LETTERS TO THE EDITOR


The article by McGarvey and Weinberg (JSHD, August 1984) on esophageal insufflation testing in nonlaryngectomized adults contains numerous statements that merit clarification. We respectfully submit our point of view in this letter to the editor because the authors chose not to incorporate any of the suggestions solicited from us during preparation of their manuscript.

The essence of our disagreement with McGarvey and Weinberg is not with the design of their study but with their extrapolation of observations made on normals to laryngectomized individuals. Stated simply, the authors observed that transnasal esophageal insufflation in nonlaryngectomized individuals produced airtight closure of the criopharyngeal sphincter and failed to produce sustained esophageal voice. They defined this as a normal response. The authors incorrectly extend this observation implying in their theoretical and clinical implications section (McGarvey & Weinberg, 1984, p. 275) and specifically in their abstract that "(b) laryngectomized patients having airtight closure of the pharyngoesophageal segment during insufflation testing exhibit a normal esophageal response" (p. 272). The two populations are not comparable considering substantial evidence that the pharyngoesophagus following laryngectomy is no longer anatomically or physiologically "normal" (Kirchner, Dey, & Shedd, 1963; Stiernberg & Bailey, 1985; Welch, Luckmann, Ricks, Drake, & Gates, 1979).

By incorrectly assuming that the inability to generate continuous esophageal voicing on transnasal insufflation is a "normal response" in both nonlaryngectomized and laryngectomized individuals, McGarvey and Weinberg confuse our work on selective pharyngeal constrictor myotomy (Singer & Blom, 1980, 1981; Singer, Blom, & Hamaker, 1981). They conclude from their observations that "unilateral selective myotomy may have been performed on laryngectomized individuals who exhibited esophageal responses similar to those elicited in nonlaryngectomized persons and, by inference, who exhibited minimal alteration or compromise in esophageal function" (McGarvey & Weinberg, 1984, p. 276). This statement suggests that we advocate pharyngeal constrictor myotomy for laryngectomized persons who exhibit what the authors consider "normal" esophageal function, which is incorrect; and the implications of their interpretation are unacceptable.

Mark I. Singer
Eric D. Blom
Indianapolis, IN

REFERENCES


Singer and Blom: A Reply

We appreciate the opportunity to reply to the brief letter of Singer and Blom have prepared in response to our recent paper. First, Singer and Blom have suggested that there is something "incorrect" about defining responses of normal individuals to esophageal air insufflation as normal responses. We fail to see any major incoherency or error in this operational procedure or definition. Moreover, we continue to interpret the results of our work "as supporting the view that normal pharyngoesophageal segment function represents an influence detrimental to the acquisition and ultimate production of functionally serviceable esophageal speech" (McGarvey & Weinberg, 1984, p. 275). Second, Blom and Singer indicate that we "incorrectly extend" our observations, although they acknowledge acceptance of our study design and findings. Specifically, they take issue with our conclusion that "laryngectomized patients having airtight closure of the pharyngoesophageal segment during insufflation exhibit a normal esophageal response" (McGarvey & Weinberg, 1984, p. 272). We fail to understand the problem. As we said in the paper, "laryngectomized patients exhibiting airtight closure of the PE segment or failure to produce voice continuously during insufflation testing exhibit responses that correspond to those exhibited by each of the nonlaryngectomized subjects" (pp. 275–276). What is unclear about this statement?

Singer and Blom take the position that the two groups (normals and laryngectomized) "are not comparable considering the substantial evidence that the pharyngo-esophagus following laryngectomy is no longer anatomically or physiologically 'normal.' " We were aware of the differences between these two groups. For example, as part of our discussion of selective myotomy we wrote,

The statement that unilateral selective myotomy may have been performed on laryngectomized individuals exhibiting functionally typical esophageal responses should not

Received October 22, 1984
Accepted January 3, 1985

287
be interpreted to imply that such individuals have anatomically normal upper esophageal segments. They do not. The upper esophageal segment of laryngectomized persons is altered, and we continue to view this segment as a surgical residue characterized by considerable variability (italics added) (Weinberg, 1980). Although the pharyngo-esophageal segments of laryngectomized individuals have been surgically altered, the voicing responses of some laryngectomized patients during air insufflation testing correspond with those exhibited by many nonlaryngectomized individuals. This correspondence suggests to us that some laryngectomized patients exhibit typical esophageal responses and minimal alterations or compromise in esophageal function. (McGarvey & Weinberg, 1984, p. 276)

Singer and Blom confuse two basic issues. Laryngectomized patients do not have anatomically normal upper esophageal segments. However, despite alteration of the upper esophagus, some laryngectomized patients exhibit normal voicing responses during esophageal insufflation testing. Stated differently, functionally normal esophageal responses are possible for some laryngectomized individuals as well as normal individuals. We believe Singer and Blom fail to understand this distinction.

Finally, there is the issue of selective myotomy. Singer and Blom indicate that we "confused" their work on selective myotomy. We explicitly described the rationale Singer and Blom have provided for advocating unilateral selective myotomy. The implications of these results and interpretations are both clear and undeniable reality.

We fail to perceive anything "confusing," "unacceptable," or "incorrect" in this concise summary of the Singer and Blom rationale for advocating myotomy.

Singer and Blom indicated that we incorrectly concluded that "unilateral selective myotomy may have been performed on laryngectomized individuals who exhibited esophageal responses similar to those elicited in nonlaryngectomized persons, and by inference, who exhibited minimal alteration or compromise in esophageal function" (McGarvey & Weinberg, 1984, p. 276). What is incorrect about this conclusion or inference? There is no question about the fact that selective myotomy has been performed on laryngectomized patients who exhibit esophageal responses comparable to those exhibited by normals. Singer and Blom may find this situation "unacceptable," but it is an undeniable reality.

Interpretations of our data were offered primarily to clarify the rationale for advocating unilateral selective myotomy. The implications of these results and interpretations are both clear and acceptable. As we stated, an acceptable rationale for advocating myotomy is to increase the likelihood of tissue oscillation and voicing. This is an important goal which provides a reasonable rationale for advocating surgical modification of the upper esophageal segment" (McGarvey & Weinberg, 1984, p. 276). This goal must be balanced against possible negative consequences such as increasing the probability of reflux and so forth. This goal does clearly differ from the one implied by Singer and Blom (1980, 1981; Singer et al., 1981) that surgical treatment is advocated to correct an abnormal physiological situation or response.

Singer and Blom indicate that we "confused" their work on selective myotomy. We explicitly described the rationale Singer and Blom have provided for advocating unilateral selective myotomy. The implications of these results and interpretations are both clear and undeniable reality.

We fail to perceive anything "confusing," "unacceptable," or "incorrect" in this concise summary of the Singer and Blom rationale for advocating myotomy.

Singer and Blom indicated that we incorrectly concluded that "unilateral selective myotomy may have been performed on laryngectomized individuals who exhibited esophageal responses similar to those elicited in nonlaryngectomized persons, and by inference, who exhibited minimal alteration or compromise in esophageal function" (McGarvey & Weinberg, 1984, p. 276). What is incorrect about this conclusion or inference? There is no question about the fact that selective myotomy has been performed on laryngectomized patients who exhibit esophageal responses comparable to those exhibited by normals. Singer and Blom may find this situation "unacceptable," but it is an undeniable reality.

Interpretations of our data were offered primarily to clarify the rationale for advocating unilateral selective myotomy. The implications of these results and interpretations are both clear and acceptable. As we stated, an acceptable rationale for advocating myotomy is to increase the likelihood of tissue oscillation and voicing. This is an important goal which provides a reasonable rationale for advocating surgical modification of the upper esophageal segment" (McGarvey & Weinberg, 1984, p. 276). This goal must be balanced against possible negative consequences such as increasing the probability of reflux and so forth. This goal does clearly differ from the one implied by Singer and Blom (1980, 1981; Singer et al., 1981) that surgical treatment is advocated to correct an abnormal physiological situation or response.
myself questioning the maxim that more is necessarily better. Large-N, multivariate studies have some serious limitations that are often not considered by investigators. For example, the measures used in these types of studies are usually not sufficiently sensitive to unravel the types of complex relationships that govern patterns of language and language-related behaviors.

Theoretic Shortcomings

One of the frequent pitfalls of large N, multivariate studies is a weak or unacknowledged theoretic context. The desire to obtain as many subjects as possible somehow manages to overshadow the importance of acknowledging underlying theoretic contexts. Thomas (1979) has written that theory provides both an indication of which facts (variables) are important to study and an explanation of how in which particular facts (variables) relate to one another. Because investigators study what they believe are important facts or variables, all research can be viewed as theory motivated. However, investigators do not always acknowledge the theory that underlies their research.

A major theoretic shortcoming in Schery's study was the absence of some acknowledgment that the composite language measure, the only dependent variable in the study, reflected Schery's theory of what language is. This measure was compiled from 14 tests that assess expressive and receptive language, reading, memory, and articulation abilities. The composite measure generally reflected high factor loadings. Yet, though the reading and articulation measures had only moderate factor loadings, they were also included in the composite measure because, as Schery put it, they added to "the concept of a general language ability" (Schery, 1985, p. 75). It is important to recognize, however, that Schery's "concept" of a general language ability represents a particular theoretic view of what language is. There are those who would question whether a general language concept should include measures of memory, reading, and articulation and not include measures of pragmatic abilities. Importantly, it does not matter so much how one defines language, so long as one acknowledges that the definition reflects a particular theoretic view of language and interprets findings relative to this theoretic view.

The absence of Schery's theoretic stance on language was compounded by the general atheoretic nature of her study. At one point in the study Schery seemed to express some surprise that "an extremely wide range of factors were represented by the IQs of the children," suggesting that "a professional staff could think about the nature and origins of language-disordered children's difficulties" (Schery, 1985, p. 80), did not account very well for language performance or gain. But as with Schery's concept of language, this range of variables reflected a particular theoretic view of the factors that contribute to language disorders. If the negative findings were interpreted within this theoretic view, Schery would have been more likely to question the usefulness and appropriateness of this theoretic view. Moreover, this type of questioning might have led Schery to consider whether a different theoretic orientation, such as one based on information processing theory or social learning theory, would not be highly correlated to language performance or language gain. In fact, given the 13-year age range of the children it seemed a statistical impossibility for anything but age and pretest language performance, which is strongly correlated to age, to be the best predictors of current language performance and language gain. The methodology of the study thus precluded Schery from discovering strong relationships between the five cluster variables and the measures of language performance.

Suggestions for Future Research

Future attempts to discover correlates of language development in language-disordered children need to take into consideration some of the points raised in this letter. I think we can all learn from Schery's efforts that bigger is not necessarily better. Smaller studies that control important sources of variability can go a long way toward unraveling the correlates of language development in disordered populations. If one has a large data pool, however, one must be particularly careful to make sure that theoretic contexts are well formulated and explicitly stated and that subject samples do not contain obvious sources of variability.

Due in part to the findings from Schery's study, I believe that we have reached a point in our thinking about language disorders that should lead us to ask somewhat different questions than the ones Schery asked. I think there is little question that the five cluster variables Schery investigated are related to language performance. Schery cites several studies attesting to these relationships. There is some question, however, about how these relationships change with age and the exact nature and function of the specific processes (e.g., perceptual, conceptual, representational, memory, symbolic) that underlie these relationships. For example, knowing that a child's cognitive abilities (IQ) predict language performance tells us little about the specific cognitive processes the child is tapping to solve cognitive and linguistic problems. I think that future studies should be designed to address these types of questions.

Let me conclude by saying that I think Schery's study was an important one, not so much in what it found, but in what it did not find. As is evident in this letter, the negative findings that emerged from Schery's ambitious research effort have led me to question the value of large N, multivariate studies as well as to think about some different questions to ask about the correlates of language development in language-disordered children. Negative findings should always be so stimulating!
Reply to Kamhi

I think that it is important to respond to Alan Kamhi’s critique of my paper, “Correlates of Language Development in Language-Disordered Children” (JSCHD, February 1985), because, in my experience, his view represents a widespread misunderstanding in our field about the kind of research this study represents. Philosophers of science distinguish the logic of discovery from the logic of confirmation (or hypothesis testing). One valid type of study is to have a structured theory and then to test it. The strength of this approach is its contribution to confirming or falsifying ideas/hypotheses that are already well developed. The weakness is that it is limited only to the few variables specified by theory and does not examine (i.e., allow for discovery of) other important factors not represented in a particular theory. Another type of study begins with few theoretical preconceptions and casts a broad net of variables in search of relationships. The strength of such an approach is the opportunity for discovery of patterns that can become the basis for new theory. Its weakness is that it offers no direct test of any given theory. Both approaches are important. We shouldn’t debate whether one is universally better, but rather, how we can combine approaches in the optimal form to advance our understanding of issues.

Kamhi’s observation that my study has some implicit theory is correct, of necessity, but the theory is of a very general nature, and the purpose of the study was not to test it but to use it as a general frame for discovery. Moreover, this implicit theory was not an idiosyncratic construction on my part but an amalgamation of assumptions commonly made by practitioners in the field during the decade of the 1970s as they acted to decide what information was needed to diagnose, serve, and evaluate progress in language-disordered children. My study took the broad, “discovery” approach using variables based upon these widely held suppositions about what was relevant to these children. The fact that few relationships were “discovered” may be disappointing (and I certainly agree with Kamhi here), but it hardly invalidates the entire approach. Indeed, it usefully challenges the implicit assumption some make that we practitioners (at least up to the 1980s) have identified the important predictive variables in our work with language-disordered children.

