, because the perception of sinusoidal vowels was affected by the range of tone variation in precursor sentences, listeners may be said to organize three-tone replicas as if the signals were produced vocally. The frequency range effect is, in fact, a vocal-tract scale effect, though certainly the "vocal tract" that issues sinusoidal speech is an abstract one, indeed.

The factors that regulate the perceptual apprehension of phonetic information from time-varying patterns remain to be specified. A full account will also require (a) an exposition of the perceptual organization of the component tones as a single coherent stream, and (b) a means of relating these rather special listening conditions to the perceptual organization of natural speech. In the meanwhile, the finding that speech perception can endure the absence of short-time acoustic elements characteristic of vocal sound production places definite constraints on models of speech perception.

References

Ainsworth, W. A. (1974). The influence of precursive sequences on the perception of synthesized vowels. *Language and Speech, 17,* 103-109.

Bailey, P. J., & Summerfield, Q. (1980). Information in speech: Observations on the perception of [s]-stop clusters. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 536-563.

Bailey, P. J., Summerfield, A. Q., & Dorman, M. (1977). On the identification of sine-wave analogues of certain speech sounds. *Haskins Laboratories Status Report on Speech Research, SR-51/52,* 1-25.

Best, C. T., Morrongiello, B., & Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics, 29,* 191-211.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century.

Bricker, P. D., & Pruzansky, S. (1976). Speaker recognition. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 295-326). New York: Academic Press.

Chistovich, L. A., Sheikin, R. L., & Lublinskaja, V. V. (1979). "Centres of gravity" and spectral peaks as the determinants of vowel quality. In B. Lindblom & S. Öhman (Eds.), *Frontiers of speech communication research* (pp. 143-157). New York: Academic Press.

Cutting, J. E. (1974). Two left hemisphere mechanisms in speech perception. *Perception & Psychophysics, 16,* 601-612.

Darwin, C. J., & Baddeley, A. D. (1974). Acoustic memory and the perception of speech. *Cognitive Psychology, 6,* 41-60.

Dechovitz, D. R. (1977b). Information conveyed by vowels: A negative finding. *Journal of the Acoustical Society of America, 61,* S39.

Dechovitz, D. R. (1977a). Information conveyed by vowels: A confirmation. *Haskins Laboratories Status Report on Speech Research, SR-51/52,* 213-219.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America, 27,* 769-773.

Delattre, P. C., Liberman, A. M., Cooper, F. S., & Gerstman, L. J. (1952). An experimental study of the acoustic determinants of vowel color. *Word, 8,* 195-210.

Disner, S. F. (1980). Evaluation of vowel normalization procedures. *Journal of the Acoustical Society of America, 67,* 253-261.

Eimas, P. D., & Miller, J. L. (1978). Effects of selective adaptation on the perception of speech and visual patterns: Evidence for feature detectors. In R. D. Walk & H. L. Pick, Jr. (Eds.), *Perception and experience* (pp. 307-345). New York: Plenum.

Elman, J. L., & McClelland, J. L. (1985). Exploiting lawful variability in the speech waveform. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 360-385). Hillsdale, NJ: Erlbaum.

Fant, C. G. M. (1959). Acoustic analysis and synthesis of speech with applications to Swedish. *Ericsson Technics, 1,* 3-108.

Fant, C. G. M. (1962). Descriptive analysis of the acoustic aspects of speech. *Logos, 5,* 3-17.

Fant, C. G. M. (1966). A note on vocal tract size factors and nonuniform F-pattern scalings. *Speech Transmission Laboratory Quarterly Progress and Status Report 4/66*, (Pp. 22-30). Stockholm, Sweden: Royal Institute of Technology. (Reprinted in 1973 as Chapter 4, in *Speech sounds and features* (Pp. 84-93). Cambridge, MA: MIT Press.)

Fourcin, A. J. (1968). Speech-source inference. *IEEE Transactions on Audio and Electroacoustics, ACC-16,* 65-67.

Gatewood, E. L. (1920). The vocality of fork, violin and piano tones. *American Journal of Psychology, 31,* 194-203.

Gerstman, L. (1968). Classification of self-normalized vowels. *IEEE Transactions on Audio and Electroacoustics, ACC-16,* 78-80.

Grunke, M. E., & Pisoni, D. B. (1982). Some experiments on perceptual learning of mirror-image acoustic patterns. *Perception & Psychophysics, 31,* 210-218.

Jenkins, J. J., Strange, W., & Edman, T. R. (1983). Identification of vowels in "vowelless" syllables. *Perception & Psychophysics, 34,* 441-450.

Joos, M. (1948). Acoustic phonetics. *Language, 24* (Supplement), 1-137.

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics, 7,* 279-312.

Klatt, D. H. (1982). Prediction of perceived phonetic distance from critical-band spectra: A first step. *Proceedings of the ICASSP, Paris,* 1278-1281.

Klatt, D. H. (1985). A shift in formant frequencies is not the same as a shift in the center of gravity of a multiformant energy concentration. *Journal of the Acoustical Society of America, 77,* S7.

Köhler, W. (1910). Akustische Untersuchungen, II. *Zeitschrift für Psychologie mit Zeitschrift für Angewandte Psychologie und Charakter-kunde, 58,* 59-140.

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America, 66,* 1668-1679.

Ladefoged, P. (1967). The nature of vowel quality. In *Three areas of experimental phonetics* (pp. 50-142). London: Oxford University Press.

Ladefoged, P. (1980). What are linguistic sounds made of? *Language, 56,* 485-502.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America, 29,* 98-104.

Liberman, A. M. (1970). The grammars of speech and language. *Cognitive Psychology, 1,* 301-323.

Lieberman, P. (1984). *The biology and evolution of language.* Cambridge, MA: Harvard University Press.

Markel, J. D., & Gray, A. H., Jr. (1976). *Linear prediction of speech.* New York: Springer Verlag.

Massaro, D. W. (1972). Perceptual images, processing time, and perceptual units in auditory perception. *Psychological Review, 79,* 124-145.

Modell, J. D., & Rich, G. J. (1915). A preliminary study of vowel qualities. *American Journal of Psychology, 26,* 453-456.

Nearey, T. M. (1978). *Phonetic feature systems for vowels.* Bloomington, IN: Indiana University Linguistics Club.

Papçun, G. (1980). How do different speakers say the same vowels? Discriminant analyses of four imitation dialects. *UCLA Working Papers in Phonetics.*

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24,* 175-184.

Rand, T. C. (1971). Vocal tract size normalization in the perception of stop consonants. *Haskins Laboratories Status Reports on Speech Research, SR-25/26,* 141-146.

Remez, R. E. (in press). Units of organization and analysis in the perception of speech. In M. E. H. Schouten (Ed.), *Psychophysics of speech perception.* Amsterdam: Martinius Nijhoff.

Remez, R. E., & Rubin, P. E. (1984). On the perception of intonation from sinusoidal sentences. *Perception & Psychophysics, 35,* 429-440.

Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science, 212,* 947-950.

Rubin, P. E. (1971). *Psychological space of vowel perception.* Unpublished senior thesis, Brandeis University, Waltham, Massachusetts.

Rubin, P. E. (1980). *Sinewave synthesis.* Internal memorandum, Haskins Laboratories, New Haven, Connecticut.

Samuel, A. G. (1982). Phonetic prototypes. *Perception & Psychophysics, 31,* 307-314.

Shankweiler, D., Strange, W., & Verbrugge, R. R. (1977). Speech and the problem of perceptual constancy. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing.* Hillsdale, NJ: Erlbaum.

Shepard, R. N. (1972). Psychological representation of speech sounds. In E. E. David, Jr., & P. B. Denes (Eds.), *Human communication: A unified view* (pp. 67-113). New York: McGraw-Hill.

Stevens, K. N., & Blumstein, S. E. (1981). The search for invariant acoustic correlates of phonetic features. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives in the study of speech* (pp. 1-38). Hillsdale, NJ: Erlbaum.

Summerfield, A. Q., & Haggard, M. P. (1973). Vocal tract normalization as demonstrated by reaction times. *Speech Perception: Report on Work in Progress* (Queen's University of Belfast), *2:3*, 1-26.

Suomi, K. (1984). On talker and phoneme information conveyed by vowels: A whole spectrum approach to the normalization problem. *Speech Communication, 3,* 199-209.

Sussman, H. M. (1986). A neuronal model of vowel normalization and representation. *Brain and Language, 28,* 12-33.

Syrdal, A. K., & Gopal, H. S. (1985). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America, 79,* 1086-1100.

Thompson, C. L., & Hollien, H. (1970). Some contextual effects on the perception of synthetic vowels. *Language and Speech, 13,* 1-13.

Verbrugge, R. R., Strange, W., Shankweiler, D. P., & Edman, T. R. (1976). What information enables a listener to map a talker's vowel space? *Journal of the Acoustical Society of America, 60,* 198-212.

Wakita, H. (1977). Normalization of vowels by vocal-tract length and its application to vowel identification. *IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-25*, 183-192.

Williams, D. R., Verbrugge, R. R., & Studdert-Kennedy, M. (1983). Judging sine wave stimuli as speech and as nonspeech. *Journal of the Acoustical Society of America, 74,* S36.

Zue, V. W., & Schwartz, R. M. (1980). Acoustic processing and phonetic analysis. In W. A. Lea (Ed.), *Trends in speech recognition (pp. 101-124). Englewood Cliffs, NJ: Prentice-Hall.*

THE STOP-GLIDE DISTINCTION: ACOUSTIC ANALYSIS AND PERCEPTUAL EFFECT OF VARIATION ON SYLLABLE AMPLITUDE ENVELOPE FOR INITIAL /b/ and /w/*

Susan Nittrouer and Michael Studdert-Kennedy†

Abstract. *Amplitude change at consonantal release has been proposed as an invariant acoustic property distinguishing between the classes of stops and glides (Mack & Blumstein, 1983). Following procedures of Mack and Blumstein, we measured the amplitude change in the vicinity of the consonantal release for two speakers. The results for one speaker matched those of Mack and Blumstein, while those for the second speaker showed some differences. In a subsequent experiment, we tested the hypothesis that a difference in amplitude change serves as an invariant perceptual cue for distinguishing between continuants and noncontinuants, and more specifically, as a critical cue for identifying stops and glides (Shinn & Blumstein, 1984). Interchanging the amplitude envelopes of natural /bV/ and /wV/ syllables containing the same vowel had little effect on perception: 97% of all syllables were identified as originally produced. Thus, although amplitude change in the vicinity of consonantal release may distinguish acoustically between stops and glides with some consistency, the change is not fully invariant, and certainly does not seem to be a critical perceptual cue in natural speech.*

Introduction

Much research in speech perception has focused on discovering acoustic properties invariantly associated with phonological categories. For the most part, this work has not been successful. Rather, it has tended to show that the acoustic attributes associated with the perception of a particular category vary with phonetic context. However, there are exceptions to this trend. For example, Stevens and Blumstein (1978) found that the spectral tilt of the aperiodic energy

---

* *Journal of the Acoustical Society of America*, 1986, *80*, 1026-1029.

† Queens College and The Graduate School, The City University of New York

Aknowledgment. This research was supported by NICHD Grant HD-01994 to Haskins Laboratories, by NIH Grant NS-07237 to the first author through Haskins Laboratories, and by a Spencer Fellowship to the second author through The Center for Advanced Study in the Behavioral Sciences. We are grateful to D. H. Whalen and Philip Rubin for their help with programming, to Mindy Sirlin and her students at Iona College for serving as listeners, and to Bruno Repp, Richard McGowan, Sheila Blumstein, and an anonymous reviewer for their comments on earlier drafts.

immediately following syllable-initial stop release was invariantly related to the perception of place of articulation. Kewley-Port (1983) also reported invariant features distinguishing among places of articulation for syllable-initial stop consonants; these features were dynamic in nature.

Further, Mack and Blumstein (1983) reported a possible invariant cue for the stop-glide manner distinction: the ratio of the rms energy of a brief acoustic segment immediately following "release offset" to the rms energy of a brief acoustic segment at "release onset."[1] This ratio is large for stops and small for glides. In other words, stops have large and rapid increases in amplitude in the vicinity of consonantal release, while glides exhibit no such rapid amplitude change. The amplitude change itself largely results according to the acoustic theory of speech production from the narrowing of the pharyngeal cavity as the jaw lowers, and the consequent, correlated rise in frequency and amplitude of the first formant. The amplitude change in the first formant also, of course, spreads to the second and higher formants (Fant, 1960). In fact, Mack and Blumstein attribute their success in identifying an apparent invariant acoustic property to their use of a measure encompassing the whole spectrum of natural speech.

As a possible explanation for why earlier perceptual experiments failed to establish invariant relations between the acoustical signal and linguistic description, Mack and Blumstein (1983, pp. 1739-1740) suggest:

> "...in their search for a particular attribute of the signal which categorizes the stop-glide phonetic contrast, researchers have typically used highly stylized and schematized test stimuli to the extent that these stimuli have shared all attributes save the particular dimension in question. The use of such stimuli may involve failing to manipulate or ignoring critical acoustic attributes for stops and glides present in natural speech."

While this criticism may be reasonable, it should be stressed that discovering an invariant acoustic property is not the same as demonstrating that the property has an invariant perceptual effect. If it were perceptual studies would be supererogatory, and we would confine our attention to acoustic analysis. Mack and Blumstein's (1983) results therefore offer strong support for the well-known fact that differences in amplitude change in the vicinity of consonantal release are invariantly associated with the different productions of stops and glides, but do not address the question of whether these amplitude patterns invariantly lead to the corresponding percepts.

In an attempt to address this latter question, Shinn and Blumstein (1984) conducted a perceptual experiment using stimuli synthesized to approximate the natural tokens measured by Mack and Blumstein (1983). Two separate continua were constructed, one with formant values

---

[1] Mack and Blumstein computed the rms energy for 15-ms segments, except at stop releases where 5-ms segments were used. Relevant segments were defined as follows:

stops: (1) release onset: that point, clearly visible in the waveform, where aperiodic noise occurs; (2) release offset: that point, also visible in the waveform, where glottal pulsing begins.

glides: (1) release onset: the portion of the utterance characterized by a "visually perceptible increase in waveform amplitude and/or complexity" (Mack and Blumstein, 1983, p. 1741); (2) release offset: the point where F2 appears as distinct from F1 in the LPC analysis.

The use of different definitions for stops and glides raises problems that we have ignored. We simply followed the procedures of Mack and Blumstein.

appropriate for /i/, and one with formant values appropriate for /a/. Three amplitude envelopes were given to both the /bi-wi/ and the /ba-wa/ continua. For one set, the amplitude envelope varied systematically between the /b/ end and /w/ end of the continua, according to values derived from Mack and Blumstein (1983). In a second set, all tokens were given a /b/ amplitude contour, and, in the third set, all tokens were given a /w/ contour. Combined results of forced choice and free identification tasks led to the general conclusion that the amplitude envelope invariantly distinguishes between continuants and noncontinuants, and is a critical perceptual cue for the specific distinction between stops and glides. This conclusion was summarized by the observation that the amplitude envelope was "...sufficiently strong to override the formant frequency, rate, and duration cues..." (Shinn & Blumstein, 1984, p. 1249).

One potential problem with this study is that, although the manipulated acoustic attributes of the stimuli were derived from analysis of natural speech, they were nonetheless synthetic. Whether they could be accurately characterized as "highly stylized and schematized" (Mack & Blumstein, 1983, p. 1739) is debatable, but clearly they were not natural. Accordingly, the possibility exists that some critical attribute for stops and glides present in natural speech may have been ignored (cf. Mack & Blumstein, 1983, p. 1740).

We therefore undertook a small study to determine, first, if we could replicate the acoustic measurements of Mack and Blumstein (1983), and second, to see what perceptual effect appropriate manipulation of the amplitude contours would have in natural samples.

Acoustic Analysis

Two male speakers were recorded reading CV syllables consisting of either /b/ or /w/ followed by one of the vowels /i,e,a,o,u/ using a TEAC recorder and Sennheiser microphone. Five samples of each syllable were obtained, yielding 50 tokens per speaker (25 each of /b/ and /w/). Tokens were digitized on a DEC Vax computer using a 10-kHz sampling rate and a 4.9-kHz low-pass filter setting. Release onset and offset were identified using the criteria described by Mack and Blumstein (1983). Amplitude ratios between rms values obtained immediately after release offset and at release onset were computed for each syllable. Table 1 reports mean values for samples for the two speakers in this study (column 3, "Original") as well as mean values and ranges obtained by Mack and Blumstein for four speakers (column 2).

It can be seen that, in large part, we replicated Mack and Blumstein's findings. All /b/ amplitude ratio means are above the criterial value of 1.37 proposed by Mack and Blumstein, and all but two /w/ amplitude ratio means are below the criterial value of 1.36. More precisely, 96% of speaker 1's tokens and 66% of speaker 2's tokens were correctly discriminated by Mack and Blumstein's metric. (Percentages ranged between 86% and 98% for individual speakers in their experiment.) Thus, although our second speaker does not exhibit perfect invariance, [2] the data

---

[2] The term "invariance" may be used slightly differently in this paper and in that of Mack and Blumstein. We have used the term in its literal sense to refer to something that "...remains the same irrespective of the context, something that is free of coarticulatory effects..." (Suomi, 1985, p. 268). This may not be the sense applied to the term by Mack and Blumstein. (See Suomi, 1985, for a complete discussion.)

confirm the fact that stops generally exhibit large amplitude increases in the region of consonantal release while glides generally exhibit little or no amplitude increase in this region.

Table 1

Mean ratios of rms energy at release offset to rms energy at release onset. Data (means and ranges of four speakers) from Mack and Blumstein (1983), and from the present study (means of five tokens from each of two speakers). "Original" refers to mean values of original syllables. "Modified" refers to syllables with amplitude envelopes transposed to match those of appropriate templates.