Some specific comments are in order regarding Kamhi’s notions of the methodological shortcomings of the study. Available social science research tools are such that broad discovery approaches generally must rely on correlational methodologies, whereas confirmation or hypothesis testing relies on experimen-
eloquently summarized by Messick (1980), an eminent authority in this field. The admonishment is based upon the centrality of construct validity and its ethical ramifications in evaluating and using human assessment instruments and procedures. Basically, evaluations of tests in terms of psychometric parameters that have not been shown to be legitimate extensions of construct validity are of questionable value. Indeed, such evaluations can yield results that may contradict pertinent underlying constructs. Moreover, such evaluations raise major issues about the ethics of assessment (Messick, 1980).

Compelling or Not?

The purpose of this letter is to substantiate that McCauley and Swisher knowingly ignored this admonishment, resulting in a "review" of language and articulation tests that was not only of questionable value but was potentially harmful by supporting some tests with doubtful construct validity.

McCauley and Swisher (1984a) stated:

Three kinds of validity have usually been considered important for any test that measures behavior and is used to make inferences about underlying abilities—construct, content, and criterion-related validity (APA, 1974, pp. 24–55). Although Messick (1980) has presented a compelling case for considering construct validity as the keystone of test development, these three kinds of validity have often been defined independently and evidence about them for a given test is weighed jointly.

Construct validity refers to the degree to which a test measures the theoretical construct it is intended to measure (Anastasi, 1976, p. 151). Construct validity is examined by a careful comparison of a test author's delineation of the construct to be tested to the test's actual content. Because the evaluation of construct validity is difficult and somewhat subjective, it was not conducted for tests in this review. (p. 35)

Thus, they indicate that an evaluation of construct validity was not conducted for the tests in their review; yet, they cite Messick (1980) as presenting "a compelling case for considering construct validity as the keystone of test development."

Expediency and Holy Trinity

The plot thickens on two grounds, expediency and the Holy Trinity problem. McCauley and Swisher (1984a) turned to expediency to dismiss considerations of construct validity for the tests under consideration. They said, "Because the evaluation of construct validity is difficult and somewhat subjective, it was not conducted for tests in this review" (p. 35). Recently, Ringel, Trachtman, and Prutting (1984) discussed the scientific status of human communication sciences. They were concerned that expediency may be a predating factor for some research. McCauley and Swisher (1984a) is a case in point.

McCauley and Swisher held that subjectivity associated with construct validity is a sufficient reason to dismiss considerations of construct validity. Such reasoning is capricious. It should be acknowledged that all of validity is subjective and that subjectivity is inherent in the scientific enterprise (Prutting, 1983). It is rigor and discipline that define its value. The subjectivity of validity can be appreciated by the following comment from Messick (1980): "It is important to note that validity is itself inferred, not measured." (p. 1014).

Regarding the Holy Trinity problem, McCauley and Swisher (1984a) again provide a case in point. They cite APA (1974, pp. 25–55) and state that three kinds of validity have usually been considered important for any test. These kinds of validity are: construct, content, and criterion related. They continue by citing Messick (1980) which, interestingly enough, specifically cites the Holy Trinity problem about such practices. Then, they say that these three kinds of validity have often been defined independently, and evidence about them for a given test is weighed jointly.

Evidently, McCauley and Swisher believe that common (usual, often) practice a decade ago is sufficient reason to override current paradigms even in the face of contradictory admonishments (Guion, 1977, 1980; Messick, 1980). We could call this the same-song-second-verse logic; but unfortunately, the appropriate literature had rewritten the song that defines the McCauley and Swisher review as discordant with the literature on test evaluation.

The Holy Trinity problem deals with misunderstandings about "types" of validity and presumed options for dealing with some types while laying construct validity aside. Drawing upon Dunnette and Borman (1979) and Guion (1980), Messick (1980) summarized this problem.

Dunnette and Borman (1979) lament, to "perpetuate a conceptual compartmentalization of 'types' of validity—criterion-related, content, and construct...the implication that validities come in different types leads to confusion and, in the face of confusion, oversimplification" (p. 483). One consequence of this simplism is that many test users focus on one or another of the types of validity, as though any one would do, rather than on the specific inferences they intend to make from the scores. There is an implication that once evidence of one type of validity is forthcoming, one is relieved of responsibility for further inquiry. Indeed, the "Uniform Guidelines" seem to treat the three types of validity, in Guion's (1980) words, "as something of a Holy Trinity representing three different roads to psychometric salvation. If you can't demonstrate one kind of validity, you've got two more chances!" (p. 1014)

The problem was amplified further by Messick (1980): "I emphasized the importance of construct validity for test use as well, arguing 'that even for purposes of applied decision making, reliance upon criterion validity or content coverage is not enough' (Messick, 1975, p. 956)." (p. 1013).

This is precisely what McCauley and Swisher did. They dismissed construct validity, then turned to content validity (face validity), criterion-related validity (concurrent and predictive validity), and other psychometric descriptions. Such dismissal of construct validity created the Holy Trinity problem. Indeed, Messick (1980) pointed out that so-called content validity is not validity anyway: "In the case of content coverage, they are not validity at all" (p. 1015).

Implications: Ethics

So what if McCauley and Swisher's (1984a) review (a) relied on expediency, (b) exemplified the Holy Trinity problem, (c) relied on "usual" practices of the past, and (d) indicated subjectivity. Does it matter? Unfortunately, it does matter. As Messick (1980) pointed out, such problems have ethical or social consequences. These ethical or social consequences are that claims are made about the values of tests, but the tests themselves may actually not warrant such claims. This becomes an ethical dilemma because unwarranted claims lead to erroneous assessments of an individual's capacity. The ethical or social consequences of such problems are evident in the findings of the McCauley and Swisher (1984a) review, and they have been pointed out elsewhere (Muma, 1983, 1984). McCauley and Swisher used Table 2 to show how their reviewed tests meet each of 10 psychometric criteria. Based upon these findings the TOLD test meets eight criteria, and the ITTP and PPVT-R tests meet each of five criteria. These were the top three tests in meeting these criteria. The implications are that these three tests are better than the others considered.

The irony, of course, is that these three tests are notoriously weak in construct validity. Had McCauley and Swisher abided by the pertinent literature on test evaluation, their findings would have been substantively different. The use of these, and the other similar tests, would have been seriously challenged. But, unfortunately, McCauley and Swisher (1984a) created the
problem they themselves cautioned against. "When given no information about a psychometric characteristic, the test user is realistically left to wonder whether or not a test is invalid and unreliable for his or her purposes. Stated differently, no news is bad news" (p. 41). One could suggest rewording their statement as follows: "When given information that does not consider construct validity, the test user is realistically duped about a test's actual value, which in turn may raise ethical issues when such tests are used." To use their phrase, "no news is bad news."

Given this circumstance, it is especially interesting that McCauley and Swisher (1984b) have a second paper on the use and misuse of norm-referenced tests in clinical assessment. This recent effort appears to be a belated attempt at theory brokerage (Volpe, 1981). I would hope such attempts would be predicated on the contemporary views of both pertinent fields, psycholinguistics and measurement and evaluation. If so, a theory broker would forward a posteriori descriptive communicative assessment approaches and the centrality of construct validity.

The following quotes amplify the centrality of construct validity and underscore the point that a consideration of construct validity is not optional in the evaluation of tests and assessment procedures. "All measurement should be construct referenced" (Messick, 1975, p. 957; 1980, p. 1015). "All validity is at its base some form of construct validity... It is the basic meaning of validity" (Guion, 1977, p. 410).

Operationism

Coombs (1954) used the term "operationism in reverse" (p. 476). Underlying constructs are supposed to provide the substantive basis for generating an assessment instrument. However, some measures may be generated by expediency, then claims might be made about their presumed construct validity. The PICA is an example. As Coombs put it, operationism in reverse is "endowing the measures with all the meanings associated with the concept" (p. 476). Bartlett (1975) raised a similar point regarding language intervention programs. It was that many programs, we could say many tests such as those reviewed by McCauley and Swisher (1984a), Darley (1979), and Aram and Nation (1982), merely make clinicians operational, but their actual values for clients remain to be demonstrated.

Muma (1985) surveyed several popular language assessment tests (PPVT, ITPA, PICA, CELF, ACLC, TACL, DSS, TOLD). Without going into specifics, two general statements can be made. One is that only two tests entertained the notion of a theoretical basis; moreover, the issue of criteria for "operationizing" underlying constructs was less than adequately met by these tests. Such criteria included: relativity, conditionality, complexity, dynamism, ecology, individual differences, and a priori as contrasted to a posteriori (Beckwith, Rispoli, & Bloom, 1984) assessment. In short, an examination of the formal language tests has revealed a "naked scarecrow," to borrow Bruner's (1981) term, McCauley and Swisher's review notwithstanding.

As a footnote, let us correct some other less important errors in the McCauley and Swisher (1984a) article. They said,

Despite recent attacks on the appropriateness of norm-referenced tests for any assessment device (Muma, 1981; Muma, Lubinski, & Pierce, 1982; Muma & Muma, 1979), such tests are generally considered appropriate and necessary for the accomplishment of the first of these objectives—the identification of a speech or language impairment (Bloom & Lahey, 1978, p. 331; Launer & Lahey, 1981; Schery, 1981; Stark, Tallal, & Mellits, 1982). (p. 34)

This is a major misrepresentation of my work, and some relevant information of some of the others has been omitted that makes us return to their theme, "no news is bad news."

Although it is true that I have cited the psycholinguistic literature to advocate descriptive assessment procedures in language assessment, McCauley and Swisher are inaccurate when they claim that I do not see any appropriateness for norm-referenced tests. They need only refer to the references that they themselves cite to learn that I explicitly state that such tests are useful in determining the problem/no problem issue. Furthermore, I would quibble with their claim that "such tests are generally considered appropriate and necessary" (McCauley & Swisher, 1984a, p. 34). One merely needs to consider what the major authorities in psycholinguistics do when they try to assess a child's skill. It is instructive to learn that the major scholars (Bates, 1979; Bloom, 1973; Bowerman, 1973; Brown, 1973; Bruner, 1981; Greenfield & Smith, 1976; Nelson, 1973, etc.) rarely use such tests, but employ descriptive procedures. Further, McCauley and Swisher (1984a) implicate a considerable cleavage between my work and Bloom and Lahey, for example. They should have said that in addition to norm-referenced tests, Bloom and Lahey, Muma, and others argue for the need for a function-based language assessment. As it stands, McCauley and Swisher leave a rather misleading impression.

Also, if I had not considered the use of norm-referenced tests as they intimate, then how do they explain my discussion of assessment power and precision of such tests in the references they cite? My discussion is very close to theirs on page 36. Then, we agree on a curious issue on page 39 where they dispense with a consideration of norm-referenced tests of auditory discrimination because Locke (1980) and Rees (1973) raised some questions about their value. Although I agree these tests are of questionable value, some of the tests they reviewed are also of questionable value once construct validity is considered. Why were they not similarly dispatched?

Finally, there is a contradiction between Table 2 and Figure 1 in the McCauley and Swisher (1984a) article. Figure 1 indicates that 2 of the 30 reviewed tests meet six and seven criteria, respectively. However, both the text and Table 2 do not verify this.

The Bottom Line

McCauley and Swisher make two constructive suggestions about the impoverished state of norm-referenced tests. They suggest that test users become more aware of psychometric principles and the flaws of the tests they are using. Presumably such knowledge will reduce the impact of these tests and particular reviews of language tests on clinical decisions. I heartily agree.

Second, they point to the potential influence of test buyers. They urged consumers to purchase tests that provide empirical evidence about reliability and validity. They infer that reliability and validity have been expendable luxuries in the current assessment literature; yet, curiously, they, themselves, placed the central issue—construct validity—on the expendable list.

John R. Muma
Texas Tech University, Lubbock

REFERENCES


BLOOM, L. (1973). One word at a time: The use of single-word


Received March 26, 1984
Accepted March 1, 1985

Reply to Muma (1985)

The purpose of the report by McCauley and Swisher (1984a) was to "stimulate discussion of the psychometric characteristics of language and articulation tests rather than to serve as a definitive psychometric review" (p. 34). It is clear from Muma's letter that such discussion has begun.

Muma's chief criticism of our report centers around our failure to include in the review procedure a criterion related to construct validity. In order to emphasize the fact that our report was not a "definitive psychometric review," we directed the reader's attention to its limited scope: "The criteria used in this review are a selected sample of a larger number of important psychometric criteria. Adherence to more numerous and, in some cases, stricter guidelines is commonly considered necessary for a well-developed norm-referenced test" (McCauley & Swisher, 1984a, p. 34). Although two criteria related to validity were included in the review procedure, neither was directly related to construct validity. We decided not to include a criterion directly related to construct validity because psychometric characteristics addressed in our review procedure were chosen not only "because of their recognized importance and relevance to tests of language and articulation," but also "because they could be translated into relatively objective decision rules" (McCauley & Swisher, 1984a, p. 37). Anastasi (1976) describes construct validity as involving "a broader, more enduring, and more abstract kind of behavioral description" than other aspects of validity and notes that its evaluation "requires the gradual accumulation of information from a variety of sources" (p. 151). Thus, although Muma labels three tests included in our review as "notoriously weak in construct validity" without describing the basis for that label, we decided not to use a small number of decision rules to address the "difficult and somewhat subjective" (McCauley & Swisher, 1984a, p. 35) evaluation of construct validity in our review.

Muma's criticism of the lack of a criterion addressing construct validity was in part related to its leading to the support of "some tests with doubtful construct validity." In making this claim, Muma overlooked the fact that the data in our paper were displayed so as to highlight the criteria being considered rather than the performance of individual tests. The names of tests meeting any specific number of criteria were not mentioned in the text and could only be determined through the reader's retabulation of data in Table 2, or, in one case, reference to the Appendix. Although we called attention to the small number of criteria met by most tests, we did not single out tests meeting more criteria as possessors of spotless psychometric virtue. Thus, we did not fall into the practice, described by Messick (1980) and referred to by Muma, of considering any single piece of evidence regarding validity as tantamount to proof of validity. In summary, the precautions we exercised in displaying and discussing our data were within the realm of good practice and were consonant with our stated purpose of stimulating "discussion of the psy-
chometric characteristics of language and articulation tests" rather than providing a "definitive psychometric review" (McCauley & Swisher, 1984a, p. 34).

In addition to the criticisms addressed above, Muma referred to what he considered to be two "less important errors" in our report. The first of these was a reference to several of his works as containing "attacks on the appropriateness of norm-referenced tests for any assessment objective" (McCauley & Swisher, 1984a, p. 34). Although we welcome his clarification of his views and although we, too, believe that norm-referenced tests are inappropriate for some assessment purposes (e.g., see McCauley & Swisher, 1984b), we stand by the reasonableness of our original interpretation. Thus, we believe that we are correct in interpreting passages such as the following as "attacks": "The psychometric model violates basic principles of validity whereas the descriptive model provides a better means for appropriately dealing with validity" (Muma, 1981, p. 70). "In short, the descriptive procedures will provide relevant data—evidence—whereas the psychometric battery will provide merely numbers that have relatively little relevance—data" (Muma, Lubinski, & Pierce, 1982, p. 145).

The second "less important error" that Muma purports to have found in our paper was "a contradiction between Table 2 and Figure 1." A reexamination of the legend for Figure 1 should resolve this misreading. That figure is a bar graph that indicates the frequency with which tests met one or more criteria, two or more criteria, and so forth. Thus, a test meeting eight criteria is represented in the bars for one or more criteria met, two or more criteria met, and so forth. Muma's comments suggest that he believed the graph contained simple frequency information.

We welcome Muma's discussion of positions held by those who address the primacy of construct validity in considerations of test validity and development (e.g., Guion, 1977; Messick, 1975, 1980). Particularly valuable is his condemnation of what might be termed an "interchangeable parts" approach to the collection of evidence suggestive of validity. As Messick (1980) pointed out:

Many test users focus on one or another of the types of validity, as though any one would do, rather than on the specific inferences they intend to make from the scores. There is an implication that once evidence of one type of validity is forthcoming, one is relieved of responsibility for further inquiry. (p. 1014)

Although Muma attempted to associate us with such an interchangeable parts approach, we did not present this view in our paper and join him in decrying such a simplification of the complex construct of validity. We are, however, concerned about the impression that Muma's comments may leave regarding the prevalent thinking about validity. Authorities on psychometric theory and practice, even as they support the primacy of construct validity, are not monolithic in their abandonment of either the terms nor the concerns that have been addressed under the terms content and criterion-related validity (e.g., Anastasi, 1976, Salvia & Ysseldyke, 1981, pp. 101–109; Thorndike, 1982, pp. 184–203). We call to the reader's attention a summary of a common contemporary view of validity: "Content, criterion-oriented, and construct validities are not separate types of validity, but aspects of a broad process of validating the interpretations and uses we make of test scores" (Nikko, 1983, p. 435).

In his conclusion, Muma supported as "constructive" our suggestions that test users become more aware of psychometric principles and of the flaws of the tests they are using and further that they act on that knowledge in their role as test consumers. We would further recommend that test users rely on no single source for their information regarding the psychometric characteristics of tests or regarding psychometric principles. Clearly, the concepts are important and complex enough to warrant further reports and further discussion.

Rebecca J. McCauley
John F. Kennedy Institute
Baltimore
Linda Swisher
University of Arizona
Tucson

REFERENCES


Received May 9, 1985
Accepted May 12, 1985

Some Problems in the Clinical Application of Phonological Theory

During the past several years, a number of studies in JSHD have cited Parker (1976), among others, as support for applying phonological theory and, in particular, distinctive feature (DF) theory to the analysis of misarticulations. Our purpose here is to point out several ways in which the apparent interpretation of current phonological theory reflected in these studies departs radically from the views presented in Parker (1976) and other relevant works and also to discuss several related problems with methodology and the interpretation of results.
Summary of Relevant Phonological Theory

Phonological theory recognizes at least four levels of representation: (a) systematic phonemic, (b) classical phonemic, (c) systematic phonetic, and (d) physical phonetic. (Classical phonemic, also known as "axonomic" or "autonomous," is not actually a level of representation in the strictest sense of the word; see Chomsky, 1964, pp. 88-91, for discussion.) The first three levels consist of segments defined in terms of DFs and are related by phonological rules, also stated in terms of DFs (Parker, 1976, pp. 28-31). Most importantly, however, all three of these levels and the rules that relate them are part of the speaker’s psychological or mental system. That is, systematic phonemic, classical phonemic, and systematic phonetic units (i.e., segments and DFs) are not isomorphic with parameters of speech production or with properties of the acoustic signal (Parker, 1977). On the other hand, the fourth level, the physical phonetic level, constitutes a physical or physiological description of physical speech production and/or the resulting acoustic signal. This level may be described in terms of physical production features (Parker, 1976, pp. 34-38) or physical properties of the signal. Thus, there are two related but distinct components of phonology relevant to this discussion: psychological (i.e., the systematic phonemic, classical phonemic, and systematic phonetic levels) and physical (i.e., the physical phonetic level).

As an illustration of these four levels, consider a partial characterization of the English words pups and pubs. (The relevant distinctions between levels are indicated by dotted lines.) (See below.) These four levels of representation are posited in order to account for a range of phenomena.

**Systematic phonemic level.** The description of the final segments in pups and pubs as /z/ accounts for the fact that speakers of English treat both of these segments as instances of the plural morpheme; thus, they are represented identically on this level of representation.

**Classical phonemic level.** The description of the same two segments (i.e., the final segments in pups and pubs) as /n/ and /z/, respectively, accounts for the fact that speakers of English can judge these two segments as different classical phonemes, through introspection alone. The apparent paradox that instances of the plural morpheme are the same on one level but different on another is reconciled by hypothesizing a phonological rule that (in this example) maps the systematic phonemic level onto the classical phonemic level and vice versa: namely,

\[ [+\text{voice}] \rightarrow [-\text{voice}] / [-\text{voice}] \rightarrow #. \]

(A [+voice] segment becomes [-voice] when it occurs at the end of a word and is preceded by a [-voice] segment.)

**Systematic phonetic level.** The description of the two instances of the classical phoneme /p/ in pups as [pʰ] and [p], respectively, accounts for the fact that speakers of English treat instances of /p/ in varying contexts as different. That is, for example, the representation of pups as [pʰpʰpʰpʰ] is grammatical in English, whereas [papʰpʰpʰ] is not. As above, the apparent paradox that instances of /p/ are the same on one level but different on another is reconciled by positing a phonological rule that (in this example) maps the classical phonemic level onto the systematic phonetic level and vice versa: namely,

\[ \begin{array}{c}
\text{[continuant]} \\
\text{[voice]}
\end{array} \rightarrow [+\text{aspirated}] / \text{[+delayed onset of vocal cord vibration]} \]

(A voiceless stop is further specified as [+ aspirated] when it occurs in syllable-initial position before a stressed vowel.)

**Physical phonetic level.** The description of each of the [p]s in pups in different physical terms accounts for the fact that the production of an acoustic signal corresponding (more or less) to prestressed [pʰ] in English involves (among other things) a delay in the onset of vocal cord vibration following the release of lip closure, whereas poststressed [p] involves a different vocal tract configuration. The motivation for distinguishing this level of representation from the systematic phonetic level is provided by the fact that no one-to-one correspondence exists between psychological units (segments and DFs), on the one hand, and either parameters of speech production or the physical properties of the signal that contribute to the perception of such units, on the other. For example, the DF [+aspirated] is not invariably cued by the production feature [+delayed onset of vocal cord vibration]. Note, for example, that in potato [pʰətəʊ] vocal cord vibration may not begin until halfway through the stressed syllable, yet the [p] is not aspirated. That is, a delay in the onset of vibration (a physical dimension) does not always coincide with aspiration (a psychological dimension).

Let us summarize the main points made in this section. First, the study of phonology can be divided into two separate components: the study of psychological phenomena and the study of physical phenomena. A complex relation holds between the two, yet they are nonetheless distinct, at least in theory. The psychological component is the set of rules that serves as input to the physical production system; the physical component is the production system itself. Second, phonologists have posited four levels of representation: systematic phonemic, classical phonemic, systematic phonetic, and physical phonetic. These levels are postulated for one simple reason: In order to account for phonological phenomena, it is necessary to treat two items as the same on one level and different on another.

**Discussion**

We will now discuss three problems that frequently arise in studies that apply DF theory and other aspects of current phonological theory to the analysis of misarticulations.

1. **DF theory is not adequately motivated in the study of physically based articulation disorders.** Speech pathologists often distinguish between two types of articulation disorders. Functional or phonemic disorders are those "characterized by the lack of a neurological or physiological basis for the disorder," whereas phonetic disorders are "those that may stem from problems in motor sequencing, structural anomalies, or from muscular or neuromuscular disorders" (Ruder & Bunce, 1981, p. 59). Note that these levels correspond not to the phonemic (systematic or classical) and systematic phonetic levels of representation but rather to the mental and physical components of
phonology. Phonemic disorders appear to reflect problems in the speech-hearer’s psychological system—that is, at the systematic phonetic level or above. Problems stemming from identifiable physical disorders, on the other hand, correspond to the physical phonology. Phonemic disorders appear to reflect problems in the psychological (i.e., mental) organization. Consequently, it is a systematic or classical), and systematic phones exist in the speech signal and/or in the physical production mechanism. The fact is that all existing research on the subject, as early as Twaddle (1938/1957) and as recently as Hammarberg (1982), argues against this view. Speech production cannot logically be described in terms of DFs, segments, or the phonological rules that relate them, because these constructs are part of a theory of psychological (i.e., mental) organization. Consequently, it is a contradiction in terms—and in theory—to talk about the “production” of a DF or of a segment or to refer to a segment as a “sound.” At best, one can talk about properties of the physical signal (i.e., acoustic cues) that lead to the perception (a psychological phenomenon) of a segment (a psychological construct). Likewise, one can talk about properties of the speech production mechanism in physiological terms (e.g., lip and jaw movement, tongue movement, etc.) but not in psychological terms (DF, segment, etc.). See Parker (1977) and Repp (1981) for a more detailed discussion of this issue.

Such terminological and theoretical confusion may arise from the lack of a commonly agreed-upon set of terms for distinguishing acoustic cues from psychological constructs. When trying to accurately describe the relationship between the speech signal and the psychological percept, one may have to resort to such unwieldy phrases as “the subject’s articulatory configuration facilitated the production of a speech signal containing the acoustic cues that led to the perception of the segment /k/”—surely an unsatisfactory alternative. To avoid such stylistic anomalies, authors might employ phraseology such as “the subject produced the acoustic correlate of /k/.” Still another alternative is to add a new symbol, one for production. Currently, double slashes are sometimes used to represent systematic phonemes (e.g., /k\), slashes are used to represent phonemes (e.g., /k/), and brackets are used to represent phones (e.g., [k]). Perhaps parentheses should be adopted to identify the production correlates of the constructs (e.g., [k]). A new notation would then clarify the level of representation to which the writer is referring. (A similar point is made by Repp, 1981.) At the very least, slashes and brackets should be avoided when discussing a subject’s productions; for example, Hodson et al. (1983) include a table in which “phonetic transcriptions” are enclosed in slashes (p. 65).

DF theory does not account for the therapy strategy and its results. Our final criticism of the articles that attempt to utilize DF theory is that it is not at all clear that DF theory provides an adequate account of either the subjects’ problems or the results of therapy. Take the cases reported by Ruder and Bunce (1981).

With respect to K. E., they state that following her “production” of /s/ and /k/, “the phoneme /t/ was selected as the target sound for the study” (p. 60). Their rationale is that “/t/ should emerge as a consequence of training on /s/ and /k/, which, when taken together, provide training on all the distinctive features of /t/ without direct training on it” (pp. 60–61). A similar rationale is offered for R. R.’s therapy (p. 63).

One problem with this therapy strategy concerns the selection of “training phonemes.” We are told that, for K. E., “training on /k/ would provide . . . experience” on the features [high] and [anterior] (p. 61). Apparently, Ruder and Bunce (1981) wanted K. E. to acquire an awareness of the values [+high] and [−anterior], values for which /k/ is specified. However, K. E. already had a segment with these values in her repertoire, namely /j/. Thus, it is unclear why /k/ was chosen as a training phoneme. As Ruder and Bunce point out themselves, “If generalization of a feature could be made with training on only one phoneme, then training only one phoneme would be more economical” (p. 65). (Again, a similar problem arises in the case of R. R.’s therapy.)

A more fundamental problem is that the theoretical assumptions that apparently underlie Ruder and Bunce’s procedures fail to predict their results. As far as researchers know, exposure to a single segment in isolation does not provide sufficient grounds for a language learner to infer the feature matrix for that segment. Take the extreme hypothetical case of a child exposed only to the signal corresponding to [p]. In this case, the child could not infer the DFs of [p] because he or she would not have been exposed to any segments that are differentiated from [p]. A more plausible assumption—that apparently held by Ruder and Bunce—is that a DF can be inferred only from a pair of segments that have contrasting linguistic status within the language, for example, [p] and [s]. Upon exposure to such segments, the language learner presumably infers some set of dimensions upon which these segments differ. Perhaps the most plausible hypothesis is that a language learner will infer a feature difference just in case two segments differ by only one acoustic feature (e.g., [p] and [b]). However, this is clearly not the hypothesis that Ruder and Bunce are
testing as /k/ and /s/ are not minimally different from each other or from /t/, the target phoneme.