		Present study	
Syllable	Mack and Blumstein	Original	Modified
/bi/	4.28	6.62	1.06
	(3.39-5.98)	3.76	1.35
/be/	4.22	5.54	1.05
	(2.62-6.02)	2.63	1.28
/ba/	3.47	7.62	1.14
	(1.70-6.93)	2.03	0.86
/bo/	3.77	5.48	1.01
	(2.15-5.33)	2.54	1.31
/bu/	2.49	6.77	1.10
	(1.37-3.21)	1.45	0.97
/wi/	1.12	1.24	7.07
	(1.05-1.16)	1.64	5.35
/we/	1.16	1.00	5.64
	(1.09-1.31)	1.31	3.30
/wa/	1.13	1.09	10.70
	(1.06-1.20)	1.11	7.46
/wo/	1.20	1.31	9.47
	(1.06-1.32)	1.00	1.96
/wu/	1.10	1.12	4.47
	(1.05-1.11)	1.70	2.48

Perceptual Experiment

The general method used for the perceptual study was based on the hypothesis that if the amplitude envelope in the vicinity of the release is an invariant cue to either the continuant-noncontinuant or the stop-glide distinction, then changing this characteristic should have a pronounced effect on perception. Specifically, if stops and glides (or stops and continuants) are distinguished by the amplitude contours described by Mack and Blumstein, then transposition of a stop-vowel amplitude contour onto a glide-vowel syllable should shift the listeners' judgments

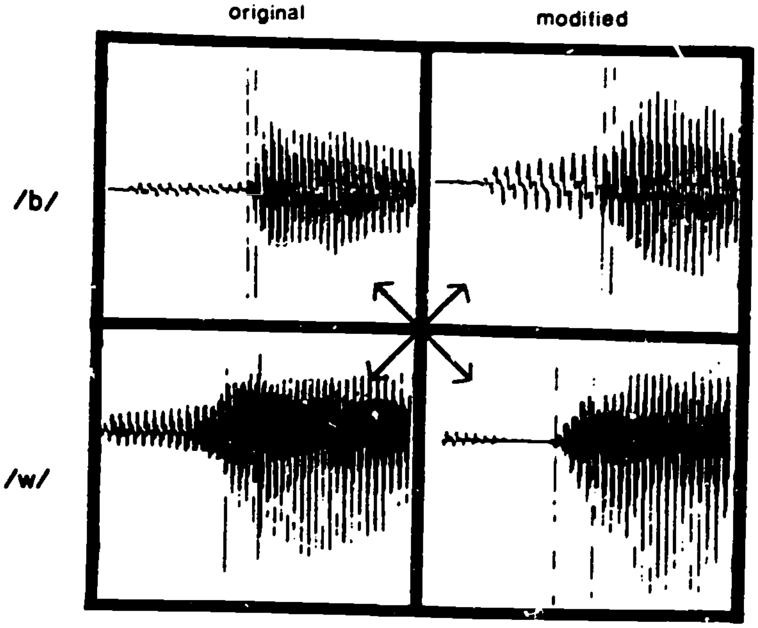


Figure 1. Samples of a /bV/ and /wV/ syllable produced by speaker 2, both before and after amplitude modification. Vertical lines indicate release onset and release offset, as defined by Mack and Blumstein.

from glide to stop, and vice versa. To accomplish this manipulation, an amplitude-matching program was used to modify the amplitude contour of a syllable to match that of a template. Tokens to serve as templates were selected by picking syllables closest to the mean ratio values for a given context. For example, the /bi/ syllable for speaker 1 with an amplitude ratio closest to the mean of the speaker's five /bi/ tokens was the template imposed on all /wi/ syllables for speaker 1. The two exceptions to this method were the glide templates for speaker 2's /bi/ and /bu/ tokens: Since the mean amplitudes of the /wi/ and /wu/ syllables for this speaker exceeded Mack and Blumstein's criterial value (see Table 1), the glide-vowel syllables with the lowest amplitude ratios were used as templates in these cases. The template and the syllable to be modified were aligned at release onset to preserve the critical amplitude characteristics in the vicinity of the release. Figure 1 displays samples of a /bV/ and /wV/ syllable produced by speaker 2 both before and after amplitude modification. The vertical lines indicate points of release onset and release offset, as defined by Mack and Blumstein. Comparing samples diagonally illustrates that modified tokens had amplitude envelopes similar to original tokens of the alternative type. Modified tokens were the length of either the template or the original syllable, whichever was shorter. This manipulation did not create any systematic variation in length because there was no systematic length variation in the original tokens.

To check the effectiveness of the amplitude transposition, amplitude ratios were computed for the modified tokens using the spectrally defined acoustic segments of the original syllables.[3] Mean ratios for modified syllables are given in column 4 of Table 1. In all cases, modified /b/ tokens exhibited ratios appropriate for glides (i.e., <1.36), and modified /w/ tokens exhibited ratios appropriate for stops (i.e., >1.37).

A test tape was made consisting of all the original and modified tokens in random order. Using a free identification task (Shinn & Blumstein, 1984), 19 undergraduate students listened to the tape under headphones and wrote the consonant they heard at the beginning of each syllable. The total number of responses for each syllable type—e.g., original /b/ syllables—was 2 speakers × 25 tokens × 19 listeners=950. A response was scored "correct" if it matched the consonant of the original syllable. Therefore, if amplitude contour was an invariant cue to the stop/glide distinction, a 100% *error* rate would be predicted for modified syllables. Results are listed in Table 2 as percentages of total responses identified as the original syllable.

Table 2

Percentages of responses ($N = 950$) given as *original* syllable, by 19 subjects.

	/b/		/w/	
	Original	Modified	Original	Modified
/i/	100.0	100.0	99.5	97.9
/e/	100.0	98.9	97.4	93.2
/a/	98.9	99.5	95.3	81.6
/o/	99.5	98.9	99.5	90.0
/u/	98.9	99.5	100.0	91.1
$\bar{X}$	99.5	99.4	98.3	90.8

These results indicate that, overall, amplitude transposition had little effect on perception. The mean score across all subjects and tokens was 97% correct recognition. Modifying /b/ syllables to match the amplitude contour of glides had essentially no effect: Five errors were made to original /b/ syllables, and six were made to modified /b/ syllables. The /w/ syllables were identified less accurately overall. Sixteen errors were made to original /w/ syllables (2% of the 950 responses). All errors involved the identification of the consonant as /b/, errors were evenly divided between speakers, and all but two of these errors were made to syllables ending with /a/ or /e/. Thus these 16 errors cannot be attributed to speaker 2's /wi/ and /wu/ tokens, which demonstrated amplitude ratios above the 1.36 criterial value. The highest error rate was on

---

[3] Because the acoustic segments used in computation of amplitude ratios for the original syllables were spectrally defined, it was impossible to *redefine* these segments for the modified syllables in exactly the same way. Modified syllables maintained the spectral characteristics of the original syllables. However, if amplitude ratios were computed for modified syllables using the same absolute time locations as the templates, these ratios obviously would be the same as those of the templates. Thus, no matter how amplitude ratios were computed, they met the criterion of Mack and Blumstein.

the modified /w/ syllables with 88 incorrect responses (9% of total). Of these 88 errors, 35 (40%) occurred for the vowel /a/ and another 36 (41%) were evenly spread between the back vowels /o/ and /u/. Seventy-nine of these errors to modified /w/ sylalbles (90% of the 88 incorrect responses) consisted of /b/ responses. The other nine errors were spread between /h/ responses and "no consonant heard."

It might be suggested that modifying the amplitude envelopes as we did failed to provide critical information in the glide prevoicing and, therefore, was not an appropriate test of the hypothesis. That is, because speaker 1 did not prevoice his /b/ syllables when these tokens were given /w/ contours, they lacked the prevoicing information normally present in glides; conversely, when his /w/ tokens were given /b/ contours, prevoicing information was destroyed. However, this suggestion can be refuted on both theoretical and empirical grounds. Theoretically, Mack and Blumstein's hypothesis contains no provision for glide prevoicing. Empirically, if this concern were valid, then speaker 1's modified /w/ tokens should strongly bias responses toward "b" since they contain no prevoicing, and speaker 2's modified /w/ tokens should bias responses toward "w" since they retain prevoicing (speaker 2 prevoiced both /b/ and /w/ in his original utterances). In fact, 10% of responses to speaker 1's modified /w/ tokens and 7% of responses to speaker 2's modified /w/ tokens were "b." Thus there seems to be little difference in responses as a function of prevoicing.

Discussion

To a great extent, the acoustic results reported here replicate Mack and Blumstein's (1983): Stop-vowel syllables generally demonstrated large amplitude changes in the vicinity of consonantal release, while glide-vowel syllables demonstrated little or no amplitude change. However, the pattern was not fully invariant for one of two speakers.

In the perceptual experiment, the results were very different from those of Shinn and Blumstein (1984). In the present study, the amplitude envelope certainly did not override other perceptually relevant acoustic attributes of the signal. Consequently, perception of both the continuant-noncontinuant and the stop-glide distinction was affected only slightly by modifying the amplitude envelope. Perhaps the different results of the two studies reflect their different stimulus materials. While we manipulated natural tokens, Shinn and Blumstein used synthetic stimuli and this may have "...involve(d) failing manipulate or ignoring critical attributes for stops and glides present in natural speech" (Mack & Blumstein, 1983, p. 1740). As to what these attributes may be, the present study is silent. However, much previous work (e.g., Liberman, Delattre, Gerstman, & Cooper, 1956; Miller & Baer, 1983; Miller & Liberman, 1979; Pisoni, Carrell, & Gans, 1983; Schwab, Sawusch, & Nusbaum, 1981) suggests that critical information for the syllable-initial stop-glide (/b/ vs. /w/) distinction is carried by the rate and extent of formant frequency change over the first 10 to 100 ms. In general, these results support the conclusion of Shinn et al. (1985): Many experimental effects disappear when stimuli are made to resemble natural speech. Thus it would seem that natural speech provides the listener with a complex of information for making phonological decisions, any one of which may be severely degraded or inappropriately manipulated without seriously affecting perception.

References

Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America, 73*, 322-335.

Liberman, A. M., Delattre, P. C., Gerstman, L. J., & Cooper, F. S. (1956). Tempo of frequency change as a cue for distinguishing classes of speech sounds. *Journal of Experimental Psychology, 52*, 127-137.

Mack, M., & Blumstein, S. E. (1983). Further evidence of acoustic invariance in speech production: The stop-glide contrast. *Journal of the Acoustical Society of America, 73*, 1739-1750.

Miller, J. L., & Baer, T. (1983). Some effects of speaking rate on the production of /b/ and /w/. *Journal of the Acoustical Society of America, 73*, 1751-1755.

Miller, J. L., & Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonants and semi-vowels. *Perception & Psychophysics, 25*, 457-465.

Pisoni, D. B., Carrell, T. D., & Gans, S. J. (1983). Perception of the duration of rapid spectrum changes in speech and nonspeech signals. *Perception & Psychophysics, 34*, 314-322.

Schwab, E. C., Sawusch, J. R., & Nusbaum, H. C. (1981). The role of second formant transitions in the stop-semivowel distinction. *Perception & Psychophysics, 29*, 121-128.

Shinn, P., & Blumstein, S. E. (1984). On the role of the amplitude envelope for the perception [b] and [w]: Further support for a theory of acoustic invariance. *Journal of the Acoustical Society of America, 75*, 1243-1252.

Shinn, P., Blumstein, S. E., & Jongman, A. (1985). Limitations of context conditioned effects in the perception of [b] and [w]. *Perception & Psychophysics, 38*, 397-407.

Stevens, K., and Blumstein, S. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America, 64*, 1358-1368.

Suomi, K. (1985). The vowel-dependence of gross spectral cues to place of articulation of stop consonants in CV syllables. *Journal of Phonetics, 13*, 267-285.

THE DEVELOPMENT OF LANGUAGE STRUCTURE IN CHILDREN WITH DOWN SYNDROME*

Anne E. Fowler

I. INTRODUCTION

Studying children with Down syndrome (DS), Lenneberg (1967) would argue, should be a wonderful way to study language development. Rather than tracking the rapid progress of non-handicapped children through developmental stages, one could choose a level of interest and examine it at a leisurely pace, for exactly the same effect. Despite substantial delays in the onset and timing of language development, the language structures that are acquired by children with DS have standardly been described as normal, unremarkable in the sequence of their development, and with no evidence of deviant forms not observed at some point in normal development (Bloom & Lahey, 1978; Evans & Hampson, 1968; Rondal, 1975; Rosenberg, 1982; Ryan, 1975).

The evidence presented here does not challenge this standard view; like prior investigators we find no evidence for aberrant constructions in children with DS. The aim here is to move beyond the descriptive account of "delay without deviance" towards a better understanding of the nature and extent of the language delay associated with Down syndrome. In this paper, I argue that much of the confusion and heterogeneity generated in this early research can be—indeed has been—alleviated by a more sophisticated developmental approach to designing and interpreting language studies. Recent studies have begun to yield consistent and coherent conclusions regarding the course of language learning in children with DS by acknowledging or removing the variability

---

* To appear in D. Cicchetti and M. Beeghly (Eds.), *Down syndrome: The developmental perspective*. New York: Cambridge University Press.

Aknowledgment. This research was supported by a Program Project Grant (NICHHD 1 PO-HD-10965) to D. Premack, R. Gelman, and D. Kemler and by a grant (12-113) from the March of Dimes Birth Defects Foundation to R. Gelman, L. Gleitman, and A. Fowler. An NICHD post-doctoral fellowship (1-F32-HD-06543) funded the author in the writing of this paper. The data presented here derive from a master's thesis submitted to the University of Pennsylvania. Portions of these data were presented at the Boston University Conference on Child Language Development (Fowler, Gelman, & Gleitman, 1980) and at the Society for Research in Child Development (Fowler, Gleitman, & Gelman, 1981). I am grateful to Rochel Gelman, Lila Gleitman, and members of their research labs for their collaboration on all parts of the study; the assistance of Renee Baillargeon is acknowledged in particular. I thank the administrators, teachers, children and parents of St. Katherine's Day School whose cooperation made this work possible. Lila Gleitman, Susan Brady, Rochel Gelman, and Marjorie Beeghly provided beneficial comments on earlier drafts of this paper; these were much appreciated. M. Loukides assisted ably with data reduction and analysis.

contributed by chronological age (CA), by the particular component of language being studied and by the language level at which the child is functioning—factors hypothesized to be important in normal language development. By teasing apart these factors, which have all too often been grouped together in the past, a developmental approach to the specific case of Down syndrome not only provides us with a much sharper descriptive picture than any alternative approach, but also allows us to move beyond description to shed light on the process of learning language in nonhandicapped populations.

II. THE "STANDARD VIEW"

To set the stage for the discussion of language in children with DS, I will first briefly review the general conclusions derived from the voluminous literature concerning the effects of mental retardation (MR) on language learning, without regard to etiology. Despite the enormous variability inherent in these studies in terms of subject population, experimental procedures, and level of language under study, there is much agreement that MR language is similar in kind to that of normals, differing only in a delay at least commensurate with the degree of retardation. (See Cromer, 1974; Rondal, 1975; Rosenberg, 1982; Ryan, 1975, and Yoder & Miller, 1972, for extensive reviews.)

A review of the literature on spontaneous usage suggests that MR children use the same structures as younger non-MR children and even talk about the same things with the same basic vocabulary. They appear to learn language facts in the same order as non-MR children and, as has been noted in virtually every published study, the order of difficulty on any given set of task items parallels that observed normally. For example, inasmuch as syntactic complexity does arise in MR populations, it appears to follow a pattern of increasing difficulty mimicking the normal course of syntactic development. (See Lackner, 1968, for a particularly influential study in this regard.)

In studies of comprehension, as in spontaneous production, the order of difficulty posed by various tasks and grammatical constructions remains comparable for MR and non-MR children (Graham & Gulliford, 1968; Lackner, 1968; Lovell & Dixon, 1967). Indeed, two quite sophisticated studies suggest that MR children rely on strategies for comprehending sentences that are directly analogous to strategies evident in early stages of language development (Dewart, 1979; Duchan & Erickson, 1976).

In experimentally-controlled elicitation tasks in which children are asked to produce or repeat a target utterance, MR children and non-MR children make similar kinds of errors, reductions, and deletions and show comparable effects of the grammatical structures being manipulated (Anastasiow & Stayrook, 1973; Berry, 1976; Rondal, Lambert & Sohier, 1981). For example, both groups find it easier to imitate simple active affirmative sentences than sentences that are passive, negative, or embedded (Berry & Foxen, 1975). Indeed, even in highly self-conscious morphological cloze tasks ("this is a dish, here are two -----," Berko, 1958), which prove very difficult in absolute terms, both the order of difficulty and the error patterns observed in MR children fit expectations derived from normal developmental research (Dever & Gardner, 1970; Newfield & Schlanger, 1968).

Finally, looking across tasks also reveals a similar order of difficulty across groups of children. Both MR children (Graham & Gulliford, 1968; Lovell & Dixon, 1967) and non-MR children

(Fraser, Bellugi, & Brown, 1963) find it easier to repeat a construction verbatim than to produce it in a controlled elicitation procedure.

In sum, in studies looking at qualitative features of MR language, it is of great significance that there is no positive evidence for novel grammatical systems or comprehension strategies defying the normal pattern of acquisition (though see Cromer, 1972, and Fowler, 1984, for some suggestive evidence otherwise).

Despite these well-documented commonalities across MR and non-MR children, there is considerable variability in the findings particularly with regard to degree of language delay. Extreme delays in the MR child have occasionally even led some investigators to raise the possibility of deviant language systems (e.g., Haber & Maloney, 1978; Naremore & Dever, 1975). Conclusions regarding the extent of delay appear to vary depending on the age and etiology of the subjects, the task and language area under study, and the actual analysis employed. While some carefully selected samples of MR children acquire syntactic skills commensurate with MA expectations (cf. Kamhi & Johnston, 1982, and Lackner, 1968), most MR children lag behind their mental age counterparts (e.g., Naremore & Dever, 1975). Studies looking at quantitative factors stress the differences between MR and non-MR language; those looking for qualitative differences stress the commonalities. In general, studies relying upon spontaneous speech samples stress qualitative similarities; those with less natural, more self-conscious tasks with pre-determined expectations (cloze tasks, Berko tasks, and elicited repetition) tend to stress large differences in absolute level of skill (Bartolucci, Pierce, & Streiner, 1980, as opposed to Lovell & Bradbury, 1967). The area of language being focused upon also bears heavily upon one's conclusion regarding degree of delay.

Studies relying upon vocabulary measures find MA to be a good predictor except at abstract levels (see Rondal, 1975, and Rosenberg, 1982, for summaries); those relying upon grammatical measures find a greater lag, such that the constructions produced by MR children are less grammatical and often less complete or "complex" than those of MA-matched controls (e.g., Goda & Griffith, 1962; Lyle, 1961; McLeavey et al., 1982; Naremore & Dever, 1975; Ryan, 1975). Such a disparity between grammatical and lexical knowledge appears to hold up across both production and comprehension. Although comprehension skill, in particular, is very often assessed by a single undifferentiated measure confounding vocabulary size, sheer memory demand, and syntactic difficulty, those studies that do distinguish between structural markers (e.g., tense, number, negation) and open-class vocabulary suggest that there may be a relative sparing of lexical/semantic skill in MR children. Grammatical knowledge is overall extremely limited, and it appears to lag further and further behind lexical knowledge as chronological age increases whether the task is comprehension (e.g., Bartel, Bryen, & Keehn, 1973), or production (e.g., Bliss, Allen & Walker, 1978; Ryan, 1977).

Because of these influences of language level, language area, and language task on performance level, in discussions concerning deviance in MR langauge, what appears to be disagreement may be more a matter of terminology than of contradictory results. Although, as stated above, no study has provided positive evidence for deviance in linguistic strategies or structures, some studies have claimed deviance due to the absence of a structure or strategy that is present in the normal developmental sequence (e.g., Haber & Maloney, 1978; Naremore & Dever, 1975). For example, although non-MR children go through a stage in which their language lacks inflectional markers, one might wish to term deviant a system in which such markers never developed. In such cases, it might better be claimed that MR language is deficient rather than deviant inasmuch as it is characterized by delays far exceeding what might be expected on the basis of general cognitive measures. As hinted at in this short review, in the remainder of this paper it will prove critical to distinguish carefully between language areas and to keep in mind the particular task and language level under investigation. In addition, by looking at a well-defined subgroup of MR individuals, it will be possible to remove much of the variability contributed by etiology, chronological age, and intellectual level that confounds research on more heterogeneous groups. Thus, we may ask, in children with DS, what of language learning remains consistent with the normal learning process, and what influence does the syndrome exert uniquely regarding specific areas of difficulty, the extent of difficulties observed, and the actual course of acquisition.

III. CHILDREN WITH DOWN SYNDROME: A SPECIAL GROUP WITH REGARD TO LANGUAGE?

For as long as people have been studying language in MR children, there have been language studies focusing on children with DS specifically, either side by side with a more heterogeneous group (e.g., Lyle, 1960, 1961) or standing alone to represent retardation in general (e.g., Lenneberg, Nichols, & Rosenberger, 1964). Despite this long-standing interest, conclusions vary widely regarding whether or not Down syndrome warrants special attention, whether language associated with DS differs in important ways from language associated with retardation generally. For example, in an overview of studies conducted in the 1950's and early 1960's, Evans and Hampson (1968) concluded that there was no pattern of language acquisition unique to children with DS. They reported that first words could appear any time from one year to six years of CA, and first sentences any time from one year to 17 years CA. There appeared to be little correlation between IQ and language level (see also Lenneberg, Nichols, & Rosenberger, 1964). And, except for a generally higher incidence of articulatory and voice defects, there was no consistent relationship between how children with DS compared with other MR groups of children. On the basis of these findings, they argued that further research specifically directed to language in individuals with DS would be unrewarding. (See also Bloom & Lahey, 1978, and Rosenberg, 1982.)

In fact, however, research focusing exclusively on individuals with DS consistently reports a large discrepancy between measured language skill and MA-based expectations. In the same paper quoted from above, Evans and Hampson (1968) noted that the "worst" area of development in individuals with DS is language and that they lag behind MA-matched non-MR controls in language development by as much as 50%. Although the definitions of "verbal" and "nonverbal" are often quite general and vary from study to study, a large body of evidence points to a delay in speech relative to traditional motor, intellectual, and social indices (see Gibson, 1978, for a comprehensive review). To cite but one example, using the Stanford-Binet to assess the skills of five- and six-year-old children with DS, Thompson (1963) found that not one child in 29 had language skills that were qualitatively on a par with his overall MA score; these children were most advanced in motor development and were often more mature on the Vineland Social Maturity Scale than on the Stanford-Binet. The language delay is also evident over time, as seen in a long-term longitudinal investigation of home-reared children with DS followed from birth. Share (1975) reported that, throughout childhood, development performance on language measures was well below motor and personal/social skills, as assessed by Gesell developmental norms. Several recent studies of very young children also suggest that the inferiority in language-related skills extends back to the very earliest stages of development (Dameron, 1963; Share, Koch, Webb, & Graliker, 1964; see also Greenwald & Leonard, 1979; Mahoney, Glover, & Finger, 1981, and Mervis, in press.)

Several studies (Castellan-Paour & Paour, 1971; Lyle 1959, 1960; Mein, 1961) suggest that the language handicap is even more extreme in children with DS than in other subgroups of the MR population, matched for CA and IQ. This result is consistent enough for Evans (1977) to revise his earlier position and conclude that there is "general agreement that of all sub-groups in the severely retarded population, mongols [individuals with DS] are the most severely language handicapped" (p. 103). For example, Johnson and Abelson (1969) reported not only that the lack of the "ability to communicate with others understandably" was the most noted handicap of institutionalized adults with DS, but also that this difficulty was far more widespread among individuals with DS than in others at a comparable level of retardation. Only 18.8% of the sample with DS could "communicate usefully" (n = 2606; mean CA = 21 years, mean IQ = 28.6); this compared to 35.4% of the other severely to moderately retarded samples (n =20606, mean CA = 24.5, mean IQ = 32). In contrast to their inferiority on language, the group with DS was considered to be more socially adaptive and socially competent and obtained relatively better ratings on such rudimentary self-help items as toileting, dressing, and feeding. It is unfortunate that while circumventing the problem of selection inherent in such comparisons, the only language measure was extremely general and incorporated a number of different aspects of language.

Rondal et al. (1981) found clearer differences between MA-matched groups of moderately retarded children with and without DS when compared for performance on an elicited verbal imitation task; the group with DS showed a much greater tendency to rely upon echolalic responses, indicating a general failure to comprehend the structures in question.

For all this evidence, Rosenberg (1982) concluded that "etiology does not predict linguistic performance in the mentally retarded" (p. 338) on the basis of two studies that failed to find differences between groups with and without DS (Lyle, 1961; Ryan 1975, 1977). In Ryan's study, the subjects in both groups were home-reared and were matched on MA (mean = 3;1 years), IQ (mean = 40) and CA (5;0 to 9;0 years). Although there were small differences regarding grammatical complexity that favored the mixed MR group over the group with DS, Ryan concluded that "the similarities between the two subnormal groups were more striking than the differences" (p. 275).

There is reason to suspect that the confusion concerning this question may arise in large part from the vagueness with which the "heterogeneous MR" groups are defined and derived. Such groups are standardly defined only in terms of IQ, and may or may not include children with organic brain damage, acquired retardation, etc. Obviously, to determine the relationship between language and etiology, it is necessary to compare language development in a number of well-defined subgroups.

Two such studies have been conducted, with rather interesting results. Wing (1975, cited in Gibson, 1978) compared language skill in subjects with DS to the language acquired by IQ- and age-matched MR children who had suffered organic brain damage; she concluded that seriously retarded children with clear-cut organic pathology experience even greater language problems than do children with DS.

The second study (Burr & Rohr, 1978) looked not at the development of grammatical or syntactic skills, per se, but focused instead upon the pattern of processing and expressive skills suggested as underlying language ability and assessed in the Illinois Test of Psycholinguistic Abilities (McCarthy & Kirk, 1961). Previous studies had suggested that individuals with DS exhibit a characteristic profile on this test, in which verbal encoding and auditory processing are differentially impaired relative to those components stressing visual processing or motor encoding (Bilovsky & Share, 1965; Evans, 1973; McCarthy, 1965). Burr and Rohr compared the profile for the group with DS to profiles derived for each of three other groups of MR children matched to the group with DS on IQ and CA. The three control groups were moderately to severely retarded as a result of biological brain damage (e.g., premature birth, birth hypoxia), a known environmental cause, or no diagnosed cause. The children with DS showed a consistent disadvantage in the verbal skills relative to the visual tasks; this was true even when the tasks were matched both on dimensions of receptive versus productive skill and on the level of abstraction required. This disadvantage was both more pervasive and greater in magnitude than that found for any of the other etiological groups, leading Burr and Rohr to suggest that early prenatal disruption yields a different pattern of loss than postnatal environmentally-caused retardation; specifically, they suggested that syntactic development was most vulnerable to prenatal disruption.

In sum, there is much evidence to suggest that Down syndrome is characterized by a specific impairment within the language domain; further details on the nature and scope of this deficit will be developed below. Secondly, this impairment appears to be more extreme in children with DS than in other MR subgroups, particularly those in which the retardation resulted from damage at or after birth. However, the heterogeneity of the usual control group precludes a clean answer to this question.

IV. TAKING A DEVELOPMENTAL APPROACH

I introduced this paper arguing that it is only by taking a developmentally informed approach that we can construct an accurate and detailed picture of language learning in children with DS. One component of such an approach, justified on the research presented thus far, is to restrict our focus to biologically well-defined subgroups. In this section, I introduce three other developmental principles of potential theoretical and methodological significance.

I first point out that maturational factors (e.g., as assessed by CA) exert a unique effect on the intellectual growth curves of children with DS. Gibson (1978) reviews the evidence that MA (such as assessed by the Stanford-Binet) does not increase linearly in children with DS (Zeaman & House, 1962), although it does, by definition, in the normal case. Relatively rapid growth in the pre-school years is interrupted by plateaus in growth that become both lengthier and more frequent in the school-age years, with an actual decline in MA scores apparent by adolescence (Gibson, 1966). More current work raises the possibility that early onset of Alzheimer's disease may explain the decline in IQ, at least in part (e.g., Ross, Galaburda, & Kemper, 1984; see Gibson, 1978, for a review). The methodological implication is that matching for MA without regard to CA may group together children at varying points on a lengthy plateau.

It remains to be determined whether CA exerts an effect on language development in children with DS independent of its effect on MA. However, this is the view of Lenneberg et al. (1964), who made some important claims about a critical period of language acquisition based, in part, upon the language growth curves observed in his longitudinal study of 62 children with DS. Over a three-year period, children with DS who had attained puberty failed to make any discernible progress in acquiring language structure; this was in contrast to younger ages in which some growth was observed. Lenneberg (1967) argued that at puberty language learning was no longer possible due to a loss of brain plasticity. Although Lenneberg's critical period hypothesis and the data on which it was based have since been challenged on many important details (Best, Hoffman, & Glanville, 1982; Dennis & Whitaker, 1977; Molfese, 1977; St. James-Roberts, 1981), a critical period for language acquisition remains an important theoretical construct regarding adult-child differences in second language learning (Krashen, 1975). It must still be explained why native fluency in American Sign Language is achieved only by children under the age of 6 years (Fischer, 1978; Newport, 1982), and why age plays such a crucial role in transforming pidgin languages into creoles (Newport & Supalla, 1980; Sankoff & Laberge, 1973). Further research establishing the effects of CA specific to language learning in children with DS may prove an important empirical base for testing and refining the critical period hypothesis.

The second developmental principle derives from the "modularity" hypothesis of cognition: One cannot conduct language research without at least acknowledging the hypothesis that language is acquired, processed and represented independently of cognition generally. (See Lenneberg, 1967, for an early statement of this hypothesis supported, in part, by his findings from research on language learning in children with DS; see Fodor, 1983, for a current statement of the position and for a discussion of what constitutes relevant data; see Keil, 1981, for a discussion of modularity from a developmental perspective). Indeed, the well-documented dissociation between language structure and general MA in children with DS only serves to strengthen this hypothesis.

There are several methodological points following from the modular point of view. For one thing, the modularity hypothesis provides a strong theoretical argument against MA-matching, as it suggests MA to be a rather meaningless construct, potentially averaging over several distinct domains. It becomes only sensible to rely upon a general measure from within the language domain when matching groups of children with and without DS for comparisons of internal language structures. It also follows that to make serious statements about the specificity of the language deficit in DS, it is essential to show growth in another, distinct well-defined non-linguistic domain (e.g., see work by Cornwell, 1974, and Gelman, 1982, for exploratory studies on number knowledge in children with DS). An important corollary of the modularity hypothesis is that one must in turn treat the subdomains of language and communication (morphology, syntax, lexicon, nonverbal communication, and discourse) as potentially separable and distinct domains (see also Abrahamson, Cavallo, & McCluer, 1985; Beeghly, Hanrahan, Weiss, & Cicchetti, 1985; and Mervis, in press, for relevant studies).

Finally, in order to characterize the development of grammatical structures, one must be sensitive to the various, internally coherent, stages of language development. Although Brown (1973) introduced language stages more as descriptive aids than theoretical constructs, they have proven to be very useful heuristics, indeed, in studies of language development over the years. Whether or not distinct theoretical or biological substrata underly different stages, quite distinct grammatical systems are connoted by the pre-linguistic stage and the five stages outlined by Brown (1973) moving from one-word to complex syntax. A great deal of information would be sacrificed if performance of children at several stages of development were collapsed into a single average score. One cannot, for example, average together two children at a one-word stage (Stage I) and one child with simple syntactic structures (Stage III) and claim that children with DS are on average in language stage II (consistently combining words). Even more to the point, questions asked about language in children with DS may have quite different answers depending on the language level under study.

As is discussed below, children with DS tend to cluster at a few early stages of language development, and remain there for extended periods. This slow-motion view of development may reveal theoretically important leaps in the system that constitute major obstacles for the child with DS, even while they are achieved seemingly effortlessly by the non-handicapped child.

We were guided by these principles in a study of the internal structure of language in adolescent children with DS conducted in 1980 (Fowler et al., 1980, 1981). This study is presented in some detail as an illustration of the developmental approach, as well as to provide a framework for further research. I go on to discuss the extension of these findings to other stages, ages, and tasks in our work and that of others.

V. THE LANGUAGE STRUCTURE OF CHILDREN WITH DS COMPARED TO NON-HANDICAPPED CHILDREN MATCHED ON MLU

The aim of this study was to test the implicit assumption of the currently accepted view outlined above: that retarded language development is a slow motion and limited replica of normal language development. Under this assumption, if one were to cut out a horizontal slice in the developing language of a non-MR child at some stage, S, one should be able to superimpose this slice on the developing language of a child with DS also at Stage S′. This study involved just such a comparison, establishing comparable slices, or stages, in the developing language of children with DS, and in those who are not handicapped, and then looking for internal discrepancies.

Of particular interest was the possibility of differential delays, with different structures developing at different rates. Instead of assuming that language, or even syntax, is one bounded and homogeneous system, the focus was on the interrelationships of a number of aspects of language. Under conditions of gross mental impairment, are all aspects of language affected equally? By juxtaposing performance on several internal measures of syntactic development, the intention was to identify which structures created the greatest difficulty for children with DS and thus begin to frame some hypotheses about the nature and extent of their deficit.

As the basis of comparison, children with and without DS were equated on an anchor measure of language development: mean length of utterance in morphemes (MLU), as defined by Brown (1973). This choice was justified on several grounds. First, as has been stressed in the studies reviewed above, the more common alternative, MA, greatly overestimates the language of children with DS. The quantitative differences resulting from an MA comparison are too great to uncover any more subtle qualitative differences; they also prohibit accurate comparisons of progress among various measures since each and every measure shows a decrement relative to MA.

The second rationale for using MLU rather than the MA measure derives from research on non-MR children; there too, vocabulary development (on which MA depends heavily) proceeds independently of syntactic development. In Newport, Gleitman, and Gleitman (1977), for example, MLU fails to correlate with vocabulary usage. On the other hand, as a measure of internal syntactic complexity, MLU surpasses any other single measure of linguistic competence in the early stages of language development both in its predictive powers and in its breadth of application across child language studies (Bloom, 1970, 1973; Brown, 1973; Nelson, 1973; Newport et al., 1977; Shipley, Smith, & Gleitman, 1969).

Despite its apparent superficiality as a language description, MLU apparently incorporates a number of specific and interesting aspects of developing syntax. Shipley et al. (1969, appendix), for example, calculated the correlation of a number of early language measures, and found that MLU (there ranging from 1.0 to 2.0, in children 18 to 33 months of age) correlated with each of the other candidate measures (e.g., use of verbs, pronouns, or articles) better than any of the other measures correlated with each other, making it the best available predictor of overall language growth. (See also Newport et al., 1977 for replication up to MLU 3.5). Bloom, Lightbown, and Hood (1975) found MLU to be a useful independent variable against which to predict the expression of semantic relations; Brown (1973) and deVilliers and deVilliers (1978) found MLU to be a fair index of when certain grammatical inflections and functors would become productive, and Klima and Bellugi (1966) used MLU to make some very well-defined predictions about the emerging structure of negative and interrogative constructions.

To these one might add that this measure of productive language is relatively free from the effects of motivation, attention, practice, and inference that typically plague structured tests of language comprehension or production. Indeed, to obtain a reliable MLU one has merely to collect a minimum of 100 utterances from the child in a naturalistic situation (which can be anything from playing a game to taking a walk) where the child is at ease and has something to talk about (Brown, 1973). Thus a speech sample may easily and fairly be collected from a child of any age or language level, without that child's being test-savvy or even desirous of doing well.

Hypotheses

In this study, the first hypothesis concerned the potential dissociation between the length and complexity of utterances. As noted above, the mean length of an utterance is a good indicator of the kind of morphological and syntactic structures a non-MR child has acquired and is prone to use. However, beyond the very earliest stages, length does not logically have to mirror complexity. Although *whose book did you read?* is complex by virtue of wh-movement and subject-auxiliary inversion, it includes the same number of morphemes (five) as the more straightforward construction *they walk to the store*. It is possible that children with DS are held back in utterance length by a production limit that masks their more advanced syntactic competence. If so, one might find evidence for constructions that are short enough to fit within the production limit, but more complex than ordinarily found at that MLU stage in non-MR children.

The second hypothesis was that some aspects of language are more affected than others by general intellectual factors. Especially in light of the large disparity between MA and language level, one might expect that those facets of language most closely tied to general intelligence proceed in a different fashion from the more purely linguistic facets. Of particular interest was the use of closed class items: that area of language that is at once least semantic and most vulnerable. The class of "little" words that glue sentences together includes grammatical inflections, articles, prepositions, and pronouns. These "functors" have been characterized as carrying no content, distinguishing them from the content words that include nouns, verbs, adjectives and adverbs. They also constitute a class in historical development of languages by virtue of their finite nature: the language acquires new closed class elements at an extremely protracted rate (over hundreds of years) as distinct from the open class system that acquires hundreds of words per year. And they fall together as a phonological class in that they, as opposed to open class words (nouns, verbs, adverbs, adjectives) lose stress within the total intonation of the sentence, resulting in *he's* for *he is*, *can't* for *can not*, etc. (Kean, 1977; see Gleitman & Wanner, 1982, for an in-depth overview of the open-class closed-class distinction).

The closed class system is fragile in many respects; hence, the expectation that it might be more impaired in retarded populations. It develops relatively late in child language (e.g., Bloom, 1970; Brown, 1973) and is missing both in pidgin languages (Newport, 1982; Sankoff & Laberge, 1973) and in invented language (homesign) of deaf children (Feldman, Goldin-Meadow, & Gleitman, 1978). It is selectively impaired in Broca's aphasia (Kean, 1977) and is more dependent on input characteristics than other language components (Gleitman, Newport, & Gleitman, 1984). Despite the well-attested fragility of closed class grammar and its distinctive developmental course (again, see Gleitman & Wanner, 1982, 1984, for an extended discussion), it has not been adequately determined whether this specific aspect of language is differentially affected in individuals with DS.

In sum, not all facets of language skill should be affected equivalently by Down syndrome retardation. Specifically, it was suggested 1) that length would be more compromised than syntactic complexity and 2) that the closed class system would be more affected than either open class vocabulary or sentence length. Thus, when children with DS and non-MR children are matched for utterance length, the grammar of those with DS is predicted to be more advanced in terms of syntactic complexity (as in embedding) and vocabulary measures, and less well-developed on measures of morphological development.

Subjects

The subjects included four moderately retarded adolescents in a stimulating parochial day school. The four were selected for homogeneity of age (CA 10;9 to 13;0 years, mean = 12;7), intelligence level (Stanford-Binet IQ 46 to 56; mean = 51; MA 6 to 7 years; mean = 6;3), and language level (all were in Stage III, using the criteria of Brown, 1973). This language level (MLU 2.75 to 3.25; mean = 2.98) was representative of adolescent children with DS at that school and appears to be in keeping with other studies of adolescents with DS (Mein, 1961; Ryan, 1975; Semmel & Dolley, 1971). As controls, non-MR youngsters were sought who had MLU's as close to 3.0 as possible. This was not easy; by 36 months of age, children have moved well beyond that language level and one has to "catch" younger children as they move through the level of interest. The controls in this study were 30 to 32 months of age, consistent with established norms (Bloom, 1970; Brown, 1973; Miller, 1980). The young age was surprising because the adolescents appeared, at least impressionistically, to be communicating well beyond a 3-year-old level. The resulting disparity in age (CA 12 vs. 2 1/2 years; MA 6 vs. 3 years) made dissociation all the more likely in the various analyses employed. (Refer to Table 1 for subject statistics).

Procedure

All analyses were based upon spontaneous speech samples collected under naturalistic conditions. Children interacted with the experimenter at school in a half-hour of free play/conversation in a quiet room away from the classroom. The children had met and talked with the experimenter on previous occasions. The conversation centered about a large three-dimensional house with miniature furniture and people. The children were encouraged to talk about the objects, to talk about own homes and families, and even to make up conversations using the objects as props (i.e., to play house). The sessions were videotaped and recorded on a hi-fidelity stereo tape recorder. The recordings were transcribed as soon as possible after the task by the experimenter. The utterance length measures were obtained as described below. In addition, internal analyses were performed to tap three different aspects of linguistic skill: syntactic complexity, grammatical morphology (closed class knowledge), and vocabulary (open class knowledge).