Given the assumption that Ruder and Bunce apparently do hold, let us consider how it fits with their actual results. K. E. begins with the segments /h,m,g/; training is undertaken on /k/ and /s/. Under Ruder and Bunce’s assumption, exposure to these five segments should allow K. E. to infer not just the DF matrix for /t/ but also those for all other consonants. R. R., who began with segments /g,n,j,y/ should also make the same inferences upon exposure to /k/ and /s/. To see why, consider the DF matrices for the relevant segments. (See above.) Note that, within the set of DFs that differ in the English stops, affricates, and nasals (i.e., all segments that are [+vocalic, +consonantal, and –low]), each child had adequate exposure to contrasting values for these features based solely on (a) each child’s initial consonant repertoire and (b) exposure to /k/ and /s/. Again, then, Ruder and Bunce’s assumptions would predict that this exposure would enable each child to infer the remaining [+vocalic, +consonantal, –low] segments: in K. E.’s case, /p,t,k,v,ʔ,d,n,z,l,t,f,dʒ,f,ʒ/, and /g/ in R. R.’s case, /p,b,f,v,m,t,d,θ,ð,ʃ,l,t,f,dʒ,f,ʒ/, and /y/. There is no basis for predicting that ʔ/ alone would be inferred. Therefore, under any feasible interpretation of DF theory, the results achieved by Ruder and Bunce are not predicted.

### Conclusion

We have discussed a number of articles in JS HD in which problems arise from (a) the assumption that the use of phonological theory, including DF theory, is motivated in cases where the articulation problems are apparently physical in nature and (b) the use of phonological and DF theory terminology to describe production phenomena. We have further discussed particular difficulties in reconciling the therapy and results reported by Ruder and Bunce (1981) with DF theory in general and the views expressed in Parker (1976) in particular.

We do not mean to suggest by these comments that phonological theory should be avoided in speech pathology studies; on the contrary, it constitutes a powerful system for the analysis of linguistic structure and for understanding the way that the human mind imposes order upon the perception of the physical signal. However, because phonological theory is ultimately concerned with psychological constructs, rather than with the level of physical production, a certain amount of caution must be taken in attempting to apply this theory to the remediation of disorders that are apparently predominantly physical in nature. In addition, researchers in speech pathology need to treat the concepts and terminology associated with this theory in a way that is compatible with existing research on the subject.

## References


Received November 8, 1984
Accepted May 16, 1985

**LETTERS** 297
By mere happenstance, I recently came across some remarkable observations and theoretical explanations by Charles Cooley (1902) on the development of personal pronouns in young children. Cooley was trained in economics and is considered to be one of the founders of sociology. The boundaries between the social sciences at the turn of the century were not as complete and as rigid as they now appear to be. In a chapter entitled "The Social Self: The Meaning of 'I,'" Cooley described the development of personal pronouns, making reference to the language behavior of his young daughter, M., and his son, R. Although a later publication of his observational notes (Cooley, 1908) has been cited, it appears that Cooley's work has gone essentially unnoticed by linguists, developmental psychologists, and speech pathologists. This is indeed unfortunate, for his observations are astute, and his explanations and theoretical concerns are pertinent to current research. Of particular interest to speech pathologists is Cooley's description of an unsuccessful attempt to remediate his son's pronoun reversals.

To illustrate the freshness and clarity of Cooley's style, I quote at length from his discussion of the problem of reference in the use of possessive pronouns. Given a learning theory explanation of language development, it is a paradox that young children should be able to use possessive pronouns correctly. Cooley (1902) observed of his daughter:

When she was two years and two weeks old I was surprised to discover that she had a clear notion of the first and second persons when used possessively. When asked, "Where is your nose?" she would put her hand upon it and say, "my." She also understood that when someone else said, "my" and touched an object, it meant something opposite to what was meant when she touched the same object and used the same word. Now, anyone who will exercise his imagination upon the question how this matter must appear to a mind having no means of knowing anything about "I" and 'my" except what it learns by hearing them used, will see that it should be very puzzling. Unlike other words, the personal pronouns have, apparently, no uniform meaning, but convey different and even opposite ideas when employed by different persons.

It seems remarkable that children should master the problem before they arrive at considerable power of abstract reasoning. How should a little girl of two, not particularly reflective, have discovered that "my" was not the sign of a definite object like other words, but meant something different with each person who used it? And, still more surprising, how should she have achieved the correct use of it with reference to herself which, it would seem, could not be copied from anyone else, simply because no one else used it to describe what belonged to her? The meaning of words is learned by associating them with other phenomena. But how is it possible to learn the meaning of one which, as used by others, is never associated with the same phenomenon as when properly used by one's self? (pp. 157–158)

Given this appreciation of the complexity of pronoun acquisition, Cooley would have agreed with Schiff-Myers (1983) that it should not be surprising to observe pronoun reversals in normal children.

Cooley goes on to present an explanation for pronoun acquisition that is both original and consistent with actual language behavior. Yet, his explanation has the merit and scholarship of drawing upon the psychological science of his day, in particular, upon the notion that there is an instinct for possessiveness that is central to the affective sense of Self (James, 1890). For Cooley (1902), my does not refer to the objects of ownership but to the self-feelings aroused by property appropriation and related forms of social assertiveness.

Watching her use of the first person, I was at once struck with the fact that she employed it almost wholly in a possessive sense, and that, too, when in an aggressive, self-assertive mood. It was extremely common to see R. tugging at one end of a plaything and M. at the other, screaming, "My, my," "Me" was sometimes nearly equivalent to "my," and was also employed to call attention to herself when she wanted something done for her. Another common use of "my" was to demand something she did not have at all. Thus if R. had something the like of which she wanted, say a cart, she would exclaim, "Where's my cart?"

It seemed to me that she might have learned the use of these pronouns about as follows. The self-feeling had always been there. From the first week she had wanted things and cried and fought for them. She had also become familiar by observation and opposition with similar appropriative activities on the part of R. Thus she not only had the feeling herself, but by associating it with a visible expression had probably divined it, sympathized with it, resented it, in others. Grasping, tugging, and screaming would be associated with the feeling in her own case and would recall the feeling when observed in others. They would constitute a language, precedent to the use of the first-personal pronouns, to express the self-idea. All was ready, then, for the word to name this experience. She now observed that R., when contentiously appropriating something, frequently exclaimed, "my," "mine," "give it to me," "I want it," and like. Nothing more natural, then, than that she should adopt these words as names for a frequent and vivid experience with which she was already familiar in her own case and had learned to associate it with itself. Accordingly it seemed, as I recorded in my notes at the time, that "'my' and 'mine' are simply names for concrete images of appropriativeness," embracing both the appropriative feeling and its manifestation. If this is true the child does not at first work out the I-and-you idea in an abstract form. The first-personal pronoun is a sign of a concrete thing after all, but that thing is not merely the child's body, or his muscular sensations as such, but the phenomenon of aggressive appropriation, practiced by himself, witnessed in others, and incited and interpreted by a hereditary instinct. This seems to get over the difficulty above mentioned, namely, the seeming lack of a common content between the meaning of "'my' when used by another and when used by one's self. This common content is found in the appropriative feeling and the visible and audible signs of that feeling. (pp. 158–160)

Even though these observations and explanations stand in the context of a science of 80 years ago, they are remarkably contemporary. For example, Cooley's observation that personal pronouns are first used in the possessive form to refer to acts of appropriation has recently been substantiated by Charney (1980). Furby (1980) has described the importance of contending for possessions, and most recently, Levine (1983) has presented evidence that there is a relationship between acts of appropriation, the sense of Self, and the use of possessive pronouns. At a more global theoretical level, Cooley's (1902) explanation presumes that there are cognitive prerequisites to the acquisition of language forms. Possessive first-person pronouns would require decentering from an egotistic perspective to acknowledge other persons with feelings like one's own.

In other words, the meaning of "'I" and "mine" is learned in the same way that the meanings of hope, regret, chagrin, disgust, and thousands of other words of emotion and sentiment are learned: that is, by having the feeling,
imputing it to others in connection with some kind of expression, and hearing the word along with it. (pp. 160–161)

Although Cooley apparently did not appreciate the importance of his presumption, some current writers (e.g., Rice, 1983) express the view that language development is contingent upon cognitive development. Later in his chapter, Cooley argues forcefully against research on the meaning of I that proceeds by asking children to speculate on what I might refer to (Hall, 1898). Given his own reliance on observational data, this would seem to presage a second theoretical issue, that a word's meaning can only be understood by reference to its usage in actual social contexts (Rose, 1980; Wittgenstein, 1958).

One final historical point. Cooley (1902) may have been the first to describe pronoun reversals in a child as well as an attempt at remediation. He attributed his son's pronoun reversals to personal temperament and to lack of property challenges by other children.

R., though a more reflective child than M., was much slower in understanding these pronouns, and in his thirty-fifth month had not yet straightened them out, sometimes calling his father "me." I imagine that this was partly because he was placid and uncontentious in his earliest years, manifesting little social self-feeling, but chiefly occupied with impersonal experiment and reflection; and partly because he saw little of other children by anthesis to whom his self could be awakened. M., on the other hand, coming later, had R.'s opposition on which to whet her naturally keen appropriativeness. And her society had a marked effect in developing self-feeling in R., who found self-assertion necessary to preserve his playthings, or anything else capable of appropriation. He learned the use of "my," however, when he was about three years old, before M. was born. He doubtless acquired it in his dealings with his parents. Thus he would notice his mother claiming the scissors as mine and seizing upon them, and would be moved sympathetically to claim something in the same way—connecting the word with the act and the feeling rather than the object. But as I had not the problem clearly in mind at that time I made no satisfactory observations. (pp. 161–162)

Further on in his criticisms of Hall's work, he mentions an attempt to remediate R.'s pronoun reversals by suggesting to the child that the referent of first-person pronouns is the body.

This analysis of the "I," asking one's self just where it is located, whether particular limbs are embraced in it, and the like, is somewhat remote from the ordinary, naive use of the word, with children as with grown people. In my own children I only once observed anything of this sort, and that was in the case of R., when he was struggling to achieve the correct use of his pronouns; and a futile, and that was in the case of R., when he was struggling to achieve the correct use of his pronouns; and a futile, and as I now think mistaken, attempt was made to help him by pointing out the association of the word with his body. (p. 163)

Cooley's explanation of the development of possessive pronouns, along with his case history of his son, suggests that the first child may be quite different from later siblings in the acquisition of pronouns. First children in nuclear families simply do not have the opportunities to dispute property claims with other children. This would seem to advise caution in making broad generalizations about normal pronoun acquisition based on case studies of first children. This might apply, for example, to Schiff-Myers's (1983) otherwise excellent case study and review.

In conclusion, I would recommend Cooley (1902) to those interested in the development of possessive pronouns. From a historical perspective, it is lamentable that he did not persist with developmental linguistics. His contributions undoubtedly would have been significant.

Floyd Rudmin
Queen's University
Kingston, Ontario

ACKNOWLEDGMENT

During the preparation of this manuscript, support was provided by the Social Sciences and Humanities Research Council of Canada, Doctoral Fellowship #452-83-5578.

REFERENCES


Received February 2, 1984
Accepted May 17, 1985

Role of Iconicity in Sign Acquisition: A Response to Orlansky and Bonvillian (1984)

With the increased use of augmentative and alternative communication (or nonspeech communication) in programming for individuals with severe communication problems as a result of cognitive, motor, and/or psychological causes, one of the major questions raised has been why nonspeech symbols are successful when the more common oral approaches have failed. At the March 1977 Gatlinburg Conference on Research in Mental Retardation, Fristoe and Lloyd (subsequently published in 1979) hypothesized that iconicity was a major factor contributing to the success of many individuals. The results of our studies since
1977 support the hypothesis that iconicity is an important factor in learning many types of non-speech communication symbols (Doherty & Lloyd, 1983; Goossens, 1984; Lloyd & Fristoe, 1978; Luftig, Gauthier, Freeman, & Lloyd, 1980; Luftig & Lloyd, 1981). Also, in June of 1977 at the NAD National Symposium on Sign Language in Chicago, Brown (subsequently published in 1978) hypothesized that iconicity was a major factor in learning manual signs and presented some interesting post-hoc analyses on early sign learning by chimpanzees and by autistic children, as well as experimental data from unimpaired children in support of this hypothesis. The recent JSHD paper by Orlansky and Bonvillian (1984, pp. 287-292) provided important data further supporting the hypothesis that iconicity is an important factor in learning manual signs. Although the authors indicated iconicity may be an important factor, unfortunately, in their abstract and in the paper itself, they stated that their study “suggests that the role of iconicity in young children’s acquisition of signs may be overrated by some investigators” (p. 287). This statement may lead some readers (especially those who still have doubts) to conclude that iconicity is relatively unimportant. Therefore, it seems appropriate to offer further discussion on the Orlansky and Bonvillian paper.