Table 1

Subject Characteristics

	Children with Down Syndrome (n=4)	MLU-matched controls (n=4)	T
Chronological age (years)			
mean	12;3	2;7	18.25***
range	10;9-13;0	2;6-2;8	
Mental age (years)			
mean	6;3		
range	6;0-6;9		
IQ			
mean	51		
range	48-50		
Utterance length statistics (based on spontaneous utterances)			
Number utterances produced	166.5	195.0	-0.86
MLU: mean length in morphemes	2.98	3.32	-1.32
Upper bound in morphemes	7.50	8.75	-1.56
Longest 10%: MLU in words	5.35	5.50	-0.29
Proportion one-word utterances	0.15	0.12	0.68
Word: morpheme ratio	0.95	0.92	1.63

Note: Groups were compared using independent-sample 2-tailed t-tests.
***$p < .001$; ** $p < .01$; *$p < .05$; +$p < .10$.

Utterance Length Measures

The guidelines outlined in Brown (1973) served as the basis for the calculation of MLU, with some variations. The primary deviation from his procedure was to incorporate in this analysis only those utterances initiated by the child; responses, imitations, and long lists without internal structure (e.g., counting or recitation of family names) were excluded. Additionally, only five tokens of any exact phrasing were included in the MLU analysis in order to relieve any possible bias introduced by repeated use of stock phrase such as *what's that?* Finally, analyses in this study were based upon all utterances produced rather than on only the first 100 utterances of a transcript. (See Fowler, 1984, for justification of these variations).

Brown's (1973) upper bound serves as an index of optimal—as opposed to mean—performance. This is simply the length of the single longest utterance produced in the session, measured in morphemes. To the same end, the MLU for the longest 10% of the utterances was also calculated. The fourth and final utterance length measure concerned the distribution of utterance lengths. Given an average length of utterance, the goal was to determine whether, perhaps as a function of usage factors, children with DS rely to a greater extent on very short utterances than other children with comparable syntactic skill.

Results. By design, all subjects had MLU's indicating they were in Brown's Stage III or IV. The average MLU of the group with DS and the non-MR group was 3.0 (range = 2.6-3.3) and 3.3 (range = 2.8-3.8), respectively (see Table 1). As intended, the difference between groups was not significant. In keeping with a generally low usage of inflectional morphology, MLU in words rather than morphemes was slightly, though comparably, lower in both groups (2.8 and 3.0 for the children with DS and the non-MR children, respectively); the word:morpheme ratios were .95 and .92, again not significantly different.

Brown reports the upper bound (in morphemes) for Stage III to be 9 morphemes. In this study the upper bound of the children with DS ranged from 7 to 9; of the normally developing children from 7 to 10, for a non-significant difference. When only the longest 10% of the utterances were taken into account, the MLU's (in words) of the two groups were nearly identical (children with DS: 5.35; non-MR children: 5.50). Although it might have been expected that the children with DS would be more variable in sentence length, alternating perhaps between long and very short utterances, this was not the case. The shape of the distribution of utterance lengths was similar across both groups. In each group, two and three word utterances made up more than 50% of the total and there was no utterance longer than 9 words. One-word utterances made up 15% of the utterances produced by the children with DS and 12% of the utterances produced by the non-MR children; this difference was nonsignificant.

Syntactic Measures

Measures of internal complexity. As a first look at the internal structure of the utterances produced, counts were made of the number of major constituents (noun phrases and verbs) per utterance and, in turn, of the number of words and morphemes entering into each constituent category. Length may derive from a greater number of sentential constituents (e.g., *see you help boy*) and from the internal length of these constituents (e.g., *going to the store*). These measures were developed by Newport et al. (1977), who found them to correlate highly with MLU in very young non-MR children.

This analysis was based upon a sample of 50 consecutive utterances from each child. Candidate utterances required a main verb other than *be* (explicit or not). A new MLU based on this sample was also computed. For this analysis, nouns were defined as nouns or pronouns, whether in subject or object position; they could also occur as objects of prepositions. Morphemes per noun phrase also included articles, inflectional markers; words per noun phrase excluded inflections. Verbs were defined to include only immediate verbal constituents: semi-modals, auxiliaries, *not* and the verb. Sentential elements not included in this count were optional elements such as adverbs as well as particles and prepositions.

Results. The mean length of non-copulaic utterances was identical across populations; in addition, there were virtually the same number of major constituents (noun phrases and verbs) per utterance (see Table 2). There was, however, one internal difference: the group with DS relied upon significantly fewer morphemes per noun phrase. As will be supported in other measures to be presented, this is due, in part, to the fact the children with DS relied heavily on pronouns rather than upon expanded noun phrases.

Table 2

Measures of Internal Complexity (following Newport et al., 1977)
(Based upon 50 consecutive nonequational utterances)

	Children with Down Syndrome (n=4)	MLU-matched controls (n=4)	T
Mean words per utterance	3.68	3.73	-0.18
Major constituents per utterance			
Nounphrases/utterances	1.37	1.37	0.10
Verbs/utterance	1.07	1.01	1.81
Other (particles, adverbs)	0.36	0.55	-1.43
Internal structure of major constitutents			
Words/nounphrase	1.62	1.72	-1.63
Morphemes/nounphrase	1.71	1.85	-2.44*
Morphemes/verbphrase	1.45	1.50	-0.40

Note: Groups were compared using independent-sample 2-tailed t-tests.
$^{**}p < .01$; $^{*}p < .05$; $^{+}p < .10$

Expression of thematic relations. As a context-dependent measure of syntactic development, productions were analyzed for presence or absence of arguments made obligatory by certain verbs (or implied verbs). Does a child with DS choose to encode and delete the same thematic arguments as the non-MR child, when both are equally limited in productive language skill? Attention was focused upon locative expressions for two reasons. First, they were elicited in quantity by the task at hand—placing objects in a model house. And second, normative developmental data on locatives were collected by Bloom, Miller, and Hood (1975); their analysis was the basis of the coding and scoring scheme relied upon here. Bloom et al. investigated three locative sentence frames. "Agent-Locative Action" utterances were defined as expressing the movement of an object (Patient) by an independent Agent, as in *I'm trying to put this back in here*. This sentence type requires an Agent (syntactic subject), Patient (object), Place, and Verb. The "Mover-Locative Action" sentence type, e.g., *you can come my house*, includes those utterances specifying a movement in which the Agent was also the Patient moved, hence the "Mover." This argument structure requires a Mover (subject), Verb, and Place. Bloom et al.'s third locative category, "Locative State," specifies a state rather than a movement, e.g., *sleep, sit, be, belong* and *go*, as in *this goes here*. This category also requires a Patient (subject), Verb, and Place.

Results. At the language level under study, children in both groups were explicit in expressing most of the obligatory thematic relations tapped (see Table 3). Of the 10 relations under study,

Table 3

Oligatory Thematic Relations Expressed in Locative Utterances (Proportion of times an argument or verb is expressed where obligatory) (following Bloom et al., 1975)

	Children with Down Syndrome (n=4)	MLU-matched controls (n=4)	T
Locative Utterance Category			
Agent Locative (e.g., *I'm trying to put this in here*)			
Mean # in category	6.75	11.75	-1.27
Agent	0.39	0.51	-1.67
Verb	1.00	0.99	0.67
Patient	0.76	0.96	-1.07
Place	0.87	0.99	-1.25
Mover Locative (e.g., *Daddy going trolley*)			
Mean # in category	12.25	12.25	0.00
Mover (Agent *and* Patient)	0.65	0.69	-0.23
Verb	0.65	1.00	0.00
Place	0.97	0.83	1.31
Locative State (e.g., *She sleep in bed*)			
Mean # in category	12.25	23.75	-1.01
Patient	0.93	0.84	0.80
Verb/copula	0.79	0.62	0.80
Place	0.94	0.96	-0.31
Average score for 10 thematic categories	0.85	0.84	0.38
Proportion Syntactic Categories Expressed[a]			
Nouns	0.72	0.75	-0.45
Subject	0.70	0.68	0.30
Object	0.76	0.96	-1.07
Verbs	0.93	0.92	0.22
Stative	0.79	0.62	0.93
Non-stative	1.00	0.99	0.00
Place terms	0.95	0.92	0.47

[a] Syntactic scores were obtained by averaging across thematic categories, according each sentence type was equal weight.

Note: Groups were compared using independent-sample t-tests (2-tailed) tests). There were no significant differences across groups.

85.1% were expressed by the group with DS and 83.7% by the non-MR group. When scores were compared across individual and averaged categories, there were no significant differences across the two groups.

Although small numbers and potential differences in scoring procedures preclude statistical comparison with Bloom et al.'s findings, the subjects in this study seemed to be performing as well as, if not better than, the children at the close of the normative study (MLU=3.0). In both studies, non-stative verbs were expressed virtually 100% of the time. In addition, the children in this study consistently included the place term across the three locative sentence types (DS 95%; non-MR 92% expressed); this was well beyond the 55% average quoted by Bloom et al. Bloom also reported that stative locative verbs (*sits, belongs*, etc.), were consistently expressed in her population (near 100%), just as were the non-stative verbs. In this study, however, non-stative verbs were often only implicit in both groups; frequently producing utterances such as *the doll ____ upstairs* while placing it in a dollhouse implying a verb such as *goes* or *belongs*; the difference between groups was not significant (children with DS 79% expressed; non-MR children 62%). In part, of course, this may arise from a failure to produce the contractible copula (*the doll's upstairs*).

When attention is restricted to obligatory nominal arguments, more variability is apparent. Overall, nominal arguments entering into locative relations were supplied 72% of the time by children with DS and 75% of the time by the non-MR group. The largest tendency to omit an obligatory nominal involved the Agent (subject) of "Agent locative utterances," as in *I'll put this right here*. Although the tendency not to express this form was somewhat greater among the DS group (39% supplied) than among the non-MR group (51% supplied), not one of the eight subjects supplied it more than 75% of the time. (Bloom et al. reports an average score of 54%.)

The children were more apt to express the "Mover" in the Mover Locative category. Although this category is similar to the Agent category above in also serving as sentence subject, thematically it also serves as Patient. The two groups performed comparably (64.8% DS; 68.8% non-MR) with two children in each group supplying this argument less than 75% of the time, and only one child (non-MR) supplying it more than 90% of the time.

Sentence subjects were most consistently supplied when functioning as the Patient in the Locative State utterances (e.g., ____ *belongs in the kitchen*). Each child supplied this argument a minimum of 75% of the time, which was the mean performance reported by Bloom. The high scores observed (92.8% DS; 83.5% non-MR) may be artificially inflated by coding procedures requiring that a minimum of two arguments be supplied in order to be included in the analysis. Thus *here*, meaning *this goes here*, would not be included in the analysis.

Bloom reports that the Patient in object position, as in Agent Locative utterances (e.g., *I'll put ____ right here*), was supplied consistently (70%) from MLU 1.2 on. In this study, as anticipated, each child in the non-MR provided it at least 90% of the time (group average 96%). Although the average score obtained by the group with DS (75.5%) did not differ significantly from the non-MR score, there was variability in the group with DS. Whereas two of the children with DS, like the non-MR children, provided the Patient consistently, a third supplied it only 80% of the time, and the fourth consistently failed to express it (22% supplied).

Bloom et al. focused on children at the earliest stages of language; their observations, presented in thematic terms, were concluded at just the level of interest here. Although the children in this study performed like Bloom et al.'s youngsters in being more apt to express a nominal argument that incorporates both Agent and Patient than one that is Agent alone, the more obvious facts seem to concern grammatical categories. Although differences between groups regarding tendency to express sentence subject (children with DS 70%; non-MR children 68%) or grammatical object (children with DS 76%; non-MR children 96%) failed to reach significance, a significant interaction indicates a greater split between these two categories in the non-MR group than in the group with DS.

Indices of complex syntax. The two previous syntactic measures are useful in analyzing language samples produced from the onset of two-word combinations in Stage I. They should, and do, reveal a good deal of positive evidence for internal linguistic structure and regularity. In the level of interest here, children are on the verge of mastering aspects of complex syntax. Most notably, they must acquire the intricacies of the verbal auxiliary system (both movement rules and grammatical terminology) to produce correct negative and interrogative structures. Additionally, children just beyond this stage acquire many means of coordinating and embedding sentence clauses such as conjunctions, relative clause structures, or preposing; here, too, they must master movement rules and grammatical markers. To measure progress made in this direction, a count was made in this study of some of the earliest appearing aspects of complex syntax, such as have been noted by researchers such as Menyuk (1969) or Bellugi (1967). Constructions of interest included subordination and coordination, subject/auxiliary version in yes/no and wh-questions; passive voice constructions with *got* or *be*, and choice of negative markers.

Confirmation of syntactic development was sought via a normed measure of syntactic complexity devised by Scarborough (1985): The Index of Productive Syntax (IPSyn). This measure, based upon a 100-utterance sample, awards points for the occurrence of 56 kinds of morphological and syntactic forms. Normative data are available, based upon 48 transcripts from children aged 24 to 48 months; the scale is of potential value for children below and beyond this age range. Because full credit is awarded if a construction occurs twice in a 100-utterance transcript, this analysis serves as a measure of optimal more than mean performance.

Results. There was very little evidence for use of complex constructions in either group; what small differences existed between groups failed to reach significance (see Table 4a). Little use was noted of the passive construction, subject auxiliary inversions, possessives or conjoined clauses. In both groups the primary means of expressing negation was through negative modals, fitting with descriptions in the literature of negation at this stage (e.g., Klima & Bellugi, 1966). Primitive forms like *he have no chin* or *this not fit* were observed primarily in the non-MR group, while the children with DS tended to produce very few negatives overall.

Multi-verb utterances were also rare, comprising less than 5% of the utterances in either group, and consisting primarily of conjunctions and concatenations. Similar multi-verb utterances were produced in both samples: a DS subject produced *put it on get more*, but *do that fix this* came from a non-MR child, *I want it shut* was produced by a child with DS and *and keep a door closed* by a non-MR child. Preverbal adjectives occurred rarely and in both samples consisted of such common constructions as *big truck* or *little boy*.

Table 4a

Indices of Grammatical Complexity (Presented as average number of occurrences per 100 utterances)

	Children with Down Syndrome (n=4)	MLU-matched controls (n=4)	T
Utterances with 2 or more sentence nodes	3.83	2.38	1.39
Conjunction			
Noun + noun	1.15	0.17	1.42
Verb + verb	0.40	0.00	1.00
Sentence + sentence	0.42	0.60	0.27
Pre-nominal adjectives	1.13	2.13	-1.62
Negative forms			
No + verb	0.00	0.98	-1.31
Not + verb	0.20	1.18	-1.70
Negative modals	3.40	2.58	0.44
Passive forms			
Got + verb	0.50	0.00	0.73
Be + verb	0.00	0.00	0.00
Use of possessive form			
Noun's noun	0.85	0.60	0.37
Auxiliary inversion in yes/no questions	0.25	0.20	0.18

Note: Groups were compared using independent-sample t-tests (2-tailed tests). There were no significant differences between groups.

Table 4b

Performance on Index of Productive Syntax (Based on Scarborough, 1985)

	Children with Down Syndrome (n=4)	MLU-matched controls (n=4)	T
Subscale			
Noun phrase	16.25	15.75	0.29
Verb phrase	17.50	16.75	0.70
Questions and negatives	11.50	9.75	0.63
Sentences	15.50	12.00	2.18+
Total	60.50	54.25	1.89

Note: Groups were compared using independent-sample t-tests (2-tailed tests).
$^{**}p < .01$; $^{*}p < .05$; $+p < .10$

The two groups also performed comparably in overall performance on Scarborough's (1985) IPSyn measure (children with DS 60.5; non-MR children 54.25). The scores of the group with DS were consistent with those reported by Scarborough for her 30-month-old group of nonhandicapped children; the non-MR group here was somewhat behind this average. Although the two groups in this study were highly comparable on the three subscales of this test tapping complexity of noun phrases, verbal auxiliary development, and devices for constructing negative and interrogative sentences, there was a notable difference regarding sentence complexity. This last subscale looks at means of embedding and coordinating sentence clauses; however, following Miller (1980), it is more concerned with movement rules and word order than it is with whether the particular grammatical markers are expressed. Thus, the child who produces *want my mommy come here* is credited with being able to produce infinitival sentences with a subject distinct from the matrix subject despite the fact that the infinitival marker *to* was unexpressed. Although both groups, and Scarborough's as well, had this advantage, as shown in Table 4b, the group with DS were relatively more advanced on such constructions than were the non-MR children.

The Closed Class System

It was suggested above that the acquisition of closed class vocabulary might be more affected in individuals with DS than would the acquisition of open class vocabulary. In particular, it was hypothesized that grammatical morphology might lag behind other aspects of syntactic development. This question was addressed in part by the *Word/morpheme ratio* presented in Table 1; this measure of use of bound inflections is calculated as the ratio of the number of "words" (defined phonologically) to the number of morphemes. The difference is made up by nominal and verbal markers such as the plural or past tense marker. The ratio of word to morpheme was very high across groups; neither group made much use of inflectional morphology (DS 0.95; non-MR .92). The nonsignificant advantage that does occur for the non-MR group is a function of one child with DS who used very few inflections at all, and one non-MR child who was more advanced on this measure. The two analyses presented next look in more detail at closed class knowledge.

Closed class vocabulary. To assess the development of closed class vocabulary, a count was made of all tokens of the five major closed class categories: pronouns, prepositions, modals, wh-forms, and demonstratives. This figure was compared to the total number of words in each speech sample. To supplement this measure of dependence on closed class vocabulary, we also made a measure of lexical sophistication within these categories by looking at the diversity of vocabulary types falling into these and other categories of closed class markers.

Results. Pronouns, prepositions, modals, wh-forms and demonstratives made up a similar proportion, overall, of the lexical items produced by children with DS (31%) and by non-MR children (30%). Differences between groups within individual categories were not remarkable, although there was a marginally significant tendency for the non-MR children to use more demonstrative terms (*this*, *that*); this was complemented by a non-significant tendency on the part of the children with DS to use a greater number of pronouns (see Table 5). In terms of diversity of the closed-class items employed, children with DS had a slight, but non-significant, advantage over the non-MR children. This difference was particularly evident with modals and wh-forms, where it neared or attained significance (see Table 6).

Table 5

Usage of Closed Class Vocabulary–Tokens
(mean proportion of words falling into each of 5 closed class categories)

	Children with Down Syndrome (n=4)	MLU-matched controls (n=4)	T
Mean # words per corpus	468.25	474.75	-01.18
Closed class categories			
Pronouns	16.56	13.81	1.57
Prepositions	3.66	4.52	-0.59
Modals/semimodals	3.23	1.97	1.21
Wh-terms	3.54	2.78	1.17
Demonstratives	4.01	6.68	-2.02+
Total	31.00	29.75	0.51

Note: Groups were compared using independent-sample t-tests (2-tailed tests).
$^{**}p < .01$ $^{*}p < .05$ $+\, p < .10$

Use of grammatical morphemes in obligatory context. For a more fine-tuned measure of morphological development, an assessment was made of use in obligatory context of the first 14 grammatical morphemes to occur regularly in early child language. These include prepositions, verbal auxiliaries, and nominal and verbal inflections. Brown was able to identify the contexts where these morphemes should appear, calculate the proportion of cases where they do in fact appear, and track the development of each over the language-learning period. His methods were adapted for cross-sectional study by deVilliers and deVilliers (1978); their procedures were adopted here.

Results. Sufficient information was available for comparison across groups for eight of the fourteen morphemes (i.e., there were at least four identifiable obligatory contexts per subject); averaging across these morphemes the overall percentage supplied for each group was virtually identical: 68% for the group with DS, 66% for the non-MR group (see Table 6). There appear, however, to be some differences regarding the pattern of acquisition for individual morphemes. This is evidenced in the scores achieved for the progressive *(-ing)*, the first morpheme typically acquired. Although the non-MR children provided this form quite consistently (86% supplied), the children with DS were much more variable (61% supplied, $p < .10$). There are eight cases of full mastery of the first four morphemes (90% usage in obligatory context) among the four non-MR children in this study, while there are only four such cases among the non-MR group. This difference is masked when one looks across the full range of the 14 morphemes: there, the overall score is nine in each group. In some sense, the non-MR group is "more normal," mastering the earlier morphemes first and only later going on to acquire the more difficult morphemes. While one cannot infer the developmental sequence from this single point data, it appears that the DS group has moved on to more difficult morphemes (notably, the copula) without having fully mastered the simplest ones.

Table 6

Usage of 14 Grammatical Morphemes in Obligatory Contexts
(following Brown, 1973 and deVilliers & deVilliers, 1978)
(# subjects reaching 90% criterion presented in parentheses)

Grammatical morphemes	Children with Down Syndrome (n=4) Proportion of contexts filled	MLU-matched controls (n=4) Proportion of contexts filled	T
Progressive marker - ing	0.61(0)	0.86(2)	-2.16+
Preposition on	0.84(1)	–[a](1)	–
Plural -s	0.84(2)	0.92(2)	-0.59
Preposition in	0.73(1)	0.75(3)	-0.07
Past tense irregular	0.82(0)	0.63(0)	1.20
Articles a & the	0.43(0)	0.49(0)	-0.46
Possessive 's	–(0)	–(0)	–
3rd person irregular	–(0)	–(0)	–
Contractible copula	0.82(2)	0.62(0)	1.31
3rd person regular -s	0.25(0)	0.36(0)	-0.71
Past tense regular	–	–	–
Contractible auxiliary BE	–(1)	0.39(0)	–
Uncontractible auxiliary BE	(0)	–(0)	–
Overall proportion supplied[b]	0.68	0.66	0.16
Mean number morphemes acquired per child	2.25	2.25	0.00

[a] Averages were not calculated when there were fewer than four obligatory contexts for three out of four subjects.

[b] Overall proportion based upon just those 8 morphemes for which there was sufficient data for both groups.

Note: Groups were compared using independent-sample t-tests (2 tailed tests).
$+p < .10$

It may well be that what non-MR children learn, they learn fully, hence rapidly reaching near 100% on the earliest morphemes. Children with DS, in contrast, may work on more constructions simultaneously but never acquire fully—or use consistently—even the earliest rules acquired.

Open-class Vocabulary Measure

At the language level under study, it is difficult to make a fair assessment of productive vocabulary, especially in a single spontaneous speech sample. Nonetheless, a gross measure of vocabulary usage was derived by calculating the ratio of different vocabulary types produced to the total number of words in the corpus (following Nelson, 1973). While this measure clearly does not explore the depth of the child' lexical knowledge for open class items, it probably does provide a fair index of the closed class lexicon, which is, after all, both finite and frequently used. The type/token ratio also serves as a measure of information value independent of the syntactic structure employed.

Table 7

Measure of Lexical Diversity
(Presented as number of different vocabulary types per 100 words)

	Children with Down Syndrome (n=4)	MLU-matched controls (n=4)	T
Number of words in corpus	468.25	474.75	-0.180
Open class categories			
Nouns	5.94	9.48	-3.69*
Verbs	5.52	5.79	-0.27
Adjectives	0.90	1.33	-0.70
Adverbs	3.05	2.54	1.08
Total	15.41	19.13	-2.10+
Closed class categories			
Pronouns	3.09	2.56	1.20
Prepositions	1.33	1.48	-0.66
Modals & semimodals	1.36	0.69	2.46*
Wh-forms	1.08	0.68	2.31+
Demonstratives	0.48	0.48	0.00
Quantifiers	0.93	0.87	0.22
Logical forms	0.25	0.42	-1.47
Total	8.50	7.18	1.82
Overall Total	23.90	26.47	-1.03

Note: Groups were compared using independent-sample t-tests (2-tailed tests).
$^*p < .05, +p < .10$

Results. Non-MR children produced an average of 19.13 different open-class vocabulary types per 100 words produced; they appeared to rely upon a wider range of open-class vocabulary than did the children with DS ($mean = 15.41, t = -2.10, p < .10$) This nearly significant difference derives almost entirely from the significantly greater tendency of the non-MR group to use different noun types (see Table 7). Thus, although PPVT scores indicate the children with DS should have access to a higher vocabulary, they tended not to rely upon a large (especially nominal) vocabulary in this spontaneous speech sample. As was observed in the analysis of closed class vocabulary, the child with DS tended to rely on pronouns where possible.

Summary

In sum, despite the fact that this set of measures was chosen specifically with the aim of uncovering plausible differences, similarities across groups of children at the same language stage were more striking than were any small differences. Performance on each measure was within the expectations derived from the normal literature relevant to that stage, and was confirmed on the basis of measures made on a non-MR control group. The children with DS in this study appear to be at a linguistically stable, if extremely restricted, point of development, with no syntactic or grammatical measure deviating from that stage one way or the other. Overall, the combination of the syntactic measures employed reveal DS subjects to be at a level of simple phrase structure grammar—one that cannot be reduced to semantic generalizations perhaps (see Slobin, 1980, for a discussion of the normal course), but one that precedes, across the board, the dramatic changes required to build the complex syntax with its associated verbal auxiliary system, sentential embedding, and movement rules.

Despite the small sample sizes, some differences between groups were large and consistent enough to approach or obtain a level of significance. In contrast to our initial expectations, where differences did occur, it was usually the child with DS who lagged behind the normally developing child despite attested higher verbal MA: they supplied early grammatical morphemes and grammatical objects less consistently, and produced less complex noun phrases. And, despite relatively advanced receptive vocabulary scores, in actual usage, the children with DS relied more heavily on pronouns than nouns.

Even where children with DS maintain an advantage, there are important disclaimers. For instance, far from being altogether insensitive to the little words of language, they were actually somewhat more advanced in the acquisition of different closed class terminology than non-MR children at this language level. And yet, other analyses make it apparent that the children with DS are extremely limited in their ability to use these forms appropriately and consistently to serve syntactic/grammatical functions. Similarly, although they produced complex sentences of appropriate length and word order, our results suggest that this level of syntactic complexity is not at all supported by appropriate grammatical markers.

Although the results of this study are at best suggestive, it serves as a good starting point in arguing several points about the development of language structure in children with DS. These arguments have gained considerable support in subsequent research, and have been extended to children at different ages, at different language stages, and in different experimental tasks.

VI. DISCUSSION: THE NATURE OF LANGUAGE IN CHILDREN WITH DS

VI.1. Internal Structure as a Function of Language Stage

There are, by now, many studies reiterating the main finding of this study: when appropriate matching procedures are employed, internal analyses fail to distinguish the language structures employed by children with DS from those produced by non-MR children at that language stage. This is true at the language level discussed above as well as at earlier stages of development, whether it is receptive or productive knowledge being assessed. This conclusion also appears to hold in the few longitudinal studies that have been undertaken to date. In most of these studies, MLU has been the matching criterion of choice. The consistency of language levels within DS has resulted in a number of studies focusing on a few language stages, thus yielding a more detailed picture of language development than currently available for other, more mixed, groups.

The virtue of comparing language structure of children with DS and without DS as a function of general language level has recently been demonstrated in several studies of mildly to moderately retarded children with DS at Brown's (1973) Stage I. At this early stage, marked by the onset of two-word utterances, the focus has been on determining the "relational" or "thematic" meanings encoded in two-word combinations (Buium, Rynders, & Turnure, 1974; Coggins, 1979; Dooley, 1977; Fowler, 1984). Although these studies vary in the semantic coding systems employed, they are in full agreement that children with DS choose to encode much the same thematic relations as non-MR children at the same language stage. Coggins (1979), for example, classified all two-word utterances of four children with DS (aged 3;10 to 6;2) to find that they showed as much diversity as non-MR children studied in Bloom, Lightbown, and Hood (1975) and Schlesinger (1971). Approximately 70% of their utterances fell into the nine "semantic categories" most frequently encoded by non-MR children at Stage I. Paralleling normal development, Coggins observed a shift from Early Stage I to Late Stage I in the tendency of the children with DS to rely increasingly on locative and stative categories. Coggins also noted that the children with DS, like non-MR children at this stage, failed to make conditional or hypothetical statements, refer to past or future events, or use grammatical morphemes productively. Within this same language stage, Harris (1983) reports that there is an effect of chronological age (CA 2;6 - 6;9 years) on language structure. The younger children with DS were indistinguishable from MLU-matched non-MR children, but the older children employed somewhat different means to expand their utterances.

A common feature of non-MR speaker at Stage I is the spontaneous imitation of adult utterances. Coggins and Morrison (1981) found that, like non-MR children studied by Bloom, Hood, and Lightbown (1974), children with DS vary considerably in the amount of spontaneous imitations produced. Of primary interest to these researchers was the finding that children with DS were selective in their choice of structures to imitate: like non-MR children, they chose to imitate words that did not occur frequently in their spontaneous productions. (See Shipley et al., 1969, for the normal case.) Coggins and Morrison suggested that this selective imitation points to a strategic attempt to resolve differences between the child's productions and the adult input.

Rondal (1980) found spontaneous imitations of children with DS to be comparable in both number and complexity to those produced by non-MR children, not only at Stage I (MLU 1.0-1.5 and 1.75-2.25), but at a somewhat higher MLU level (2.5-3.0) as well. Language level was a significant factor in both groups: in children with DS as in non-MR children, the absolute number of imitations decreased and the length and complexity of imitations increased as language level improved. In both groups, there was a tendency to imitate the end of a maternal utterance, which Rondal suggests shows a "perceptual centration" on the final part of adult utterances. (See Slobin, 1973, for the normal case.) Indeed, of a wide number of measures taken, the only difference distinguishing the children with DS from the non-MR children was a greater tendency to imitate modifiers from the mother's speech.

Two studies have been conducted, looking at the internal structure of language as it develops over time, starting at Stage I. Dooley (1977) conducted a year-long observational study of two children with DS (IQ 51 and 44; starting CA 3;10 and 5;2); both children were in Brown's (1973) language Stage I throughout the study. Over the year, one child made approximately one month's progress (MLU 1.48 to 1.75) relative to non-MR children studied in Brown (1973); the other child actually declined somewhat in MLU (1.84 to 1.73). With the exception of the fact that they failed to change significantly over the period, the children were similar to Stage I non-MR children on internal measures (semantic relations, grammatical morphemes, utterance diversity and size of lexicon). The only difference of note was a greater tendency on the part of the children with DS to rely heavily on routinized expressions and proforms (*it*, *they*, *here*, *there*, *do*) in keeping with our findings above. Dooley, however, notes that this tendency also varies in the normal population.

Fowler (1984) reports on a child with DS, Rebecca, (IQ 57; CA 4;3 to 7;5) who was observed on a monthly basis over a three year period (CA 4;3 to 6;11 years); Rebecca's language moved from Stage I to Stage III/IV over this time period. Despite a slow start in Stage I, Rebecca grew consistently from MLU 1.4 to 3.5 in a normally rapid period of time (13 months). This was followed by an extended plateau (10 months) at the 3.5 period. Near the close of the study, Rebecca was veering into Stage V (MLU = 4.0), but with considerable fluctuation. On internal analyses (semantic relations, grammatical morphology, negative and interrogative constructions), Rebecca's language was indistinguishable from the norm between the stages of I and III, but for a slight advance in semantic relations relative to MLU. Upon reaching the 3.5 threshold, progress on these internal measures halted and interesting overgeneralizations were observed in the verbal auxiliary system and wh-interrogation. Although modest further progress has been made in her syntactic development, this development appears to have diverged from the normal course both in rate and character.

At a higher language level, Ryan (1975, 1977) compared children with DS with non-MR and non-DS MR children on general measures of internal structure evident in spontaneous speech; she too found no differences between groups when matched on MLU. In a study on thematic relations extending the work above with Stage I children, Layton and Sharifi (1979) found that school-aged children with DS (CA 7;4 to 12;2 years; MLU 3.5-5.5) used verb types similar to those found in non-MR children matched on MLU. However, they did find different proportions of usage. Most notable was a much lower use in the DS group of "process" verbs involving a change of stage, e.g., *John cut the paper*, to explain which they invoked a conceptual deficiency. Rondal (1978a) found no overall differences between children with DS and language-matched controls on syntactic complexity as determined by Lee's (1974) Developmental Sentence Scoring procedure. Like Wiegel-Crump (1981), however, he reports that at the higher MLU levels advanced constructions fail to develop in the child with DS at a level commensurate with normal expectations.

Studies of comprehension skill as a function of general language stage suggest that structural language deficits cut across comprehension equivalently. Bridges and Smith (1984) compared the performance of children with DS (CA 4;4 to 17;1) with that of non-MR children (CA 1;11 to 4;4), when asked to act out active and passive semantically biased and neutral sentences. Rather than relying upon MA, they used a general standardized comprehension measure (Reynell, 1969) as their criterion for selecting normal controls. There were six groups of matched pairs ranging in VCA (Verbal Comprehension Age) from 2;6 to 5;0 years. In this study, groups did not differ significantly in overall percentage correct response. The pattern of errors was comparable across groups; the groups were affected equivalently by manipulations of the semantic plausibility of the sentences presented. The children with DS were, however, six to twelve months behind the non-retarded children (as measured by VCA) in acquiring syntactic comprehension strategies.

Fowler (1984) also used an object manipulation procedure to compare the syntactic comprehension strategies and skills of schoolchildren with DS (CA 6;9 to 16;8) with those of young toddlers (CA 2;3 to 3;2) matched for two different productive language levels (MLU 3.15 to 3.55 and 4.0 to 4.65); the DS group had significantly higher MA's as assessed on the PPVT-R (Dunn & Dunn, 1981). In this study, too, there were no differences in mean performance overall, or in the semantically reversible condition. The effect of syntactic type (actives, passives, datives, and three relative clause types) was constant across groups. However, there was a significant interaction between group and effect of semantic plausibility, suggesting that the children with DS may depend more heavily on semantic inferencing skills. The children with DS performed better in the plausible conditions, and worse in the implausible conditions, than did the non-MR controls. The apparent contradiction between Fowler (1984) and Bridges and Smith (1984) is likely a function of the matching procedure. The VCA measure employed by Bridges and Smith incorporates lexical as well as syntactic knowledge; consistent with the lexical/syntactic split discussed below, it appears to have overestimated syntactic knowledge in the DS group, but to have been a fair measure of semantic inferencing skill. On the other hand, the measure selected by Fowler (MLU) turned out to be an appropriate index of syntactic comprehension knowledge, but underestimated lexical knowledge. Differences between groups in Fowler (1984) relevant to plausibility factors appear to be a function of large differences in vocabulary level. (See Dewart, 1979, for comparable results with a group of MR children without DS).

VI.2. Differential Delays Within the Verbal/Communication Domain

I reviewed evidence above demonstrating that children with DS experience difficulties with verbal measures that are incommensurate with their general developmental status. In these studies, however, language skills are often not well-defined or contained and the contrast skill or score is often a vague all-purpose measure. In this section, I go one step further to argue that even relative to other well-defined measures within the verbal/communication domain, the child with DS appears to be especially hampered in acquiring grammatical and syntactic structures. A large number of studies point to a split between lexical and structural knowledge; a growing literature suggests that syntactic development may also lag behind other closely related well-defined skills such as communicative function and symbolic play. These disparities are evident from the onset of syntax, with the gap only increasing as the child becomes older.

Lexical versus structural development. The study presented in detail above, the superior MA's of the children with DS relative to their MLU scores support the hypothesis that lexical knowledge in children with DS is spared relative to structural knowledge as assessed by measures of syntax and morphology. This particular hypothesis has received considerable, if indirect, support. In studies of lexical knowledge, whether the task is to produce, recognize, or identify vocabulary items, children with DS are on a par with, or only slightly behind, MA-matched controls, retarded and not (e.g., Bartel et al., 1973; Blount, 1968; Lyle, 1960; Ryan, 1975; Spreen, 1965). On the other hand, when grammatical or syntactic indices are taken, children with DS standardly perform on a lower level than their MA peers (Lyle, 1960; Paour & Paour, 1971 [cited in Gibson, 1978]; Ryan, 1975; Spreen, 1965). Most convincing are several studies in which grammatical and lexical measures have been made on the same subjects, demonstrating a greater delay for grammatical than for lexical forms, relative to MA. This disparity is maintained whether the task involves production or comprehension structures and appears to increase with chronological and mental age (Bless, Swift, & Rosen, 1985; Hartley, 1982; Lyle, 1960; Rogers, 1975; Ryan, 1975).

Evans (1977) provided evidence for such a lexical/structural split in a factor-analytic study using a wide range of measures including verbal and nonverbal intelligence, CA, several structural and fluency measures derived from spontaneous speech samples, and performance on the subtests of the ITPA (McCarthy & Kirk, 1961). In Evans' data there was a consistent pattern of clustering of measures according to three largely distinct factors. Significantly, there was a sharp dichotomy between the "general verbal ability" factor and the "structure of speech" factor. Stanford-Binet Mental Age was the highest loading on the verbal factor, but was insignificant on the "structural" factor; CA, in contrast, loaded moderately but significantly on the "structural" factor, but not on the "verbal" factor. Those measures loading on the structural factor included sentence length, complexity, ratio of nouns to verbs and other parts of speech, and a task of grammatical morphology. These were largely independent from the general "verbal" measures including vocabulary, general intelligence tests, and most of the ITPA measures. Those measures tapping fluency of speech correlated only among themselves and nonsignificantly with the other two factors.

Hartley (1982) also found that DS difficulties in syntax surpassed those in the lexicon and went on to show that this disparity was more extreme in children with DS than in a mixed MR group. Her conclusion was based on patterns of performance on *The Token Test for Children* (Di Simoni, 1978) which she administered to children with DS (mean CA 11;0 years), to MR children of other etiologies (mean CA 10;7 years), and to non-MR children (mean CA 4;7 years). The three groups were matched on receptive vocabulary score: mean MA on the PPVT was 4;2 years for each group. The Token Test, which requires the subject to manipulate wooden objects of varying shapes and colors, incorporates conditions that allow the experimenter to manipulate memory demands directly, as in (1) versus (2) below, as well as syntactic requirements, such as (2) versus (3). In this analysis, Hartley further divided the tasks standardly labeled syntactic into those she considered to be primarily syntactic and those which she thought were really more lexical or "spatial," illustrated by (3) versus (4).