Definition of Iconicity

As Orlansky and Bonvillian pointed out, there are numerous definitions and interpretations of iconicity. They chose to restrict their definition of iconicity to only those signs that are clearly guessable or transparent and stated that this corresponded to the categories of pantomime and imitation suggested by Stokoe, Casterline, and Cronenberg in 1965. All other signs in the Orlansky and Bonvillian study were classified as metonymic (signs incorporating a relatively minor feature of the referent) or arbitrary (signs with no relationship to the referent). However, during the past decade numerous investigators have used broader definitions (E.g., Luftig & Klima, 1976; Brown, 1977, 1978; Fristoe & Lloyd, 1977, 1979; Griffith, Robinson, & Panagos, 1981; Klima & Bellugi, 1979; Mandel, 1977; Robinson & Griffith, 1979). For example, Bellugi and Klima (1976) included two aspects of iconicity: transparency (i.e., the degree to which nonsigners can guess the meaning of a sign) and translucency (i.e., the degree to which nonsigners see some basis for the relationship between a sign and its meaning). Brown's classic 1977 NAD paper used a similar definition of iconicity that included both transparency (or guessability, which approximates the Stokoe et al., 1965, iconic category) and translucency (which is similar to the Stokoe et al., 1965, metonymic category). This broader definition allows one to make sense out of what on the surface may have appeared to be conflicting results on the importance of iconicity in early lexical development. Iconicity can be said to encompass both the iconic and metonymic categories in the Orlansky and Bonvillian study, and instead of their subjects' vocabularies containing approximately 30% iconic signs, the percentage of signs bearing a relationship to their referent was actually closer to 65% when both categories were combined (i.e., the preponderance of signs in the vocabularies was iconic). Using their own logic, their data present a stronger case for the importance of iconicity. This is especially true when we consider the small number of "iconic" or onomatopoetic spoken words in initial vocabularies of children. Although there do not appear to be any published reports of specific analysis, Brown (1977) mentioned that learning an initial speech lexicon is an arbitrary task.

Degree of Iconicity in Sign

Orlansky and Bonvillian stated that there is not yet consensus on the exact percentage of easily guessable or transparent signs and used Hoemann's (1983) estimation that ASL transparency ranged from approximately 10% to 30%, but they do not refer to

more recent papers that result in a more conservative estimate. One is cautioned that Hoemann's upper percentages are based on liberal scoring of responses and that for some items judged transparent, only 4 of 52 subjects (i.e., 8%) correctly guessed. Concerned about Hoemann's criteria and that such a broad range was resulting in the perception that ASL signs were more transparent than the data actually indicated, we analyzed the data presented in the paper and concluded that when one used a realistic criterion of transparency requiring at least 50% of the subjects to guess an item correctly, only 13% of Hoemann's items could be considered transparent. This is in general agreement with a 10%-15% transparency range found in other studies (e.g., Bellugi & Klima, 1976; Kirschner, Algozzine, & Abbott, 1979; Luftig & Fristoe, 1978). Because Orlansky and Bonvillian's own data showed 30.8% (ranging from 20% to 60%) of the 10-sign and 33.7% (ranging from 13.5% to 46.7%) of the 18-month vocabularies were transparent, this clearly indicates that the proportion of transparent signs in this sample is higher than the 10%-15% one would expect from a more general sample of ASL signs. Therefore, the Orlansky and Bonvillian data support the role of iconicity in sign acquisition. These findings with normally hearing children of deaf parents are in agreement with Launer's (1983) recent dissertation, which found the early lexical acquisition of deaf children of deaf mothers was influenced by semantics, iconicity, and material input.

Orlansky and Bonvillian conclude from their subjects' early use of iconic baby signs, which are replaced with more arbitrary adult signs, that iconicity is not important in early lexical development. Rather, we would agree with Launer's (1983) observation that the early use of iconic signs further demonstrates the importance of iconicity in early lexical development.

Importance of Iconicity to Nondeaf Learners

Numerous investigations of the role of iconicity in learning non-speech communication symbols by autistic, retarded, and nonhandicapped subjects have clearly demonstrated its importance for comprehension (Goossens', 1984; Griffith & Robinson, 1980; Luftig et al., 1980; Luftig & Lloyd, 1981; Polzer, Wankoff, & Wollner, 1979; Snyder, 1978) and for production (Doherty & Lloyd, 1983; Konstantareas, Oxman, & Webster, 1979; Luftig, 1981; Polzer, Wankoff, & Fristoe, 1979). We could find only one study that failed to support the hypothesis (Kohl, 1981). Kohl's failure to support the hypothesis may be related to several critical factors including the small number of signs investigated (only 16), the small number of subjects involved (8), and the method used to determine levels of iconicity (dichotomous rather than continuous). When we compared Kohl's 16 signs with iconicity (translucency) ratings obtained on 910 signs by naive observers using a 7-point scale, it became evident that Kohl used signs spanning the continuum of iconicity rather than two distinct groups, 8 iconic versus 8 arbitrary or opaque signs. In addition to including many signs of medium iconicity in both groups, we found 1 sign in the "iconic" set (cat) and 1 sign in the "arbitrary" set (coat) were not placed in the appropriate group. It is our view that although Kohl's study provided useful data on other features of signs, it should be disregarded on the iconicity question.

Other Factors

In pointing out the importance of iconicity, we do not mean to diminish the role other features may play in learning of signs by individuals with deafness as well as other causes of severe communication impairment. We believe formational or motor features also play an important role in learning an initial sign lexicon (e.g., Fristoe & Lloyd, 1977, 1979, 1980; Karlan & Lloyd, 1983; Pennington, Karlan, & Lloyd, in press). In addition to our ongoing iconicity research, we have conducted research on the
role of formational features (Doherty & Lloyd, 1983; Lloyd & Doherty, 1983). These studies indicated that contact is a useful formational feature for some subjects and that, in addition, the number of hands used to produce a sign interacts with iconicity in facilitating sign acquisition. We are continuing our research on formational features; however, the critical point of this letter is that the Orlansky and Bonvillian data add further support to the hypothesis that iconicity is important. Thus, the conclusion we must draw is that iconicity remains one of several important factors that assist learners of sign systems and account for the relative ease with which individuals learn nonspeech symbols.

Lyle L. Lloyd
Barbara Loeding
Jane E. Doherty
Purdue University
West Lafayette, IN

REFERENCES


BROWN, R. (1977, June). Why are signed languages easier to learn than spoken languages? Part Two (An extension of a keynote address presented at the National Symposium on Sign Language Research and Teaching, Chicago).


Received October 23, 1984
Accepted March 1, 1985
Response to “Contemporary Accounts of the Cognition/Language Relationship” (Rice, 1983)

In her article “Contemporary Accounts of the Cognition/Language Relationship: Implications for Speech-Language Clinicians” (JSHD, November 1983), Mabel Rice discusses several hypotheses about the relationship between cognition and language that have appeared in the literature, and she considers the implications of these hypotheses for the clinical process. She ultimately concludes that “the cognition hypotheses have not fulfilled their early promise, but they have spawned a range of explanations that contribute helpful perspectives” (p. 355).

Though Rice has done an excellent service in reviewing and integrating a large body of literature, we do not agree that the clinician who is searching “for models that approximate reality” (p. 355) should be encouraged by that literature. Rather, we would argue that the literature is marked by logical fallacies and inconsistent definitions that make it impossible to have a meaningful discussion of the relationship between cognition and language. On the basis of Rice’s own review, we conclude that (a) the cognition and language hypotheses can be meaningfully answered, not because “the available scientific methods are crude relative to the task” (p. 355) but rather because there is a major problem in how the constructs have been defined and in how the available research evidence has been interpreted. We believe that the problem is more logical than empirical and, therefore, will not be solved by the usual appeal to more and better research.

Problems in Definition

At a minimum, in order to evaluate the claim that cognition causes language (or the opposite), it is necessary that language and cognition refer to identifiable distinct concepts. If this minimal condition is not met, we end up asserting that language causes language or that cognition causes cognition. A second minimal condition is that the terms language and cognition should be used with some consistency across the various hypotheses that attempt to account for the relationship. Rice acknowledges the problems of definition when she notes, “One obvious problem is the use of independent measures of cognitive and linguistic knowledge” (pp. 344–345) and also when she recognizes “the possibility that language-based measures of cognition imply a circularity in our reasoning about the role of cognition in language acquisition” (p. 355). Rice’s review makes it clear that the current literature lacks agreed-upon definitions of the two constructs, but she is still inclined to suggest that the literature provides useful perspectives.

It is our impression that the two constructs, cognition and language, have not been successfully separated in the literature or in Rice’s review. Early on, Rice introduces her own definition of cognition as “nonlinguistic knowledge, or mental process, as contrasted to knowledge encoded in language” (p. 348). This definition is apparently based on Piaget’s concepts of mental representations and structures. She then depicts the relationship in a diagram (Figure 1, p. 349) in which nonlinguistic knowledge is on the bottom “because it is more basic” (p. 349), and linguistic knowledge is represented as a “higher-order kind of knowing” (p. 349). (The diagram is problematic because it already seems to presume the nature of the relationship by making cognition basic and language higher order.)

By ignoring the competing versions of the cognition-first hypothesis, Rice introduces different notions of cognition that exist in the literature. In what Rice calls the local homologies model, there are apparently deep cognitive processes that drive cognition and language (an instance of cognition presumably causing cognition). In the epigenetic interactionist viewpoint, Rice points out, cognition subsumes specific processes, such as short-term memory or auditory processing, which are outside the traditional notions of mental representations and structures. We suggest that to the degree that the two hypotheses use significantly different definitions of cognition, they cannot be considered competing explanations of the same phenomenon. Furthermore, within the framework of the epigenetic interactionist viewpoint, we are not sure whether the processes subsumed by cognition are strictly cognitive or linguistic. As Rice notes, processes such as short-term memory and auditory processing have traditionally been invoked as possible causes of language deficit, but are themselves affected by language skills. Thus, short-term memory may account for a lack of expressive vocabulary, but a lack of vocabulary can also constrain short-term memory.

In summary, it seems fruitless to compare the merits of the local homologies and the epigenetic hypotheses until there is agreement concerning the terms of the hypotheses. If cognition implies auditory processing in one hypothesis and an entirely different Piagetian process in another, the hypotheses cannot meaningfully be compared. Although, of course, terms can be made to mean whatever the user intends, a scientific base of understanding cannot be built when the terms are not closely related to the same events.

Another issue that requires conceptual clarification is the way in which cognition may be assumed to cause language. One form of the relationship might be that before any language development can occur, there is some minimal level of cognitive development that must be reached, for example, that children are not able to acquire language until they have reached means-end mastery. A second form of the relationship would be that there is a specific cognitive attainment for each form of linguistic development, for example, that children can’t learn to pluralize until they first master the concept of numerosity. A third possibility combines the first two: Children must meet some minimal cognitive requirement for any language to develop, and from that point onward certain linguistic structures require specific cognitive prerequisites, although others do not. Each of these possibilities invites a different kind of analysis, and all of them require a sound and consistent definition of the basic terms of the analysis.

In summary, our first concern in attempting to establish the relationship between cognition and language is that (a) the constructs underlying these terms have not been defined with precision, (b) the definitions change across arguments and hypotheses, and (c) the constructs themselves do not satisfy the minimal condition of independence. In addition, the presumed relationship can take different forms that make different demands on how closely cognitive and linguistic events are calibrated.

Problems With Interpreting the Evidence

Our second major concern is with how the empirical evidence should be interpreted. Let us grant for now that cognition and language can be separated so that it is meaningful to look for a causal relationship between them. Because the evidence concerning the relationship is invariably correlational, there is little hope of determining the direction of causality. For example, Rice cites the fact that children use words in a nonadult manner as support for the cognition hypothesis, arguing that the errors indicate that children have different cognitive understandings than do adults. This interpretation may be correct, but the data might also indicate that although children’s concepts are exactly like those of adults, children are specifically deficient in their ability to represent underlying concepts in linguistic terms. Both interpretations are plausible given the empirical evidence. The same finding can be used to support either a cognition-first or a language-first account of the relationship.

As further evidence for the cognition-first hypothesis, Rice reviews the evidence that children use the terms now and yet with a past tense verb before they have productive command of
the present perfect tense and that they comprehend location before producing locative terms. However, temporal sequence is not a sufficient condition to establish causality. The fact that comprehension is shown to precede production does not mean that it causes production, any more than the fact that children vocalize before they point establishes vocalization as a prerequisite for pointing.

At another point, Rice comments,

If the mental deficits of language-disordered children are linked with their language performance, we would expect them to perform below chronological age matches, below mental age matches, and commensurate with language-matched (MLU) language-normal children. However, in studies with MA- and language-matched samples, those expectations are not always confirmed. (pp. 351-352)

It should not be surprising that these expectations are not always confirmed because they reveal a logical misunderstanding. The implicit argument that underlies the expectations reduces to: If mental retardation (A) causes language deficits (B), then whenever language deficits are observed, there must be an underlying mental retardation. The problem with this logic is that it assumes that because B is sometimes caused by A, it is always caused by A. However, it is logically quite possible that mental retardation may be one of many factors that result in language deficits. It is important to notice that the issue is neither empirical nor factual. We are not disputing Rice’s premise that there are multiple causes of language disorder; but if there are, as Rice seems ready to believe, then there is no reason to expect that all language-disordered children will show signs of mental deficits. The data that Rice summarizes provide “an annoying complexity” only if one misinterprets a logical rather than a factual relationship. Once again, this is a problem that will not be solved by accumulating more and better evidence; the solutions lie in how the evidence is used.

Finally, in her conclusion, Rice comments that there is a temptation to cast all cases of language impairment as a consequence of one common causal factor. The diversity of explanations, however, suggests a diversity of contributing factors as well; some children’s problems may be a consequence of one limitation, whereas other limits operate in other cases; or different patterns may be involved across children. Such possibilities would focus on the heterogeneity of language-disordered children instead of the assumption of homogeneity, and would move theoretical formulations toward clinical reality. (p. 355)

Rice seems to suggest that the search for a common causal factor underlying language disorders is out of touch with clinical reality. One could equally assert that the failure to discover a common causal factor reveals the poverty of our theories, rather than a fundamental truth about language disorder. Diverse explanations do not necessarily suggest diverse causes. They may simply reveal that we have not yet developed an adequate theory to account for the variety of language disorders. It would be unfortunate if the diversity of explanations for any of the speech and language disorders was taken as evidence that there are no general explanations for those disorders or that the search for general explanations is inappropriate.