(1) *Give me the red circle*

(2) *Pick up the big red circle and the small green square*

(3) *Before picking up the yellow circle, touch the red square*

(4) *Put the white circle in front of the blue square*

Both groups of MR children were significantly less successful on all parts of the test than non-MR children matched for MA. Children with DS, in turn, performed less well than other non-DS children of similar age and IQ. Only on the spatial items was there no difference between the DS and the non-DS children; this was in direct contrast to the significant difference found between MR groups on the sentences stressing syntactic skill. Although Hartley hypothesized that children with DS process "simultaneous," but not "sequential" information, her results can also be handled by invoking a lexical/syntactic split that holds up even when the lexical items under study are highly relational spatial terms.

A more recent body of developmentally informed research has examined the development of language structure, as indexed by MLU, within the context of more general communication skills. This literature has focused on Stage I of syntactic development. The general finding is that children with DS exhibit communicative skills at a level equivalent to, or more advanced than, communicative behavior of non-MR children at Stage I (Coggins, Carpenter, & Owings, 1983; Coggins & Stoel-Gammon, 1982; Leifer & Lewis, 1984; Owings & MacDonald, 1982; Scherer & Owings, 1984). Beeghly and Cicchetti (1985) found not only that the communicative skills of children with DS were more mature than those of MLU-matched controls, but went on to show that they were equivalent to those of MA-matched controls whose MLU well exceeded that of the group with DS. Related findings are reported by Weiss, Beeghly and Cicchetti (1985) in regard to symbolic play, often thought to rest on the same cognitive basis as language behavior. There too, on indices of symbolic play, children with DS perform at a level on a par with general cognitive development and are generally more advanced than their language level (Stage I) alone would predict. Other studies of older individuals with DS suggest that functional communicative skills exceed verbal abilities, but these have not been as well controlled (Leuder, Fraser, & Jeeves, 1981; Nisbet, Zanella, & Miller, 1984; Price-Williams & Sabsay, 1979). One must be careful, however, to keep in mind that a relative advantage on measures of communicative skill such as turn-taking or very early gestural skill need not imply that children with DS would have any real advantage in acquiring a structured sign language over spoken language. Although there is an advantage for gestural over spoken communication at the earliest stages of development in all children, this advantage appears to drop out for both non-MR children and children with DS prior to the onset of syntax (Abrahamsen et al., 1985; Petitto, 1985).

The emergence of language. Thus far, I have focused upon children who have already begun to combine words (Stage I). Recently, however, there has been a surge of interest attempting to place the onset of syntax—and of first words—within a well-defined developmental context. This research on the pre-linguistic and early linguistic child with DS is in keeping with research at other stages regarding a lack of qualitative differences. Children with DS speak with the same early vocabulary (Gillham, 1979), follow highly constrained rules of lexical acquisition (Mervis, in press), and select from similar kinds of nonverbal gestures as those used by normal children (Greenwald & Leonard, 1979). Within narrow and well-defined domains, development proceeds in an orderly, if not rapid fashion.

Also consistent with the findings reviewed thus far, confusion is generated at the first attempt to examine the interrelationship between one and another subdomain of language or, between general cognitive development and the emergence of language. Although Lenneberg et al. (1964) reported that the onset of first words and first sentences in children with DS intercalated nicely with major developmental milestones of walking and running, other studies conducted at that time suggest that children with DS show an inferiority in verbal skill relative to more general developmental measures even at early ages (Dameron, 1963; Share et al., 1964). A few recent studies have made a more careful exploration of the relationship between the onset of first words and phrases, and measures of cognitive development hypothesized to underlie language skill (Greenwald & Leonard, 1979; Mahoney et al., 1981, and Mervis, in press) to yield a rather complicated story.

Greenwald and Leonard (1979) focused on the prelinguistic stage, comparing performance on a sensorimotor task with measures of communicative behavior in situations designed to induce either an imperative or declarative response. Children with DS and normally developing children were selected as consistently functioning either at stage 4 or 5 on three scales of the Uzgiris and Hunt (1975) Ordinal Scales of Psychological Development. At stage 4, the children with DS ranged in age (CA) from 10 to 19 months (mean IQ 52); these were compared to non-MR children between 7 and 9 months of age (mean IQ 110). At this emphatically pre-linguistic level, no group differences emerged in communication measures: both groups were equally adept at using "imperative" gestures and neither group displayed any evidence of making "declarative" gestures. More advanced levels of communicative behavior were found at stage 5 in both groups. However, although the imperative task continued to yield nonsignificant group differences, the non-MR children (CA 9-13 months; mean IQ 125) by sensorimotor Stage 5 had achieved a significant advantage over the children with DS (CA 16-26 months; mean IQ 68) in the kinds of declarative gestures relied upon. The major source of this advantage was a greater tendency on the part of the non-MR children to express themselves verbally. Indeed, the group with DS relied almost totally on nonverbal gesture whereas vocalization frequently replaced or accompanied gesture in the non-MR group.

A separate subgroup of older pre-schoolers with DS (CA 31 to 54 months; mean IQ = 62), although still at Stage 5 on the sensorimotor scale, obtained more advanced communication scores and were more apt to use a verbal response than either of the younger groups. The authors suggest this shows some independence between the cognitive and linguistic tasks, perhaps facilitated by the massive linguistic intervention provided for these children. While it seems that the Stage 5 level of sensorimotor development may be associated with a wide range of communicative skill, even within the groups with DS some of the variability may arise from grouping as a single category both verbal and nonverbal communication skills (see Beeghly & Cicchetti, 1985, for a cleaner separation of these skills in Stage I children). Greenwald and Leonard's findings that declarative, but not imperative, gestures are delayed in the infant with DS are of particular interest in light of research on referential looking in the pre-linguistic infant with DS. Miller (in press) reviews several studies, most notably Jones (1977, 1980) in which children with DS (MA 8 to 22 months) exhibited significant deficits in referential looking behavior relative to general developmental level. Jones (1980), however, did not find that the children with DS vocalized significantly less than the MA-matched non-MR children. Further disentanglement of what pre-linguistic behaviors do and do not correlate with later linguistic patterns will surely aid our theoretical understanding of what elements constitute precursors to language learning.

Mahoney et al. (1981) employed more explicitly linguistic (as opposed to communication) measures to report that language skill lagged behind sensorimotor functioning in very young children with DS. They matched children with DS (CA 24 to 38 months) and non-MR children (CA 12 to 19 months) on the Bayley Mental Developmental Scale (Bayley, 1969) such that each group approximated a developmental age of 17 months (DS 13 to 23 months, mean = 16.8; non-MR 12 to 23 months; mean =17.1). Although the children with DS scored comparably or higher than the non-MR children on 5 out of the 6 subscales of Uzgiris-Hunt scales of general sensorimotor development, the non-MR group had a substantial advantage on the vocal imitation subscale and on both receptive and expressive measures of the Receptive and Expressive Early Language Test (Bzoch & League, 1970).

Finally, in a longitudinal study bridging the gap between pre-language and language, Mervis (in press; see also Cardoso-Martins, Mervis, & Mervis, 1985) examined the ongoing relationship between cognitive development and linguistic development. Consistent with the findings of

Lenneberg et al. (1964), she reports that the emergence of referential language, both in comprehension and production, is in keeping with developmental expectations. In children with DS, referential comprehension was first demonstrated at a mean MA of 14.5 months (Bayley) in children with DS, compared to a mean MA of 13.8 months in non-MR children. Similarly, the onset of referential production occurred at a mean MA of 18.9 months in the group with DS compared to a mean MA of 19.5 months in the non-MR group. These results were supported by a separate analysis that showed children with DS and non-MR children to be comparable in their stages of sensorimotor development at the onset of referential communication.

On the other hand, Mervis reports vocabulary size is less well associated with sensorimotor measures. When groups were matched on sensorimotor development at four distinct points (all within stages 5 and 6), vocabulary size was consistently smaller in the children with DS; the greatest difference in vocabulary size was apparent at the final point, upon complete attainment of stage 6. Mervis did not find differences in vocabulary size when she equated groups on interpolated MA levels (Bayley, 1969), but stresses that such differences were apparent in a larger-scale study conducted by Strominger, Winkler, and Cohen (1984). In that study, 36-month-old children with DS showed large deficits in vocabulary size relative to their mental age. Although they had a mean mental age of 20.6 months on the Bayley scales, their vocabulary size (mean = 18.5 words; range 0 to 85) was closer to that of a normally developing child of 18 months (mean vocabulary = 22) than to the average 21 month old. By 21 months, the non-MR child has typically undergone a vocabulary explosion and has an average vocabulary of 118 words. Mervis concludes that "these results suggest that a major reason for the large differences in size of vocabulary between children with DS and non-handicapped children is that the vocabulary spurt for the child with DS does not begin at the mental age that would be expected, based on the findings concerning non-handicapped children" (p. 43).

Given the close relationship between the vocabulary spurt and the onset of syntax in the non-handicapped child (Gleitman & Wanner, 1982; Lenneberg, 1967), this particular delay may presage directly the structural deficits already discussed.

VI.3 Homogeneity of Language Structure: A Ceiling on Language Development?

In recent years, much attention has been paid to the fact that, despite a poor prognosis for language development, some children with DS acquire considerable linguistic maturity and may even learn to read and write. Consider, for instance, the celebrated account of a mildly retarded DS boy (IQ 60) who learned to read and write, and kept a diary of his life from CA 11 to 45 years (Seagoe, 1965). At his most mature stage (age 24), his language, analyzed for syntactic complexity, was at the level of a bright five year old, except for a high redundancy of words of "low rank" and few words at a "high rank."

Despite the well-documented fact that language in DS can range from mutism to linguistic maturity, seemingly independent of IQ (Evans & Hampson, 1968; Lenneberg et al., 1964), such reports obscure the fact that individuals with DS by and large fail to move beyond the most rudimentary stages of syntactic development. Gibson (1978) reviews two very early studies on DS language that report that DS children cluster at extremely low level of language development. Muir (1903) reported that the development of language in children with DS, although paralleling that of normal children in very early vocabulary development, did not move beyond the stage of acquiring a few verbs even in his "most accomplished mongol." Similarly, Brousseau and

Brainerd (1928) stated that children with DS were "disinclined" to use sentences. More evidence derives from Thompson (1963) and Wiegel-Crump (1981), both of whom remarked upon the reliance of subjects with DS on short, simple sentences. Wiegel-Crump (1981) used Lee's (1974) Developmental Sentence Scoring procedure to assess the language skills of 80 children with DS (CA 6;0 - 12;7 years; MA 2;0 to 6;11 years). These children exhibited a more homogeneous pattern of syntactic usage than MA-matched non-MR children; they relied almost exclusively on low-level syntactic structures with very little variety of construction: this changed little across the MA-range observed.

This clustering of DS children at very low language levels is echoed in a number of recent studies, also focusing on syntactic skills, but using more experimental procedures. Semmel and Dolley (1971) made a study of comprehension and imitation skills, relying upon the procedures and theoretical framework underlying the work of Lovell and Dixon (1967), Lackner (1968), and Graham and Gulliford (1968), but restricted themselves to children with DS (CA 6 to 14 years; IQ 22 to 62). Although qualitative similarities held across these studies in regard to the role of syntactic type in these tasks, Semmel and Dolley found that most of the DS subjects could comprehend and reproduce only simple active declaratives; they were at base level for negatives, passives, and negative passive sentences. Individuals with DS seemed to need contextual cues to make any sense of these sentence types.

Analogously, Evans (1977) was completely unsuccessful in his attempt to administer the Berko morphological task to his sample of 101 DS individuals (CA 8.3 to 31.1 years; MA 2.5 to 7.8 years). Out of the 101 subjects, 96 scored 6 or fewer correct responses to the 30 items in the nonsense condition; 61 could not even do the simplest task of supplying a plural for *wug*. Even in his spontaneous language measures, Evans noted that "because of limited speech development no subject reached the ceiling in either the Sentence Structure test or the Linguistic Features Count" (p. 113). In that analysis, more than a third of the utterances produced were one-word long, and the 10 longest utterances in each subject's corpus averaged only 5.2 words.

Rondal et al. (1981) also reported very low levels of skill in sentence repetition tasks, especially in their subjects with DS. Only three Down's subjects out of 19 (age range 5;2 to 12;7 years; IQ range 40 to 49) could successfully imitate a five word active declarative sentence; when the declarative was negative, still five words in length, only one subject could handle it. There were many echolalic responses. Although the constructions produced under these conditions would clearly be considered "deviant" if produced spontaneously, such constructions have been observed in very young non-MR children (24 to 30 months CA) when asked to repeat sentences well beyond their level of proficiency (Lenneberg et al., 1964; Fowler, 1980).

Finally, it should be noted that even studies ostensibly reporting a range of language skill in individuals with DS suggest a largely homogeneous pattern. In the population studied by Lenneberg et al. (1964), for example, 32 of the 35 subjects (CA 5;6 to 13;6) for which he reports data fell into his two middle categories of stage of language development ("mostly words," "primitive phrases"). Only two of the 35 were at the "sentence" level of development, while one child was still at the "babbling" stage (this observation derives from Rosenberg, 1982). Lenneberg et al. found little relation between IQ and language level, but did hint at an IQ threshold of 50, beyond which considerable syntactic competence was possible though certainly not guaranteed.

Andrews and Andrews (1977) also reported tremendous variability in the spontaneous speech samples collected from 39 DS children (CA 5;8 to 17;9; IQ 31 to 60). Although there are methodological problems with this study due to the limited quantity of speech collected per child, it appears that the variability that exists stems primarily from usage factors (quantity of speech and the like). On structural measures such as MLU, however, the children in this sample were highly consistent in obtaining low scores.

From these various studies, the generalization seems to emerge that DS may exert considerable influence on the ultimate language level that a DS individual child will attain. Although a few will reach the language level of a five year old and will have limited skills in reading and writing, the child with DS is not apt to move beyond the level of simple phrase structure grammar found in non-MR children younger than three years of age.

VII. EXPLANATORY ACCOUNTS OF THE SPECIFIC LANGUAGE DEFICIT IN CHILDREN WITH DS

On the basis of the research presented, it is quite clear that the course and limits on language learning in children with DS cannot be explained as a simple function of general cognitive development, either as assessed by MA or by more sophisticated measures of communicative skill or sensorimotor development (see also Cromer, 1974, 1976). Several explanatory accounts have been put forth in an attempt to explain why it is there is such a large discrepancy between measured MA and the final language attainments of individuals with DS, and why learning appears to stop at the point that it does. Although none of these accounts has been unequivocally ruled out, and several may play a role, I present them here as possible directions for future research.

A commonly invoked explanation for deficits in heterogeneous MR children above and beyond those predicted by general cognitive factors is that they are less motivated than non-MR MA peers in performing the task at hand (Zigler, 1969). However, motivation fails to account for normal language learning: whereas it is presumably extremely crucial for all individuals to have a rudimentary communication system for basic needs, the elaboration of such a system to incorporate complex linguistic syntactic forms (e.g., *can I have a cookie* vs. *cookie!*) lacks motivational explanation in all cases. And yet, it is just this elaboration that is lacking in the child with DS. Note too that the child with DS studied today is not the depressed institutionalized child of yesterday about whom motivational accounts were hypothesized. Today's subjects often live at home in family settings and receive special attention from very early ages at home and in the schools. Motivation is also high in adult workshops: Evans (1977) noted his subjects with DS were highly motivated to use speech and regularly did so with the support of linguistically able non-MR adults. Nonetheless these individuals continued to perform in language well below MA expectations. Perhaps most telling concerning the role of motivation in acquiring syntax are two explicit studies of the effects of institutionalization on language in children with DS. Although McNutt and Leri (1979) found that institutionalization dampened progress across many areas of development, including vocabulary and verbosity, both they and Wiegel-Crump (1981) report no differences between institutionalized and home-reared children with DS on a wide range of syntactic measures.

A second account put forth to account for the extreme linguistic deficits in children with DS refers specifically to the language environment. Given the discrepancy between the child's age/size and linguistic capacities, it is possible that the speech directed to him or her is inappropriate.

That is, whatever advantage is derived from the tendency to speak to young children in short, simple sentences (comprehensibility, perhaps), may be denied the retarded child at similar stages in the acquisition sequence. (For discussion of the facilitative effects of motherese in normal language learning, see Fernald & Simon, 1984; Furrow, Nelson, & Benedict, 1977; Gleitman et al., 1984, and Newport et al. 1977.) A number of studies have addressed this issue and rendered this hypothesis implausible; mothers interact with their children with DS in much the same way as they interact with their non-MR children at comparable language levels. Most compelling is the finding that MLU is a much better predictor than MA or CA of the characteristics of a mother's speech to her child. (See Peterson & Sherrod, 1982; Rondal, 1978b, and Rosenberg, 1982, for comprehensive reviews of the relevant literature). Even where differences in mother's style of interaction do emerge, it has not been possible to assign a specific result to that difference. Mervis (in press) reports that although mothers make fewer adjustments to accommodate to the early lexical categories of their children with DS, this difference seems to have no direct effect on the categories formed and maintained by their children.

Miller (in press) is in agreement with these results, but suggests that a lack of maternal responsiveness at the pre-linguistic level may be responsible for a delay in language. In a review of the literature, he reports that mothers are less responsive to infants with DS (and vice versa); that mothers of infants with DS tend to talk while their chid is vocalizing rather than engaging in turn-taking behavior, and that mothers are more directive, intrusive, and controlling while interacting with children with DS than are mothers interacting with non-MR children of comparable mental age (see also Beeghly, Weiss, & Cicchetti, 1984; Hanzlik & Stevenson, 1986). The results of intervention studies seeking to mitigate these differences look promising; however, thus far, they lack controls. Furthermore, it is not clear how one should go about matching pre-linguistic groups as a function of language level; we know in any case that early sensorimotor and communication measures are poor indices of how far along the child is toward acquiring verbal skill.

A third proposal to account for the specific language deficit suggests that the neurological structures underlying language are particularly impaired in children with DS. Although there is overwhelming evidence of anatomical, physiological, and neurochemical abnormalities in the brain of children with DS (e.g., Ross et al., 1984), such differences have not yet been specifically related to differences in language function. Two attempts have, however, been made to relate language differences to a lack of a dominant language hemisphere as assessed by dichotic listening tasks.

Looking for the biological underpinnings of a specific language impairment in individuals with DS, Zekulin-Hartley (1981) administered a dichotic listening task to children with DS, to heterogeneous MR children matched to the group with DS on IQ and CA, and to non-MR controls matched to the MR groups on MA. Whereas both groups of children without DS showed the expected right ear advantage in processing linguistic serially presented auditory stimuli (digits and common object labels), children with DS showed a significant left ear advantage.

Sommers and Starkey (1977) used the dichotic listening paradigm to look at differences within the syndrome comparing two groups of children representing the extremes in language functioning. All subjects, selected from a day-school population of 150 children, met a minimum criterion of being able to point correctly to each of the dichotic words and foils used in the test; all had normal hearing and were predominately right-handed. The high-language group (n =15, CA 7;0 to 19;0 years, mean IQ = 46, mean MA = 64 months) was functioning at a language age of 5;3, as assessed by the Carrow (1973) Test for Auditory Comprehension of Language. These children had a mean MLU of 7.5 words, used both simple and complex grammatical structures, and produced appropriate and intelligible verbal responses to stimulus materials. In contrast, the low performance group (n = 14, CA 7;1 to 17;8, mean IQ = 39, mean MA = 51 months) was functioning at a language level of 3;7 years, had a mean MLU of 2.5, and spoke in three- and four-word sentences at the base phrase structure level. Their speech was telegraphic, intelligibility was poor, and verbal perseveration and inappropriate responses were common. Despite the clear differences between groups in linguistic skill, and despite the great care taken to use a sensitive and reliable measure, the ear advantage in both groups was essentially zero. This contrasted with non-MR controls (3;0 to 5;5 years CA, mean MA 5;1) who showed the expected significant right ear advantage. Although dichotic listening tasks are fraught with methodological pitfalls (Bryden, 1982), the lack of a right ear advantage is at least suggestive of a seriously impaired language area. Converging evidence from other biological indices is crucial to exploring this hypothesis further.

The fourth explanation is Lenneberg's (1967) critical period hypothesis discussed above. As noted there, this hypothesis fails on many points. Nonetheless, a maturational account contains a great deal of merit. All longitudinal studies spanning the pre-school years report maximal language growth to occur by age 7. For example, Share (1975) reports on the developmental progress, measured using the Gesell developmental examination, of 76 children with DS followed from birth; at the time of the report, these children ranged from 7 to 18 years CA. In this study, most language development occurred between CA 4 and 6 years, during which time the child with DS typically moved from the developmental level of 15 months MA to two and a half or three years MA. This report is consistent with our findings (Fowler, 1984; see Fowler et al., forthcoming) and also appears to fit with the data presented in Lenneberg et al. (1964). Indirect support is provided by several cross-sectional studies in which the gap between syntactic and lexical knowledge widens with growth in CA (see Miller, in press, for a review). A revised version of the critical period hypothesis would suggest that specialized language abilities are no longer available after the age of 7 years, consistent with reports from normal language development (Newport, 1982).

Alternatively, one might wish to account for these same facts not as a function of chronological age, but of linguistic stage. If, indeed, impaired brain function sets a limit on the level of language that can be accommodated, it may well be that language skill is acquired in a timely fashion consistent with general maturational development, but beyond the limiting language level, growth will cease or differ from the normal course (see Fowler, 1984). The considerable evidence speaking to a ceiling on language skill supports this hypothesis, but from the data presented, it is not possible to distinguish the effects of chronological age from those of language stage. (See Fowler et al., forthcoming, for a full discussion of these last two hypotheses and for a more extended set of longitudinal data bearing on this issue).

VIII. OVERVIEW

In sum, the measures and procedures that have been used to study language in DS vary in many ways. But, taken together, they seem to suggest a coherent picture. Individuals with DS tend to form a well-organized human subset with regard to the prognosis for language development: They are a diminished case of the normal within the language domain and represent a particularly well-studied and straightforward case of the delay/deficit observed in MR language generally.

1) Research on children with DS serves to reinforce the view that language in MR individuals is qualitatively similar to the language of younger non-MR children, particularly in Stages I through III. This holds up across a wide number of detailed internal analyses such as the one presented here. Such studies are motivated and informed by the current developmental perspective in that groups are matched for developmental level within the developing domain of interest (here, language), reflecting a general trend toward domain-specific research (e.g. Keil, 1981). Also in keeping with this trend, the language measures taken are not presented as a single summary index: rather, the internal measures employed in research on children with DS (thematic relations, morphological usage, etc.) are treated as coherent rule-governed systems that are potentially separable from other language measures. Although work on other well-defined domains to contrast with syntactic growth is progressing, much more research needs to be done in this direction.

2) By virtue of studying a biologically well-defined subgroup, some insight has been gained into the course and prognosis of language development under one condition of retardation. Individuals with DS appear to be more consistent in terms of the degree of language delay than are more diffuse groups of MR individuals. Language consistently develops more slowly in children with DS than do other aspects of motor or cognitive development; this lag is evident in infancy and grows wider as the children with DS become older.

3) Although comparisons between biologically well-defined subgroups remain to be done, the available evidence suggests that children with DS, as a group, are more impaired in linguistic skill than are subgroups of children whose retardation is acquired, and are on a par with, or behind, subgroups with obvious organic brain-damage.

4) Working from a developmental model of normal language acquisition, and finding comparable matches in terms of language stage, it becomes possible to determine what the child with DS is or is not capable of in absolute terms, rather than relative to some arbitrary MA measure. Some children with DS do acquire substantial linguistic competence, to the level of a normally developing five year old. Although a general IQ of approximately 50 appears to be a prerequisite for, though not a guarantee of linguistic success, it has not yet proven possible to predict, on non-linguistic grounds, how to distinguish the few children who will acquire syntax from the great majority whose language will level off at an early age without acquiring the complexities of syntax or morphology. Many of these children appear to be stalled at the level of a two and a half year old non-MR child.

5) It was suggested above that MA overestimates structural linguistic knowledge in MR populations generally, which might reflect a split between lexical and syntactic knowledge. This question has been systematically addressed in DS, experimentally manipulating the demands on the child to tap one or the other aspect of language skill. Converging evidence from both comprehension and production strongly supports the conclusion that the language deficit in DS is most pronounced in the grammatico-syntactic components with a relative sparing within the lexical domain (also within nonverbal communication). On the other hand, within the structural components of language, it appears that the deficit is consistent across syntax and morphology and across comprehension and production.

6) Although there are no data available for other subgroups of MR individuals, a small data base is accruing concerning the course of language learning over time in individuals with DS, based on longitudinal and systematic cross-sectional research. On the basis of this preliminary

research, it appears that language learning does not proceed at a constant pace from birth to puberty. Rather, a great deal of language learning appears to take place by seven to nine years of age (CA); for brief periods it may proceed at a near normal pace. In regard to internal analyses of the language systems acquired, the conclusions presented above appear to hold up across the entire developmental sequence, whether defined in terms of CA, MA, or language level, with a widening of the semantic/syntactic gap as CA increases.

Much work remains in order to understand why children with DS stop where they do in the language acquisition sequence, to determine why some children do in fact make better progress in language, to explain differences in consistency and usage that emerge on detailed internal analyses, and to explore more fully the possibility of different learning mechanisms leading to similar results. It should be apparent that the research on DS, by taking a stance that is at once developmental, domain-specific, and biological has greatly enhanced our general understanding of the effects of retardation on language development.

References

Abrahamsen, A., Cavallo, M., & McLuer, J. A. (1985). Is the sign advantage a robust phenomenon? From gesture to language in two modalities. *Merrill-Palmer Quarterly.*

Anastasiow, N. J., & Stayrook, N. G. (1973). Miscue language patterns of mildly retarded and nonretarded students. *American Journal of Mental Deficiency, 77*, 431-434.

Andrews, R. J., & Andrews, J. G. (1977). A study of the spontaneous oral language of DS children. *Exceptional Child, 24*, 86-94.

Bartel, N., Bryen, D., & Keehn, S. (1973). Language comprehension in the mentally retarded child. *Exceptional Children, 39*, 375-382.

Bartolucci, G., Pierce, S. J., & Streiner, D. (1980). Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. *Journal of Autism and Developmental Disorders, 10*.

Bayley, N. (1969). *Bayley Scales of Infant Development: Birth to Two Years*. New York: Psychological Corporation.

Beeghly, M., & Cicchetti, D. (1985). Development of functional communication during Stage I of syntactic development by children with Down syndrome. Presented at *10th Annual Boston University Conference on Language Development*, October, 1985.

Beeghly, M., Weiss, B., & Cicchetti, D. (1984). Structure and style of free play behavior in Down syndrome and non-handicapped children. Paper presented at the *International Conference on Infant Studies*, New York, April, 1984.

Beier, E. G., Starkweather, J. A., & Lambert, M. J. (1969). Vocabulary usage of mentally retarded children. *American Journal of Mental Deficiency, 73*, 927-934.

Bellugi, U. (1967). *The acquisition of negation*. Unpublished doctoral dissertation, Harvard University.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150-177.

Berry, P. (1976). Elicited imitation and assessment of abilities. *Language and Speech, 19*, 363-373.

Berry, P., & Foxen, T. (1975). Imitation and comprehension in severe subnormality. *Language and Speech, 18*, 195-203.

Best, C. T., Hoffman, H., & Glanville, B. B. (1982). Development of infant ear asymmetries for speech and music. *Perception & Psychophysics, 31*, 75-85.

Bilovsky, D., & Share, J. (1965). The ITPA and Down's Syndrome: An exploratory study. *American Journal of Mental Deficiency, 70*, 78-82.

Bless, D., Swift, E., & Rosen, M. (1985). *Communication profiles of children with Down syndrome*. Waisman Center on Mental Retardation and Human Development. University of Wisconsin, Madison, Wisconson.

Bliss, L., Allen, D., & Walker, D. (1978). Sentence structures of trainable and educable mentally retarded subjects. *Journal of Speech and Hearing Research, 21*, 722-731.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Bloom, L. (1973). *One word at a time*. The Hague: Mouton.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when and why. *Cognitive Psychology, 6*, 380-420.

Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: Wiley.

Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and variation in child language. *Monograph of the Society for Research in Child Development, 40* (2).

Bloom, L., Miller, P., & Hood, L. (1975). Variation and reduction as aspects of competence in language development. In A. Pick (Ed.), *Minnesota Symposium on Child Language* (Vol. 9).

Blount, W. R. (1968). Language and the more severely retarded: A review. *American Journal of Mental Deficiency, 73*, 21-29.

Bridges, A., & Smith, J. (1984). Syntactic comprehension in Down's Syndrome children. *British Journal of Psychology, 75*, 187-196.

Brousseau, K., & Brainerd, M. G. (1928). *Mongolism: A study of the psychical and mental characteristics of mongoloid imbeciles*. Baltimore: Williams and Williams.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Bryden, M. P. (1982). *Laterality. Functional asymmetry in the Intact Brain*. New York: Academic Press.

Buium, N., Rynders, J., & Turnure, J. (1974). Early maternal linguistic environment of normal and Down's Syndrome language-learning children. *American Journal of Mental Deficiency, 79*, 52-58.

Burr, D. B., & Rohr, A. (1978). Patterns of psycholinguistic development in the severely retarded: A hypothesis. *Social Biology, 25*, 15-22.

Bzoch, K. R., & League, R. (1970). *The Receptive-Expressive Emergent Language Scale for the Measurement of Language Skills in Infancy*. Gainesville, FL: The Tree of Life Press.

Cardoso-Martins, C., Mervis, C. B., & Mervis, C. A. (1985). Early vocabulary acquisition by children with Down syndrome. *American Journal of Mental Deficiency, 90*, 177-184.

Carrow, M. (1968). The development of auditory comprehension of language structure in children. *Journal of Speech and Hearing Disorders, 33*, 99-111.

Carrow, E. (1973). *Test for Auditory Comprehension of Language* (5th ed.) Austin: Learning Concepts.

Castellan-Paour, M., & Paour, J. (1971). Attempted syntactic structure reeducation in the adolescent mongol. *Rev. Neuropsychiatr. Infant Hyg. Ment. Enfance, 19*, 449-461. [Cited in Gibson, 1978.]

Coggins, T. E. (1979). Relational meaning encoded in the two-word utterances of Stage 1 Down's Syndrome children. *Journal of Speech and Hearing Research, 22*, 166-178.

Coggins, T. E., Carpenter, R. L., & Owings, N. O. (1983). Examining early intentional communication in Down's syndrome and nonretarded children. *British Journal of Disorders of Communication, 18*, 98-106.

Coggins, T. E., & Morrison, J. A. (1981). Spontaneous imitations of Down's Syndrome children: A lexical analysis. *Journal of Speech and Hearing Research, 46*, 303-308.

Coggins, T. E., & Stoel-Gammon, C. (1982). Clarification strategies used by four Downs Syndrome children for maintaining normal conversational interaction. *Education and Training of the Mentally Retarded, 17*, 65-67.

Cornwell, A. (1974). Development of language, abstraction, and numerical concept formation in Down's syndrome children. *American Journal of Mental Deficiency, 79*, 179-190.

Cromer, R. F. (1972). Learning of linguistic surface structure cues to deep structure by educationally subnormal children. *American Journal of Mental Deficiency, 77*, 346-353.

Cromer, R. F. (1974). Receptive language in the mentally retarded. In R. L. Schiefelbusch & L. L. Lloyd (Eds), *Language Perspectives: Acquisition, Retardation and Intervention*. Baltimore: University Park Press.

Cromer, R. F. (1976). The cognitive hypothesis. In D. M. Morehead & A. E. Morehead, (Eds.) , *Normal and deficient child language*. Baltimore: University Park Press.

Dameron, L. (1963). Development of intelligence of infants with Mongolism. *Child Development, 34*, 733-738.

Dennis, M., & Whitaker, H. A. (1977). Hemispheric equipotentiality and language acquisitioin. In S. J. Segalowitz & F. A. Gruber (Eds.), *Language development and neurological theory*. New York: Academic Press.

Dever, R. B., & Gardner, W. I. (1970). Performance of normal and retarded boys in Berko's test of morphology. *Language and Speech, 13*, 162-181.

deVilliers, J., & deVilliers, P. (1978). A cross-sectional study of the acquisition of grammatical morphemes in child speech. In L. Bloom, (Ed.), *Readings in language development*. New York: Wiley.

Dewart, M. H. (1979). Language comprehension processes of mentally retarded children. *American Journal of Mental Deficiency, 84*, 177-83.

DiSimoni, F. (1978). *The Token Test for Children*. Massachusetts: Teaching Resources.

Dooley, J. (977). *Language acquisition and Down's Syndrome: A study of early semantics and syntax*. Unpublished doctoral dissertation, Harvard University.

Duchan, J. F., & Erickson, J. G. (1976). Normal and retarded children's understanding of semantic relations in different verbal contexts. *Journal of Speech and Hearing Research, 19*, 767-776.

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Text -Revised*. Circle Pines, MN: American Guidance Service.

Evans, D. (1973). *Some language abilities of severely mentally handicapped persons with special reference to Down's Syndrome*. Doctoral dissertation, Exeter University, England.

Evans, D. (1977). The development of language abilities in mongols: A correlational study. *Journal of Mental Deficiency Research, 21*, 103-117.

Evans, D., & Hampson, M. (1968). The language of mongols. *British Journal of Disorders of Communication, 3*, 171-181.

Feldman, H., Goldin-Meadow, W., & Gleitman, L. (1978). Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock (Ed.), *Action, symbol and gesture: The emergence of language*. New York: Academic Press.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*.

Fischer, S. (1978). Language and creoles. In P. Siple (Ed.), *Understanding language through sign language research*. New York: Academic Press.

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

Fowler, A. (1980). *Effects of mental retardatiion on the acquisition of language*. Unpublished masters thesis, University of Pennsylvania.

Fowler, A. (1984). *Language acquisition in Down's syndrome children: Production and comprehension*. Doctoral dissertation, University of Pennsylvania.

Fowler, A., Gelman, R., & Gleitman, L. (1980). A comparison of normal and retardate language equated on MLU. Presented at the *5th Annual Boston University Conference on Child Language Development*.

Fowler, A., Gleitman, L., & Gelman, R. (1981). Towards an understanding of "delay without deviance" in retarded language acquisition. Presented at the *Biennial Meetings of the Society for Research in Child Development*, April, 1981.

Fowler, A., Gleitman, L., & Gelman, R. (forthcoming). Constraints on language learning in children with Down syndrome.

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, productions and comprehension. *Journal of Verbal Learning and Verbal Behavior, 2*, 121-135.

Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language, 6*, 432-442.

Gelman, R. (1982). Basic numerical abilities. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.

Gibson, D. (1966). Early developmental staging as a prophecy index in Down's Syndrome. *American Journal of Mental Deficiency, 70*, p. 825-828.

Gibson, D. (1978). *Down's syndrome: The psychology of mongolism*. London: Cambridge University Press.

Gillham, B. (1979). *The first words programme*. London: George Allen and Unwin.

Gleitman, L., Newport, E., & Gleitman, H. (1984). The current status of the Motherese hypothesis. *Journal of Child Language, 1*.

Gleitman, L. R., & Wanner, E. (1982). Language acquisition: the state of the art. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art*. Cambridge: Cambridge University Press.

Gleitman, L .R., & Wanner, E. (1984). Current issues in language learning. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook*. Hillsdale, NJ: Erlbaum Associates.

Goda, S., & Griffith, B. C. (1962). Spoken language of adolescent retardates and its relation to intelligence, age and anxiety. *Child Development, 33*, 489-498.

Graham, N. C., & Gulliford, R. A. (1968). A psychological approach to the language deficiencies of educationally subnormal children. *Educational Review, 20*, 136-145.

Greenwald, C. A., & Leonard, L. B. (1979). Communicative and sensorimotor development of Down's Syndrome children. *American Journal of Mental Deficiency, 84*, 296-303.

Haber, L., & Maloney, C. (1978). Linguistic development in mentally retarded children. In D. Farkas, W. Jacobsen, & K. Todrys (Eds.), *Papers from the 14th Regional Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press.

Hanzlik, J., & Stevenson, M. (1986). Interaction of mothers with their infants who are mentally retarded, retarded with cerebral palsy or non-retarded. *American Journal of Mental Deficiency, 90*, 513-520.

Harris, J. (1983). What does mean length of utterance mean? Evidence from a comparative study of normal and Down's syndrome children. *British Journal of Disorders of Communication, 18*, 153-169.

Hartley, X. Y. (1982). Receptive language processing of Down's Syndrome children. *Journal of Mental Deficiency Research, 26*, 263-269.

Heber, R. (1961). A manual on terminology and classification in mental retardation. *American Journal of Mental Deficiency, Monogr. Suppl.*,

Johnson, R. C., & Abelson, R. B. (1969). Intellectual, behavioral and physical characteristics associated with trisomy, translocation and mosaic types of Down's Syndrome. *American Journal of Mental Deficiency, 73*, 852-5.

Jones, O. H. M. (1977). Mother-child communication with prelinguistic Down's Syndrome and normal infants. In H. R. Schaffer & J. Dunn (Eds.), *The first year of life: Psychological and medical implications of early experience*. New York: Wiley.

Jones, O. H. M. (1980). Prelinguistic communication skills in Down's syndrome and normal infants. In T. M. Fields, S. Goldberg, D. Stern, & A. M. Sostek (Eds.), *High risk infants and children: Adult and peer interactions*. New York: Academic Press.

Kamhi, A., & Johnston, J. (1982). Towards an understanding of retarded children's linguistic deficiencies. *Journal of Speech and Hearing Research, 25*, 435-445.

Kean, M-L. (1977). The linguistic interpretation of aphasic syndromes: Agramatism in Broca's aphasia, an example. *Cognition, 5*, 9-46.

Keil, F. (1981). Constraints on knowledge and cognitive development. *Psychological Review, 88*, 197-227.

Klima, E., & Bellugi, U. (1966). Syntactic regularities in the speech of children. In A. Reibel & S. A. Schane (Eds.), *Modern studies in English*. Englewood Cliffs, NJ: Prentice Hall.

Krashen, S. (1975). The critical period for language acquisition and its possible causes. *Annals of the New York Academy of Sciences, 263*, 211-224.

Lackner, J. R. (1968). A developmental study of language behavior in retarded children. In D. M. Morehead & A. E. Morehead (Eds.), *Normal and deficient child language*. Baltimore: University Park Press.

Layton, T. L., & Sharifi, H. (1979). Meaning and structure of Down's Syndrome and non-retarded children's spontaneous speech. *American Journal of Mental Deficiency, 83*, 439-445.

Lee, L. (1974). *Developmental sentence analysis*. Evanston, IL: Northwestern University Press.