Conclusions

Rice has reviewed the current state of a literature on language and cognition that is full of problems. Her review is thorough, and she is certainly aware of the general problems, as indicated in her conclusions. We disagree with Rice in how to interpret the problems. To us they are symptomatic not so much of crude scientific methods as they are of unrefined definitions and logic. We do not think the issues will be resolved by research that is more carefully designed or by greater efforts at collecting data. It seems to us that conceptual clarification is the necessary first step, with greater attention to how terms of an analysis are used and the kinds of presuppositions that are made as the evidence is evaluated. After a careful and hard look at the arguments, it may well develop that the issue of relating cognition to language in a causal way disguises a pseudargument rather than a problem of any generality. There may not be any general solutions or study-specific relations that are constrained by the instances of language and cognition that are studied. This need not be a troubling outcome. It would allow us to seek relations among specific classes of behavior rather than vague, poorly defined constructs.

Gerald M. Siegel
Jun Katsuki
Gail Potechin
University of Minnesota
Minneapolis

REFERENCE


Received April 11, 1984
Accepted November 22, 1984

Reply to Siegel, Katsuki, and Potechin

The criticisms of Siegel, Katsuki, and Potechin are at several levels. At the most general level, they question the integrity of the constructs of cognition and language and, by implication, the value of studies of the role of cognition in language development. At more specific levels, they question the interpretation of evidence by scholars working in the area of cognition and language, and, in particular, some of my interpretations and conclusions. In so doing, they assert that there are logical fallacies and inconsistent definitions.

I welcome the opportunity to provide a more complete context for my statements in the review article. A more lengthy treatment of some of the issues can be found in my recent book (Rice & Kemper, 1984).

The most important issue raised is the differentiation of cognition and language. Two points are embedded in this issue. One is the validity of the constructs of cognition and language. The other is the problem of independent measurement. In regard to the first point, the constructs of cognition (or thought) and language have been part of scientific inquiry for centuries (cf. Richer & Vetter, 1980). They are part of the tradition that attempts to differentiate qualitatively distinct domains of functioning within the individual organism. As such, they parallel other explanatory dichotomies such as intellect versus emotion, mind versus body, sick versus well, and immediate external versus internalized chains of reinforcement. Although there are challenges involved in attempts to isolate factors that interact in human functioning, there are heuristic advantages in doing so. In the case of language and cognition, it allows us to
address the extent to which language is a unique area of competence (cf. Karmiloff-Smith, 1975, chap. 1).

There is evidence that supports the conclusion that language is distinguishable from nonlinguistic knowledge and skills. A short, nondetailed list includes the following:

1. Cognition ontogenetically precedes language. Children do not learn to use the symbols of language to communicate with others until other skills have been mastered and other kinds of knowledge acquired.

2. There are well-documented differential effects of fatigue and neurologic trauma to language functioning as compared to other domains of knowledge.

3. Some aspects of verbal language do not have apparent parallels in nonlinguistic knowledge. Language structures appear to have arbitrary aspects that are not consistent with logical possibilities (cf. Chomsky’s arguments in Piattelli-Palmarini, 1980).

4. Within individuals there are differential profiles of achievements in language-related versus other skills. The most extreme discrepancies are apparent in language-disordered children, who are defined as having normal intelligence with special problems in language acquisition. The determination of normal intelligence hinges on the child's performance on the child's skill tasks that do not require linguistic skills. Our very identity as helping professionals is based on the distinction between cognition and language.

Contrary to the conclusion of Siegel et al., the basic definition of the constructs is shared by a number of contemporary child language scholars along with the conviction that it is reasonable and helpful to distinguish between cognition and language (cf. Cromer, 1981, in press). What is problematic is the second aspect of the issue, the determination of appropriate measurement (i.e., the nature of appropriate evidence).

As the contemporary literature has evolved, there has been a shift from an initial reliance upon the Piagetian model of cognition to an emphasis on aspects of cognition that are not central to Piagetian theory. It would be a mistake to characterize these shifts as fundamental changes of definition of cognition, as Siegel et al. have done. As I noted in the article (Rice, 1983, p. 348), no one, including Piaget, denies the existence and importance of psychological processes as part of the cognitive domain. What has been at issue is the level of description or abstraction at which these processes are manifested and the relationship between language and cognition. Piaget argued that parallels and ultimate explanation were to be found at the level of structure of knowledge, which is defined in abstract, metatraditional patterns of interrelationships. For Piaget, mental representation is an abstract cognitive accomplishment that can be manifested in several different ways: symbolic play; deferred imitation; recognition of patterns in objects, and possession of concepts. The notion of mental representation is not unique to Piaget (as interpreted by Siegel et al.), but its characterization is of it special (Mandler, 1983, provides further clarification as suggested in the article on p. 348). This network of interrelated skills/competencies generally precedes the child's language acquisition, hence Piaget's claim that language is only one manifestation of a previously acquired mental representation. Much of the dissatisfaction with Piaget is at the level of specific predictions and the strong claim for cognition's causal role. Many contemporary scholars accept the most general observation of Piaget's that mental representation is first evident across a wide array of tasks and contexts and is ontogenetically prior to language.

The evidence for the conclusion that meaning (mental representation) precedes language is most compelling when one observes language. Children talk about what they know. Their words and how they use them reveal underlying knowledge. That observation initiated the efforts of the last decade to account for children's earliest language acquisition (cf. Brown, 1973). No one argues that children's word meanings do not reflect underlying cognitive knowledge. The issue about circularity of reasoning hinges on the fact that language as evidence of cognition is too conservative, insofar as it does not take into account aspects of knowledge that are not represented in language or were introduced by language. Therefore, the use of language as a sole source of evidence is biased in the direction of the strong cognition hypothesis. It is understandable that the initial work with young children used language as an index of cognition, given that it is easy to evoke and is a measurable response of youngsters. Recent work has tackled the challenge of language-independent measures of cognition.

The other response to the strong cognition account of Piaget's is determination of the levels of abstraction that best capture the nature of the linkage between cognition and language. Implicit in the direction of influence issue is the extent of linkage between the domains. This combination of concerns is evident in Siegel et al.'s analysis of the way in which cognition may be assumed to cause language. A paraphrase of their point is that abstract, structural aspects of cognition may undergird language, and/or language and cognition may be linked at the level of specific meanings. The latter possibility is central to a new theory of cognition that has been developed by Kurt Fischer (1980) as an alternative to Piaget's model.

To summarize my response to Siegel et al.'s concerns about definitions, my position is as follows: (a) The constructs of cognition and language are distinct in that the constructs are empirically valid, insofar as available evidence and professional practices support the notion of distinct domains of competence; (b) although fundamental definitions remain relatively unchanged over the many years of controversy about the relationship between thought and language, emphasis upon particular aspects of thought and language has shifted as have corresponding definitions of specific competencies; (c) the shifts in emphases and definitions have been particularly salient during the past 15 years as a consequence of an unprecedented accumulation of evidence and theoretical formulation. The impetus, however, is not toward rejection of the fundamental distinctions but instead is toward greater specification of components and interrelationships. The fundamental questions of direction of influence and areas of linkage remain the same.

Siegel et al. also question the interpretation of evidence. In so doing, they provide a number of interpretations of their own. The first is that evidence concerning the relationship between cognition and language is "invariably correlational." That statement is not true if one defines correlation in the traditional statistical sense of relationship between two variables. Correlation does not entail temporal ordering or distinctions between independent and dependent variables. Interestingly, Siegel et al. call upon noncorrelational evidence to illustrate their assertion: the way that children use their words. As noted above, this evidence entails the assumption that meanings are prior to representation in words. Other noncorrelational evidence is replicable temporal relationships between cognition and language (cf. Corrigan's specification of temporal priority of cognition in her 1978 paper) and in treatment studies (Rice, 1980; Steckel & Leonard, 1981). At the very least, the qualifier "invariably" must be changed to "often."

Given that the data consist of more than simple associations, it is possible to test hypotheses about direction of influence, contrary to Siegel et al.'s pessimistic conclusions. There is little hope of using the data to determine causality. The example that Siegel et al. provide, that of children's errors with words, does lend itself to testing. They speculate that instead of indicating concepts different from adults', children's errors may instead indicate a problem with semantic representation—a specific difficulty of expressing in words what is known. That explanation is testable by careful analysis of the nature of children's word usage. It would involve consideration of children's correct uses of words, as well as errors, to determine if children have a general problem with semantic representation or if they have difficulty expressing certain concepts. For example, children's early use of color terms indicates awareness of semantic networks without underlying conceptual mastery. Children first name a small set of colors in response to the question, What color is this? That is, they reply with the name of a color. At the same time, they are
inaccurate in their selection of color terms to correspond to the color properties of objects. That is, they choose the wrong color name (Rice, 1980). In this case, children learn word-to-word relationships (semantic relationships) before mastering the meanings of words. In this case, children learn word-to-word relationships (semantic relationships) before mastering the meanings of words. The errors of color naming are conceptual, not semantic, in nature. The point of this illustration is that a comparison of patterns of correct and incorrect word usage, along with evidence from nonverbal performance, allows the testing of interpretations. The next problem is that Siegel et al.'s assertion that I included as evidence for the cognition hypothesis the observation that children comprehend locative uses of color before producing locative terms. My statement about locatives is that "children's acquisition of locative terms is paced by their nonlinguistic knowledge" (p. 351). The distinction is important because the antecedent skill is nonlinguistic knowledge (in this case, measured independent of language), not comprehension of the locative terms, as implied by Siegel et al.'s statement that comprehension before production does not indicate a causal relationship. Nowhere in the article, nor in my thinking, have I made that assertion.

Siegel et al. make the important point that temporal antecedence in and of itself does not constitute causal evidence. I agree completely. That is why in my review I tried to capture the convergence of evidence of different types that is essential to the determination of direction of influence between early cognitive achievements and language acquisition. The need for convergence of evidence is widely recognized by scholars working in the area. For example, in a recent discussion of the cognition hypothesis, Bates and Snyder (1982) call for four different types of evidence: empirical, sequencing, training, qualitative similarity, and correlational.

The next instance of misinterpretation is in regard to my discussion (pp. 351–352) of the hypothesis that the problems of language-impaired children are a consequence of certain mental deficits. Siegel et al. evidently erroneously interpreted "mental deficits" as "mental retardation." The key point here is that the distinction is about language-impaired children (i.e., those who score within normal range on nonlinguistic intelligence tests [as discussed in the preceding paragraph on p. 351 of the article]). The claim is that these children have some special problems of mental representation critical to language development, even though they have normal general intellectual functioning. In that case, the argument does entail that if there is a language deficit, there is a concurrent specific problem with mental representation. The hypothesis as originally proposed did not waffle with a sometimes-this-is-the-case clarification, nor has it been applied to mentally retarded children (as far as I know). My conclusion (p. 392) that the evidence does not consistently support the mental representation deficit in language-impaired children is empirically sound and logically consistent as well.

Siegel et al. put an impression on my use of language in their extensive quote from a paragraph about causes of language impairment (p. 355). They rightfully note that "diverse explanations do not necessarily suggest diverse causes." A more precise way of expressing my point would have been to state that the diversity of supporting evidence associated with each of the explanations suggests a diversity of contributing factors. I also concur with their observation that we need a theory of language disorders that can encompass and integrate available evidence.

To summarize my reaction to Siegel et al.'s concerns about interpretation of evidence, I find most of their assertions are misinterpretations of the available literature or of my writing. Siegel et al. conclude that a search for the nature of the relationship between cognition and language is fundamentally flawed by conceptual obscurity and is ultimately futile insofar as there is probably no general solution. Instead, they suggest that we would be better advised "to seek relations among specific classes of behavior." It is not clear to me how it may be the case that specific relations between language and cognition lack general import, although relations among specific classes of behavior do have general significance.

The issue with Siegel et al.'s conclusion that more careful definitions of cognition and language will necessarily obliterate the distinctions between the two domains and lead to trivial, nongeneralizable findings. What seems apparent in recent work is greater specificity and more evidence of subtle interactions (cf. Gopnik, 1984; Gopnik & Meltzoff, in press-a, in press-b). Evidence of specific relationships does not necessarily jeopardize generalizability. Instead, it can lead to new theoretical formulations of the structure of knowledge and its ontogenetic development (see Kurt Fischer's theory of cognitive development, 1980).

The existence of intricate interactions in the functioning of normally developing children does not entail a merging of the two constructs. Instead, it suggests the mutual interdependence of two domains and the extent of consequences in the case of diminished competence with one or the other.

I thank Siegel, Katsuki, and Potechin for the opportunity to expand my comments about the relationship between cognition and language, to clarify any inadvertent obscurities in my writing, and to contemplate one of my favorite topics.