Leifer, J., & Lewis, M. (1984). Acquisition of conversational response skills by young Down Syndrome and non-retarded children. *American Journal of Mental Deficiency, 88*, 610-618.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Lenneberg, E. H., Nichols, I. A., & Rosenberger, E. F. (1964). Primitive stages of language development in mongolism. *Research Publications, Association for Research in Nervous and Mental Disease, 42*, 119-147.

Leuder, I., Fraser, W. I., & Jeeves, M. A. (1981). Social familiarity and communication in Down syndrome. *Journal of Mental Deficiency Research, 5*, 133-142.

Lovell, K., & Bradbury, B. (1967). The learning of English morphology in educationally subnormal special school children. *American Journal of Mental Deficiency, 71*, 609-615.

Lovell, K., & Dixon, E. M. (1967). The growth of the control of grammar in imitation, comprehension, and production. *Journal of Special Education, 3*, 275-284.

Lyle, J. G. (1959). The effect of an institution upon the verbal development of imbecile children. I. Verbal intelligence. *Journal of Mental Deficiency Research, 3*, 122-128.

Lyle, J. G. (1960). The effect of an institution upon the verbal development of imbecile children. II. Speech and language. *Journal of Mental Deficiency Research, 4*, 1-13.

Lyle, J. G. (1961). A comparison of normal and imbecile children. *Journal of Mental Deficiency Research, 5*, 40-51.

Mahoney, G., Glover, A., & Finger, I. (1981). Relationship between language and sensorimotor development of DS and nonretarded children. *American Journal of Mental Deficiency, 86*, 21-27.

McCarthy, J. (1965). *Patterns of psycholinguistic development of Mongoloid and non-Mongoloid severely retarded children*. Doctoral dissertation. Urbana: University of Illinois Press.

McCarthy, J., & Kirk, S. A. (1961). *The construction, standardisation and statistical characteristics of the Illinois Test of Psycholinguistics*, Urbana: University of Illinois Press.

McLeavey, B. et al. (1982). Non-retarded and mentally retarded children's control over syntactic structures. *American Journal of Mental Deficiency, 86*, 485-494.

McNutt, J. C., & Leri, S. M. (1979). Language differences between institutionalized and noninstitutionalized retarded children. *American Journal of Mental Deficiency, 83*, 339-45.

Mein, R. (1961). A study of the oral vocabularies of severely subnormal patients. *Journal of Mental Deficiency Research, 5*, 52-62.

Mein, R., & O'Connor, N. (1960). A study of the oral vocabularies of severely subnormal patients. *Journal of Mental Deficiency Research, 4*, 130-143.

Menyuk, P. (1969). *Sentences children use*. Cambridge, MA: The MIT Press.

Mervis, C. B. (in press). Early conceptual development of children with DS. In D. Cicchetti & M. Beeghly (Eds.), *Down syndrome: The developmental perspective*. New York: Cambridge University Press.

Miller, J. F. (1980). *Assessing language production in children: Experimental procedures*. Baltimore: University Park Press.

Miller, J. F. (in press). Language and communication characteristics of children with Down syndrome. In A. Crocker, S. Paschel, J. Rynders, & C. Tinghey (Eds.), *Down syndrome: State of the art*. Baltimore: Brooks Publishing.

Molfese, D. L. (1977). Infant cerebral asymmetry. In S. J. Segalowitz & F. A. Gruber (Eds.), *Language development and neurological theory*. New York: Academic Press.

Muir, J. (1903). An analysis of twenty-six cases of mongolism. *Archives of Pediatrics, 20*, 161.

Naremore, R., & Dever, R. (1975). Language performance of educable mentally retarded and normal children at five age levels. *Journal of Speech and Hearing Research, 18*, 82-95.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monograph of the Society for Research in Child Development, 38*, Serial 149.

Newfield, M., & Schlanger, B. (1968). The acquisition of English morphology by normal and educable mentally retarded children. *Journal of Speech and Hearing Research, 11*, 693-706.

Newport, E. L., (1982). Task specificity in language learning? Evidence from speech perception and American Sign Language. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art*. Cambridge: Cambridge University Press.

Newport, E. L., Gleitman, H., & Gleitman, L. (1977). Mother I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow and C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.

Newport, E. L., & Supalla, T. C. (1980). The structuring of language: clues from the acquisition of signed and spoken language. In U. Bellugi & M. Studdert-Kennedy (Eds.), *Signed and spoken language: Biological constraints on linguistic form*. Dahlem Konferenzem. Weinheim: Verlag Chemie.

Newfield, M., & Schlanger, B. (1968). The acquisition of English morphology by normal and educable mentally retarded children. *Journal of Speech and Hearing Research, 11*, 693-706.

Nisbet, J., Zanella, K., & Miller, J. (1984). An analysis of conversations among handicapped students and a nonhandicapped peer. *Exceptional Children, 51*, 156-162.

Owens, R. E., & MacDonald, J. D. (1982). Communicative uses of the early speech of nondelayed and Down syndrome children. *American Journal of Mental Deficiency, 86*, 503-510.

169-171.Peterson, G., & Sherrod, K. (1982). Relationship of maternal language to language development and language delay of children. *American Journal of Mental Deficiency, 86*, 391-398.

Petitto, L. (1985). Are signed languages acquired earlier than spoken languages? Presented at *Biennial Meetings of the Society for Research in Child Development*, Toronto.

Price-Williams, D., & Sabsay, S. (1979). Communicative competence among severely retarded persons. *Semiotica, 26-1/2*, 35-63.

Reynell, J. (1969). *Test Manual.* Reynell Developmental Language Scales. Experimental Edition. Windsor, Berks: National Foundation for Developmental Research.

Rogers, M. G. (1975). A study of language skills in severely subnormal children. *Child Care Health and Development, 1*, 113-126.

Rondal, J. A. (1975). Developpement du langage et retard mental: Un revue critique de la litterature en langue anglaise. *Annee Psychologique, 75*, 513-547.

Rondal, J. A. (1978a). Developmental sentence scoring procedure and the delay-difference question in language development of Down's Syndrome children. *Mental Retardation, 16*, 169-171.

Rondal, J. A. (1978b). Maternal speech to normal and Down's Syndrome children matched for mean length of utterance. *Monographs of the Association of Mental Deficiency, 3*, 193-265.

Rondal, J. A. (1980). Verbal imitation by Down Syndrome and nonretarded children. *American Journal of Mental Deficiency, 85*, 318-321.

Rondal, J. A., Lambert, J. L., & Sohier, C. (1981). Elicited verbal imitation in Down's Syndrome and other mentally retarded children: A replication and extension of Berry. *Language and Speech, 24*, 245-54.

Rosenberg, S. (1982). The language of the mentally retarded: Development, processes and intervention. In S. Rosenberg (Ed.), *Handbook of applied psycholinguistics*. Hillsdale, NJ: Erlbaum.

Ross, M. H., Galaburda, A. M., & Kemper, T. L. (1984). Down's syndrome: Is there a decreased population of neurons? *Neurology, 34*, 909-916.

Ryan, J. (1975). Mental subnormality and language development. In E. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development* (Vol. 2). New York: Academic Press.

Ryan, J. (1977). The silence of stupidity. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics: Developmental and pathological*. Ithaca, NY: Cornell University Press.

St. James-Roberts, I. (1981). A reinterpretation of hemispherectomy data without functional plasticity of the brain. *Brain and Language, 13*, 31-53.

Sankoff, G., & LaBerge, S. (1973). On the acquisition of native speakers by a language, *Kivung, 6*, 32-47.

Scarborough, H. S. (1985). Measuring syntactic development: The Index of Productive Syntax, Presented at the *Biennial Meetings of the Society for Research in Child Development*, Toronto.

Scherer, N. J., & Owings, N. O. (1984). Learning to be contingent: Retarded children's responses to their mothers' requests. *Language and Speech, 27*, 255-267.

Schlesinger, I. M., (1971). The production of utterances and language acquisition. In D. I. Slobin, (Ed.), *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press.

Seagoe, M. V. (1965). Verbal development in a mongoloid. *Exceptional Children, 31*, 269-73.

Semmel, M. I., & Dolley, D. G. (1971). Comprehension and imitation of sentences by Down's Syndrome children as a function of transformational complexity. *American Journal of Mental Deficiency, 75*, 739-745.

Share, J. B. (1975). Developmental progress in Down's Syndrome. In R. Koch & F. F. de la Cruz (Eds.), *Down's syndrome (mongolism): Research, prevention and management*. New York: Brunner/Mazel.

Share, J., Koch, R., Webb, A., & Graliker, B. (1964). The longitudinal development of infants and young children with Down's Syndrome (mongolism). *American Journal of Mental Deficiency, 68*, 685.

Shipley, E., Smith, C., and Gleitman, L. (1969). A study in the acquisition of language: Free responses to commands. *Language, 45*, 322-42.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development*. New York: Holt, Rinehart and Winston, 175-208.

Slobin, D. I. (1980). The repeated path between transparency and opacity in language. In U. Bellugi & M. Studdert-Kennedy (Eds.), *Signed and spoken language: Biological constraints on linguistic form*. Dahlem Konferenzem. Weinheim: Verlag Chemie, 229-243.

Sommers, R. K., & Starkey, K. L. (1977). Dichotic verbal processing in DS children having qualitatively different speech and language skills. *American Journal of Mental Deficiency, 82*, 44-53.

Spreen, O. (1965). Language functions in mental retardation: A review. I. Language development, types of retardation and intelligence level. *American Journal of Mental Deficiency, 69*, 482-494.

Strominger, A. Z., Winkler, M. R., & Cohen, L. T. (1984). Speech and language evaluation. In S. M. Pueschel (Ed.), *The young child with Down syndrome*. New York: Human Sciences Press, 253-261.

Thompson, M. M. (1963). Psychological characteristics relevant to the education of the pre-school mongoloid child. *Mental Retardation, 1*, 148-151.

Uzgiris, I., & Hunt, J. McV. (1975). *Assessment in infancy: Ordinal scales of psychological development*. Urbana: University of Illinois Press.

Weiss, B., Beeghly, M., & Cicchetti, D. (1985). Develoment of symbolic play in children with Down syndrome and nonhandicapped children. Presented at the *Biennial Meetings of the Society for Research in Child Development*, Toronto, April, 1985.

Wiegel-Crump, C. A. (1981). The development of grammar in Down's Syndrome children between the MA's of 2-0 and 6-11 years. *Education and Training of the Mentally Retarded, 6*, 24-60.

Wing, L. (1975). A study of language impairment in severely retarded children. In N. O. O'Connor (Ed.), *Language, cognitive deficits and retardation*. London: Butterworth.

Yoder, D. E., & Miller. (1972). What we may know and what we can do: Input toward a system. In J. E. McLean, D. E. Yoder, & R. L. Schiefelbusch (Eds.), *Language intervention with the retarded: Developing strategies*.

Zeaman, D., & House, B. (1962). Mongoloid MA is proportional to log CA. *Child Development, 33*, 481-488.

Zekulin-Hartley, X. Y. (1981). Hemispheric asymmetry in DS children. *Canadian Journal of Behavioral Science, 13*, 210-217.

Zigler, E. (1969). Developmental versus difference theories of mental retardation and the problem of motivation. *American Journal of Mental Deficiency, 73*, 536-556.

PUBLICATIONS APPENDIX

PUBLICATIONS

Baer, T., Gore, J. C., Boyce, S., & Nye, P. W. (1987). Application of MRI to the analysis of speech production. *Magnetic Resonance Imaging, 5, 1*, pp. 1-7.

Baer, T., Sasaki. C., & Harris, K. S. (Eds.). (1987). *Laryngeal function in phonation and respiration*. Boston: College-Hill/Little Brown.

Barac, D., Remez, R. E. & Rubin, P. E. (1986). Perceptual differences between spontaneous and read speech. *PAW Review, 1*, pp. 19-21.

Browman, C. P. & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*, 219-252.

Boyce, S., Browman, C. P. & Goldstein, L. (in press). Lexical organization and Welsh consonant mutations. *Journal of Memory & Language*.

Browman, C. P. (in press). The hunting of the quark: The particle in English. *Language & Speech*.

Browman, C. P., & Goldstein. L. (1986). Dynamic processes in linguistics: Casual speech and historical change. *PAW Review, 1*, pp. 17-18.

Colomb, E. (1986). L'attention. *Médicine et Hygiène, 44*, 2854-2858.

Feldman, L. B., & Fowler, C. A. (1987). The inflected noun system in Serbo-Croatian: Lexical representation of morphological structure. *Memory & Cognition, 15*, 1-12.

Harris, K. S. (1987). Speech production. In H. Levitt (Ed), *Encyclopedia of deafness and deaf persons*. New York: McGraw Hill.

Mann, V. A. (1986). Phonological awareness: The role of reading experience. *Cognition, 24*, 65-92.

Mann, V. A. (1986). Distinguishing universal and language dependent levels of speech perception: Evidence from Japanese listeners' perception of "*ℓ* and *r*". *Cognition, 24*, 169-196.

Repp, B. (in press). Integration and segregation in speech perception. *Proceedings of the Eleventh International Congress of Phonetic Sciences, Tallinn, Estonia, USSR.*

Repp, B. (1986). The sound of two hands clapping. *PAW Review, 1*, pp. 24-25.

Rosenblum, L. D., & Turvey, M. T. Fluctuation measures of the maintenance tendency. *PAW Review, 1*, pp. 13-16.

Schmidt, R. C., Carello, C., & Turvey, M. T. (1986). Perceiving the region of reversability. *PAW Review, 1*, pp. 5-10.

APPENDIX

Status Report		**DTIC**	**ERIC**
SR-21/22	January - June 1970	AD 719382	ED 044-679
SR-23	July - September 1970	AD 723586	ED 052-654
SR-2	October - December 1970	AD 727616	ED 052-653
SR-25/26	January - June 1971	AD 730013	ED 056-560
SR-27	July - September 1971	AD 749339	ED 071-533
SR-28	October - December 1971	AD 742140	ED 061-837
SR-29/30	January - June 1972	AD 750001	ED 071-484
SR-31/32	July - December 1972	AD 757954	ED 077-285
SR-33	January - March 1973	AD 762373	ED 081-263
SR-34	April - June 1973	AD 766178	ED 081-295
SR-35/36	July - December 1973	AD 774799	ED 094-444
SR-37/38	January - June 1974	AD 783548	ED 094-445
SR-39/40	July - December 1974	AD A007342	ED 102-633
SR-41	January - March 1975	AD A013325	ED 109-722
SR-42/43	April - September 1975	AD A018369	ED 117-770
SR-44	October - December 1975	AD A023059	ED 119-273
SR-45/46	January - June 1976	AD A026196	ED 123-678
SR-47	July - September 1976	AD A031789	ED 128-870
SR-48	October - December 1976	AD A036735	ED 135-028
SR-49	January - March 1977	AD A041460	ED 141-864
SR-50	April - June 1977	AD A044820	ED 144-138
SR-51/52	July - December 1977	AD A049215	ED 147-892
SR-53	January - March 1978	AD A055853	ED 155-760
SR-54	April - June 1978	AD A067070	ED 161-096
SR-55/56	July - December 1978	AD A065575	ED 166-757
SR-57	January - March 1979	AD A083179	ED 170-823
SR-58	April - June 1979	AD A077663	ED 178-967
SR-59/60	July - December 1979	AD A082034	ED 181-525
SR-61	January - March 1980	AD A085320	ED 185-636
SR-62	April - June 1980	AD A095062	ED 196-099
SR-63/64	July - December 1980	AD A095860	ED 197-416
SR-65	January - March 1981	AD A099958	ED 201-022
SR-66	April - June 1981	AD A105090	ED 206-038
SR-67/68	July - December 1981	AD A111385	ED 212-010
SR-69	January - March 1982	AD A120819	ED 214-226
SR-70	April - June 1982	AD A119426	ED 219-834
SR-71/72	7July - December 1982	AD A124596	ED 225-212
SR-73	January - March 1983	AD A129713	ED 229-816
SR-74/75	April - September 1983	AD A136416	ED 236-753
SR-76	October - December 1983	AD A140176	ED 241-973
SR-77/78	January - June 1984	AD A145585	ED 247-626
SR-79/80	July - December 1984	AD A151035	ED 252-907
SR-81	January - March 1985	AD A156294	ED 257-159
SR-82/83	April - September 19857	AD A165847	ED 266-508
SR-84	October-December 1985	AD A168819	ED 270-831
SR-85	January-March 1986	AD A173677	ED 274-022
SR-86/87	April-September 1986	AD A176816	ED 278-066

Information on ordering any of these issues may be found on the following page.

AD numbers may be ordered from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22151

ED numbers may be ordered from:

ERIC Document Reproduction Service
Computer Microfilm Corporation
3900 Wheeler Ave.
Alexandria, Virginia 22304-5110

Haskins Laboratories Status Report on Speech Research is abstracted in *Language and Language Behavior Abstracts*, P.O. Box 22206, San Diego, California 92122.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER SR-88, 1986	**2. GOVT ACCESSION NO.**	**3. RECIPIENT'S CATALOG NUMBER**
4. TITLE (and Subtitle) Haskins Laboratories Status Report on Speech Research		**5. TYPE OF REPORT & PERIOD COVERED** Continuation Report October-December, 1986
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Staff of Haskins Laboratories: Michael Studdert-Kennedy, President		**8. CONTRACT OR GRANT NUMBER(s)** HD-01994 BNS-8520709 NO1-HD-5-2910 NS 13870 RR-05596 NS 13617 NS 18010
9. PERFORMING ORGANIZATION NAME AND ADDRESS Haskins Laboratories 270 Crown Street New Haven, CT 06511-6695		**10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS**
11. CONTROLLING OFFICE NAME AND ADDRESS National Institutes of Health National Science Foundation		**12. REPORT DATE** December, 1986
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) As above		**15. SECURITY CLASS.** (of this Report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

UNLIMITED: Contains no information not freely available to the general public. It is distributed primarily for library use.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

As above

18. SUPPLEMENTARY NOTES

N/A

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Speech Perception:

Lexical organisation, consonants, Welsh, mutations, cerebral asymmetries, cognitive function, language structure, children, Down Syndrome, coarticulation, fricatives, adults, perceptual normalisation, vowels, sinusoidal voices, stop-glides, acoustic analysis, syllable, amplitude envelope, initial /b/, initial /w/.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report (1 October-31 December) is one of a regular series on the status and progress of studies on the nature of speech, instrumention for its investigation, and practical applications. Manuscripts cover the following topics:

- Lexical organisation and Welsh consonant mutations
- The emergence of cerebral asymmetries in early human development:
- A literature review and a neuroembryological model
- The role of coarticulatory effects in the perception of fricatives by children and adults

DD FORM 1 Jan 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19. Key Words (Continued)

Speech Articulation:

Motor Control:
Phase-locked, modes, phase, transitions, oscillators, motion,
anatomical specificity,

Reading:
Hemispheric asymmetries, phonological processing, lexical ambiguity,
processing, visual, word recognition, orthography, word order,
Serbo-Croat, short-term memory, reading ability

20. Abstract (Continued)

- Hemispheric asymmetries in phonological processing
- Processing lexical ambiguity and visual word recognition in a deep orthography
- Some word order effects in Serbo-Croat
- Short-term memory, phonological processing and reading ability
- Phase-locked modes, phase transitions, and component oscillators in biological motion
- Beyond anatomical specificity
- Perceptual normalisation of vowels produced by sinusoidal voices
- The stop-glide distinction: Acoustic analysis and perceptual effect of variation on syllable amplitude envelope for initial /b/ and /w/.
- The development of language structure in children with Down syndrome