Mabel L. Rice
University of Kansas
Lawrence

REFERENCES


Received January 7, 1985
Accepted May 12, 1985.
In the 34-page theoretical paper (Abbs, Hunker, & Barlow, 1983) critiqued by Hixon and Hoit (1984), we reviewed over 50 papers from neurophysiology, neuroanatomy, biomechanics, and clinical neurology to support the argument that (a) the various motor systems of the body (e.g., lips, jaw, tongue, velum, rib cage, abdomen, upper limbs, etc.) are likely to manifest different kinds or degrees of motor impairment as a result of systemic or supranuclear neurological damage, and (b) evaluation of limb motor impairments associated with these disorders may offer an indirect basis for inferring speech motor system pathophysiology. From our point of view, or from the perspective of most neurologists or motor neurophysiologists, there are far from major revelations. Indeed, our basic argument was that the very significant and unequivocal neurophysiological and biomechanical differences among different speech motor subsystems and the limbs was almost irrefutable evidence that the CNS does not control the spinal and cranial systems nor their respective subsystems in the same manner. . . [and hence] if there is damage of the central nervous system at a suprabulbar/supraspinal level, the result will be impairments in movements and muscle contractions that are different among the speech production systems and the limbs. For example, hypo- and hyper-gamma motor drive to muscle spindles, loss of aberrations in recurrent inhibition, and impairment of selective influences on motorneuron pool recruitment patterns have all enjoyed some popularity as partial pathogenic explanations for spasticity, rigidity, tremor, ataxia, hypotonia, dysmetria, and asthenia. If some of these explanations are even partially correct, and the implicated physiological processes (e.g., presence of spindles, operation of recurrent inhibition) differ from one motor system to another, then the pathophysiology must differ as well. (p. 30)

In providing support for impairment differences among speech and limb subsystems, we also offered some of our own recent data published in refereed journals (Barlow & Abbs, 1983; Barlow, Cole, & Abbs, 1983; Hunker, Abbs, & Barlow, 1982), primarily from patients with Parkinson's disease and congenital spasticity. Our basic position has been substantiated further by more recent findings (cf. Barlow, 1984; Barlow & Abbs, 1984; Hunker, 1984; Hunker & Abbs, 1984, in press). The last illustrative example of speech subsystem differential impairment we provided (described in a mere three quarters of a page of text) was based upon some very suggestive respiratory data from a 71-year-old woman with cerebellar disease. Surprisingly, Hixon and Hoit (HH) focused exclusively upon those limited, illustrative respiratory data from that single cerebellar patient and attempted, upon that narrow basis, to discount the important and highly defensible hypothesis of the entire 34-page paper.1 For example, the title of HH's letter implies directly that our sole basis for suggesting differential motor subsystem impairment was the data from this single ataxic subject. In this same vein, HH reconstructed quotes from our paper (including the elimination and addition of words that changed meaning and context; Hixon & Hoit, 1984, p. 436) to make it appear that the summary and conclusions (Abbs et al., 1983, p. 48) referred only to the data of that single cerebellar patient. Hixon and Hoit did not offer any substantive neurophysiological or neuroanatomical data for their rejection of our hypothesis. However, this response permits us to clarify the arguments made in that original review paper as well as to address critical misconceptions of speech respiratory control in normal and neurologically impaired subjects upon which HH's evaluations of our data obviously were based.

In the comments that follow, it will be demonstrated that HH's interpretations of our data are both antiquated and in error, due specifically to (a) reliance upon Hixon et al.'s (Hixon, Goldman, & Mead, 1973, 1976) limited, qualitative observations in very small groups of normal, highly practiced, knowledgeable, relatively young male subjects; (b) absence of technical caution regarding the limits of earlier single dimension, magnetometer-based rib cage (RC) and abdominal (AB) measures; (c) disregard of numerous recent papers indicating that their assumptions concerning respiratory biomechanics and physiology primarily are oversimplifications; and (d) misinterpretation of our data displays. The arguments in response to HH will be developed in the same order as they are listed above.

Inadequacy of the HH Normal Data Base

Throughout their critique of our data, HH cite their previous work on AB and RC measures in speech (i.e., Hixon et al., 1973, 1976), based upon work by Konno and Mead (1967) and Mead, Peterson, Grimby, and Mead (1967). For example, HH cite that earlier work to indicate that the paradoxical RC and AB movements we interpreted as abnormal are a common manifestation in normal subjects. Obviously, the validity of HH's comments depends upon the extent to which those cited data provide an adequate, normal baseline upon which to evaluate our observations. The descriptive studies of Hixon et al. were on a total of 6 normal subjects. All of these 6 subjects were male and younger than 52 years (mean age: 37.2). None of the subjects weighed less than 70 kg (154 lbs); mean weight was 84.1 kg (185 lbs). Perhaps most significantly, all of Hixon et al.'s subjects were experienced in respiratory studies, and most had considerable firsthand knowledge of the particular experimental purposes (3 of the 6 were the authors themselves). Finally, Hixon et al.'s subjects were "guided" in their performance by instrumental monitoring of the speech output.

In line with the principles of rigorous scientific inference, it is clear that Hixon et al.'s earlier data are not acceptable as a basis upon which to evaluate the validity of speech respiratory patterns in the subject we studied who was (a) experimentally naive, (b) elderly (71 years old), (c) frail (an 8-year nursing home resident weighing 107 lbs), (d) female, and (e) suffering from cerebellar ataxia. Additionally, our observations were not contaminated by her artificially monitoring the speech output; speech was produced naturally.

The inadequacy of Hixon et al.'s data base in this case also is documented directly by other observations. Specifically, the recent respiratory literature is replete with studies indicating that (a) age and sex are significant factors in respiratory function and control (cf. Axen, Haas, Guadino, & Haas, 1984; Braun, Arora, & Rochester, 1982; Chadha et al., 1982; Faithfyl, Jones, & Jordan, 1979; Niewoehner, Kleinerman, & Liotta, 1975; Bizzotto & Marazzini, 1970; Tobin, Tejvir, Jenouri, Birch, Cazergoul, & Sackner, 1983; Turner, Mead, & Wohl, 1968); (b) subject physi-
Differences in Transduction Techniques

A second probable basis for questioning HH’s interpretations is that our RC and AB signals were obtained with an entirely different transduction system than that used by Hixon et al. Specifically, based upon their experience with a magnetometer transduction system, HH suggest that our RC and AB signals were subjected to severe low-pass filtering (200–250-ms delays).

Paradoxically, HH also suggest that these same signals contain “frequent, rapid adjustments” that reflect artifact. Despite the illogical inconsistency between these two arguments, a major problem may be in their long-term experience with the limited magnetometer transduction system. The magnetometers (used by Hixon, 1982; Hixon et al., 1973, 1976) provide an analog signal that is allegedly proportional to the anterior-posterior (A-P) dimension of the RC and/or AB, from which it appears possible (based upon empirical manipulations) to derive the respective volume contribution of these two parts of the respiratory system. By contrast, the respiratory inductive plethysmograph (RIP) we used directly transduces the circumference of the RC or AB providing signals proportional to change in both the A-P and lateral dimensions. As such, magnetometer validity is based upon the simplifying assumption that changes in the A-P dimension are parallel to changes in other cross-sectional dimensions as well. However, as noted by Sackner (1980), the cross-sectional shape of the rib cage varies in the degree to which it is more elliptical or circular with variations in (a) lung volume levels, (b) respiratory activities (expiration vs. inspiration), (c) increased airway resistance (such as during speech), and (d) increased rate of respiratory maneuvers (also present during speech).

Given these considerations, it appears quite likely that the unidimensional magnetometer measures used by Hixon et al. may only imperfectly reflect the volume contributions of RC and AB, even with empirically based calibration. That is, even with recalibration during nonspeech maneuvers (such as those used in the Hixon et al. studies), errors specific to the speech tasks (e.g., rapid volume changes or increased airway resistance) would appear unavoidable. Specifically, nonspeech calibration would not take into account the cross-sectional shape changes of the RC and AB inherent to the speech tasks. Additionally, if the lateral dimension of the RC is out of phase with the A-P dimension (cf. Godfrey, Leventhal, Qintraub, Katzenelain, & Cooper, 1972), A-P readings from the RIP yield apparent (transduced) volume contributions that are time distorted in regard to the actual volume contributions. Seemingly in response to these problems, an additional magnetometer set to measure the lateral RC dimension as well as the A-P dimension was employed in several recent studies (Sampson & De Troyer, 1982), including one of the coauthors on the Hixon et al. studies (cf. Loring & Mead, 1982a). In contrast to the magnetometers, the RIP technique we used yielded a direct measurement of mean cross-sectional diameter that is not “affected by distortions in rib cage and abdominal compartments as are magnetometers [italics added]” (Sackner, 1980, p. 527).

A second critical difference between the data obtained with magnetometers and with the RIP may lie in the relative frequency responses of these two transduction systems. The frequency response of the RIP is R~1:6 Hz (cf. Sackner, Nixon, Davis, Atkins, & Sackner, 1980). The signals presented in our paper were not, contrary to the assertion by HH, filtered in any way. However, it is probable that HH were misled by previous observations with the magnetometers that have inherent time distortions as a result of a low-pass filter that is built into the magnetometer electronics. Unfortunately, Hixon et al. did not report the transducer frequency response (nor have previous papers from Hixon’s laboratory; cf. Rothenberg, 1982). For the magnetometer system in our laboratory (designed and constructed by the same company as the system used by Hixon), there is a built-in low-pass filter (to make the signals “look clean”) to use HH’s description, which is a Butterworth design (2-pole) with a cutoff at 10 Hz (~3 dB). Our magnetometer
system is somewhat newer than that used by Hixon et al. with a carrier frequency of 3 kHz. However, Hixon’s carrier frequencies were one-half to one-quarter those of our newer magnetometer system. (Hixon et al., 1973, report carrier frequencies of 1.53 kHz for the RC and 0.69 kHz for the AB.) Based upon these lower carrier frequencies, one might estimate that the built-in low-pass filter in Hixon et al.’s system would have proportionately lower cutoffs; that is, in the range of 2.5–5 Hz, with time delays ranging from 50 to 108 ms, respectively. Unfortunately, such filter delays would not be constant but vary for fast and slow movements. Such frequency responses would yield substantial losses of movement energy at frequencies presumably within the speech respiratory range.

Although these frequency response estimates are based upon some assumptions, they would be unnecessary if dynamic calibration data were simply provided in the original Hixon et al. articles (cf. Barlow & Abbas, 1983; Cole, Konopacki, & Abbas, 1983; Müller & Abbas, 1979). Regarding the frequency response necessary for transducing RC and AB movements for speech, a bandwidth of at least 16 Hz is desirable (cf. specifications in Irvin et al., 1984, or frequency response data in Ponsin, Papon, Du Vivier, & Richalet, 1975; Schmid-Schoenbein & Fung, 1978). Some of the rapid respiratory adjustments seen in our data, and interpreted by HH as artifacts, may simply reflect a misconception of respiratory movement dynamics based upon previous observations in overfiltered magnetometer signals. The Arizona speech research group has apparently overfiltered signals in the past (e.g., Hixon & Smitherton, 1985, report that their intraoral air pressure transduction system had a frequency response of 11.5 Hz, yet at least 50 Hz is desirable; cf. Edmonds, Lilly, & Hardy, 1971; Müller & Brown, 1980, Figure 9).

Given the inherent problems with magnetometer-based observations, many of HH’s qualitative judgments merely may reflect their own questionable measures of respiratory system behavior for speech. For example, it is assumed that the alleged calculations by HH of the delays between lung volume changes and the speech waveform in our unfiltered signals were based upon previous data from the magnetometers that appear to have inherent time distortions. However, as will be apparent, HH’s evaluations are flawed additionally because of obsolete assumptions regarding respiratory mechanics.

Rib Cage-Abdominal Diaphragmatic Movements and Mechanics

The critical assumptions associated with HH’s interpretations of our data are that RC and AB movements (a) reflect the only two possible means by which changes in lung volume occur, and hence (b) provide a valid basis to derive, via electronic “manipulations,” the total lung volume and the relative contribution of that lung volume from these two parts (Konno & Mead, 1967). As noted by Hixon et al. (1973):

The combined displacements of the rib cage wall and the abdominal wall together reflect the total lung volume change. Each of the two parts seems to move as a unit during breathing and, despite their sharing a common boundary at the costal margin, there is considerable functional separation between them. (p. 80)

Inherent within this conception is the notion that one can infer the movements of the diaphragm from movements of the abdominal wall (i.e., these two parts of the respiratory system are coupled in an obligatory manner). These assumptions, which shortly will be shown to be faulty, were the basis for the earlier analyses of Hixon et al. (1973, 1976), upon which some of HH’s claims clearly were based. For example, HH infer incidences of breath holding and inspiration from our signals based upon their assumption that simple RC and AB measures can be “processed” to determine changes in total lung volume. Similarly, HH suggest that the relative gains of our RC and AB signals are in error and when appropriately “corrected” should yield a “combined” signal that reflects total lung volume.

However, in the last 5 years, studies from several prominent respiratory physiology laboratories (including the one in which the Hixon et al., 1973, 1976, measures were made) indicate unequivocally that the convenient, operational notions underlying HH’s evaluations are an oversimplification and fatally flawed. More specifically, it is apparent that (a) diaphragmatic and RC volume contributions are not independent (Loring & Mead, 1982b; Macklem, Macklem, & De Troyer, 1983; Mead & Loring, 1982), (b) diaphragmatic volume contributions cannot be inferred in any simple manner from A-P AB wall movements (Loring & Mead, 1982b; Mead & Loring, 1983; Newman et al., 1984), (c) the interdependence of the diaphragm and RC as well as the independence of AB wall and diaphragm are questionable (Loring & Mead, 1982b; Macklem et al., 1983; Mead & Loring, 1982), (d) actions of some AB muscles act to generate inspiratory changes in RC dimensions (De Troyer, Sampson, Sigrist, & Kelly, 1983), and (e) the diaphragm must be considered two separate muscles with different innervation and different muscle fiber composition (De Cramer, De Troyer, Kelly, & Macklem, 1984; De Troyer, Sampson, Sigrist, & Macklem, 1981, 1982; Riley & Berger, 1979). Most importantly, Mead and Loring (1982) point out that the errors in estimates of diaphragmatic lung volume contributions made from A-P abdominal wall movements (as was the case for HH’s “calculations” and Hixon et al., 1973, 1976) are very large. They note,

The concept of a diaphragm-abdominal pathway for lung volume displacement suggests that the diaphragm would not contribute to a breath taken without displacing the abdominal wall. On the contrary (italics added), our analysis would suggest that the fractional contribution of the diaphragm to a tidal breath in which only the rib cage expands (italics added) is 0.41 ± 0.61. (p. 732)

Although the physiological factors that contribute to these phenomena are outside the scope of this letter, it is important to note that the errors in inferring movements of the diaphragm from the anterior abdominal wall movement vary substantially with lung volume level, based upon inherently flawed time delays in the biomechanical linkages among the RC, AB, and diaphragm (cf. Mead & Loring, 1982, Figures 2 and 4; Macklem et al., 1983, Figures 7 and 8). Moreover, such errors are substantial at or above FRC, the lung volume levels at which most of Hixon et al.’s normative data base was obtained.

These recent observations thus suggest that HH’s alleged calculations on our data simply were based upon earlier and far too simple concepts of respiratory physiology. For example, HH’s attempts to determine the correctness of the relative weighting of our RC and AB signals are an unjustified empirical exercise if one cannot make unambiguous inferences from AB wall movements to volume contributions of the diaphragm. As noted specifically by Mead and Loring (1982), the Konno-Mead technique (used by Hixon et al., 1973, 1976) overestimates contributions to lung volume by the rib cage and underestimates contributions by the abdomen (cf. Mead & Loring, 1982, p. 751).

In this vein, in our paper we made no claims concerning the ability to perfectly derive total lung volume from RC and AB movements, only that the RC-AB movements in this patient were paradoxical. Indeed, such simplified derivations apparently are not possible, given the recent data cited above. Similarly, a certain degree of AB movement cannot be interpreted quantitatively or qualitatively (as argued by HH) to reflect a volumetric contribution of the diaphragm to total lung volume changes. Likewise, apparent absence of total lung volume changes from measures confined to the RC and AB cannot be interpreted unequivocally as breath holding (as was suggested by HH). Finally, in the three sentences we devoted to the respiratory data from the cerebellar patient, we never claimed that the RC and AB time-motion signals and the motion-motion signals were from the same part of the test utterances. Moreover, such synchrony is
irrelevant to paradoxical movements between the AB and RC that are obvious in both displays.

In this context, it is apparent that HH's evaluations of our data are further flawed by their not incorporating recent developments in respiratory mechanics and respiratory physiology as well as a misinterpretation of our data displays. The final issue to be addressed in this response is the argument by HH against differential subsystem impairment with cerebellar disease.

Respiratory-Orofacial Impairment Differences in Cerebellar Disease

Without even a single supporting literature citation, HH assert that our argument for differential motor subsystem involvement with cerebellar impairment is a spurious claim. Given overwhelming evidence to the contrary, this assertion perhaps reflects the most serious limitation of their critique. Moreover, HH's assertion is contradictory to their enthusiastic endorsement of the earlier paper by Kent and Netsell (1975). That is, differential motor impairment of the respiratory and orofacial systems with cerebellar disease can be extrapolated directly from Kent and Netsell's (1975) interpretations. Kent and Netsell (1975) offered three major hypotheses regarding cerebellar dysfunction in sensorimotor integration (pp. 131–133). It is possible to show that if any one or all of Kent and Netsell's hypotheses are correct, cerebellar disease or damage will yield different motor impairments in the orofacial and respiratory systems.

Kent and Netsell's first hypothesis drew on the concept that one role of the cerebellum is controlling gamma motor drive to muscle spindles (cf. Denny-Brown & Gilman, 1965; Gilman, 1969; Glaser & Higgins, 1966). More specifically, cerebellar damage yields hypotonia (Holmes, 1917, 1922), which is said to be due to reduced spindle influences on alpha motoneurons (Gilman, 1969; Glaser & Higgins, 1966). As we pointed out (Abbs et al., 1983), a major neurophysiological difference between the orofacial system and many spinal innervated muscles (viz., the respiratory system) is in the operation of muscle spindles. Namely, the diaphragm and the rib cage muscles are endowed with muscle spindles that excite motoneurons autogenically, whereas the lip and tongue muscles are not so configured (cf. review by Abbs et al., 1983; Holmes, 1922). Moreover, different contributions of primary motor (area 4) and premotor (area 6) cortices to movement control (cf. Abbs & Welt, 1985; Gilman; 1973; Holmes, 1922). Moreover, consistent with our data, the loss of intercostal muscle afferents, via dorsal rhizotomy, dramatically influences respiratory motor control (Axen & Haas, 1962; Nathan & Sears, 1960). Given these data, hypotonia should be manifest in the respiratory system and not in the lips or tongue. Although the effects of hypotonia on movement control are not fully documented, a number of hypotheses have been offered, including problems in coordinating multiple movement actions (Gilman, 1969; Holmes, 1922). Moreover, consistent with our data, the loss of intercostal muscle tone has been associated with a reduction of chest wall recoil and elasticity, leading to paradoxical movements of the rib cage and abdomen (De Troyer & Heilporn, 1980; Frimioni, 1982).

The second hypothesis of Kent and Netsell (1975) was that cerebellar impairment influences the normal capability for integration of afferent signals based on considerations of normal cerebellar function (cf. Brooks, 1969; Patton, Porpora, & Brouskart, 1970; Konorski, 1967). To this point our data indicating that the cutaneous mechanoreceptors from the orofacial region do not project to the cerebellar hemispheres or the deep cerebellar nuclei in the same manner as muscle afferents from the spinal motor system (Stanton, 1980; cf. Bloedel & Courville, 1981, for discussion). Further, loss of respiratory muscle spindle afferent information with reduced gamma drive is likely to yield an impoverished set of ascending afferent signals to the cerebellum.

An important role of the ascending information from the chest wall may be found in the area of cerebellar control of the respiratory muscles especially for differentiated amplitude control of the different muscles and for integration of respiratory activity with postural activity and voluntary movements of the trunk (p. 338)

Not surprisingly, one must surmise that with cerebellar damage, such spindle-dependent control would be differentially impaired for respiratory versus orofacial speech actions, as we originally suggested.

The final hypothesis of Kent and Netsell (1975) was Eccles's classical argument, namely that a significant cerebellar influence on motor output is implemented via a reciprocal pathway between the cerebellum and the neocortex. The major component of this cerebellar-cerebral loop is via the dentate nucleus (cf. Brooks & Thach, 1981; Meyer-Lohmann, Hore, & Brooks, 1977; Spidalieri, Bushy, & Lamarre, 1983). Not surprisingly, the cerebellar outputs for respiratory muscles project to functionally and cytoarchitectonically different regions of the cerebral neocortex than those for the orofacial region. Specifically, Schell and Strick (1984) have shown recently that the orofacial cerebellar outputs (via the caudal dentate nucleus) project to a premotor region of the precentral gyrus (cytoarchitectonic area 6), whereas cerebellar projections for spinal nerves (via the rostral dentate nucleus) project to the primary motor cortex proper (cytoarchitectonic area 4). This implies that Eccles's cerebral-cerebellar loops are critically different for orofacial and respiratory motor control, as reflected clearly in the different contributions of primary motor (area 4) and premotor (area 6) cortices to movement control (cf. Abbs & Welt, 1985; Gilman; 1973; Holmes, 1922). Moreover, consistent with our data, the loss of intercostal muscle tone has been associated with a reduction of chest wall recoil and elasticity, leading to paradoxical movements of the rib cage and abdomen (De Troyer & Heilporn, 1980; Frimioni, 1982).

The second hypothesis of Kent and Netsell (1975) was that cerebellar impairment influences the normal capability for integration of afferent signals based on considerations of normal cerebellar function (cf. Brooks, 1969; Patton, Porpora, & Brouskart, 1970; Konorski, 1967). To this point our data indicating that the cutaneous mechanoreceptors from the orofacial region do not project to the cerebellar hemispheres or the deep cerebellar nuclei in the same manner as muscle afferents from the spinal motor system (Stanton, 1980; cf. Bloedel & Courville, 1981, for discussion). Further, loss of respiratory muscle spindle afferent information with reduced gamma drive is likely to yield an impoverished set of ascending afferent signals to the cerebellum.

Summary

Overall, the above considerations point to the erroneous nature of HH's critique. Under these circumstances, and given the substantial neurophysiological evidence, it clearly is justified to conclude that the dyssynchronization observed between the rib cage and abdomen in the cerebellar patient we studied is interpretable in relation to Holmes's (1917, 1922) original data indicating decomposition of movement and its differential manifestation in the orofacial and respiratory systems. In taking this position, we do not mean to imply that additional studies of this phenomenon are not necessary. However, for the present, the...
burden of proof is upon HH to provide quantitative evidence documenting that (a) their earlier so-called normal respiratory data bases are generalizable to older, frail, female, naive subjects under natural speech conditions; (b) magnetometer measures confined to a single dimension are not invalidated or time distorted by speech task-dependent variations in lateral rib cage movements or built-in low-pass filtering; (c) multidimensional A-P abdominal and rib cage surface wall movement signals can be empirically manipulated to yield an index of total lung volume under dynamic speech conditions despite inherent biomechanical variations in the diaphragm, abdomen, and rib cage with lung volume level; and (d) substantive data can be offered to support their implicit assertion that cerebellar damage or disease will yield uniform motor impairment in the respiratory and orofacial motor systems. Indeed, unequivocal data on these issues is likely to vary before additional work on speech respiratory control is undertaken or published, especially if unidimensional measures from magnetometers are to be used.

James H. Abbs
University of Wisconsin, Madison

REFERENCES


Received December 12, 1984

Accepted May 6, 1985

Reply to Abbas (1985)

In our letter (Hixon & Holt, 1984), we pointed out that the data provided by Abbas and his colleagues (Abbs, Hunker, & Barlow, 1983; Hunker, Bless, & Weismer, 1981) failed to support differential speech mechanism subsystem impairment, differential bulbar and spinal motor system impairment, and decomposition of compound respiratory movement in the ataxic dysarthric
subject studied. Our conclusion was based on what we judged to be compelling evidence that the data published from this subject had been misinterpreted, were internally inconsistent, and were invalid. Abbs's response has done nothing to change our minds. What it does, however, is present additional misinterpretations, internal inconsistencies, and invalid conclusions that warrant comment. Our reply considers the general topics dealt with in Abbs's response: respiratory-orofacial impairment differences in cerebellar disease, the normal data base used for comparison, differences in transduction techniques, and rib cage-abdominal-diaphragmatic movements and mechanics.

Respiratory-Orofacial Impairment Differences in Cerebellar Disease

Over one third of the text of Abbs's response is unrelated to the points we made in our letter. Our letter was not, as he contends, a critique of the Abbs et al. (1983) article. Abbs et al. claimed to provide sample data on different subjects to support the central hypothesis of their article. This hypothesis is that "if there is damage to a certain part of the central nervous system at a suprabulbar/supraspinal level, the result will be impairments in movement and muscle contraction that are different among the speech production subsystems and the limbs" (p. 30). What we addressed was the fact that the data that Abbs et al. provided from their ataxic dysarthric subject did not support the statements they made concerning the nature of her disorder and its underlying mechanism. We did not reject the hypothesis of the Abbs et al. article. We rejected the data presented from the ataxic dysarthric subject and her normal control. Our concern was that Abbs et al. were claiming the documentation of a clinical entity without supporting data. We doubt that anyone working clinically with individuals with dysarthria would question the possibility of differential speech mechanism subsystem impairment, differential motor system impairment, and differential speech mechanism and limb impairment in association with suprabulbar/supraspinal lesions. All we said is that the ataxic dysarthric data in the publication of Abbs et al. must not be taken as evidence of the existence of such involvement in their ataxic dysarthric subject. If, as Abbs suggests, rejection of these data constitutes rejection of the Abbs et al. hypothesis, then so be it.

The Normal Data Base Used for Comparison

A major part of Abbs's response is devoted to a discussion of the use of normative data as a standard against which to compare the data of his ataxic dysarthric subject. As we pointed out in our letter, the displacement-time data from the normal control subject of Abbs et al. (1983) reflect problems that are similar to those seen in the data from their ataxic dysarthric subject. That is, their normal control data reveal three inspirations, three instances of breath holding, and nearly two dozen relative volume contribution changes between the rib cage and abdomen during about 1.0 s of utterance. We noted in our letter that when these normal data are graphed into displacement-displacement charts, the resultant tracings are unlike any reported elsewhere for normal subjects. Pertinent to present concerns, it should be pointed out that when the data from the Abbs et al. normal control subject and ataxic dysarthric subject are graphed into appropriate displacement-displacement charts, they share such a large number of similar aberrations that they appear to be from the same population. This being the case, one must ask why the Abbs et al. claim of disordered respiratory behavior in the ataxic dysarthric subject is not also made for the normal control subject.

An issue we did not address in our letter is the contradiction between the data published by Abbs et al. (1983) and those published by Hunker et al. (1981). The same normal control subject (GW) was used to generate the data for both publications. However, the data provided in the two reports are highly dissimilar. Abbs et al. provide kinematic data characterized by inspirations, breath holdings, and frequent relative volume contribution changes. By contrast, Hunker et al. provide kinematic data that do not show inspirations, breath holdings, and frequent relative volume contribution changes. Concerning their data, Hunker et al. state that "the chest wall kinematic behavior of this subject was consistent with previously reported data from normal subjects (Hixon et al., 1982c)" (p. 3). The normal control subject demonstrated "consistent well-defined chest wall movement patterns . . . [in which] . . . expiratory excursions were nearly constant" (pp. 4–6) (meaning without abrupt chest wall adjustments). Which of these versions of the data—Abbs et al. or Hunker et al.—for the same normal control subject are we to accept? Another contradiction has to do with the position Abbs takes concerning what is acceptable as a normal standard of comparison for his ataxic dysarthric subject's data. Abbs contends that it is inappropriate to compare this subject's data to those of any normal control subject who differs in age, sex, physique, and level of experimental sophistication. He, in fact, admonishes that "the principles of rigorous scientific inference" are violated by using other than a closely matched normal control subject for comparison purposes. We wholeheartedly agree. Abbs describes his ataxic dysarthric subject as being a 71-year-old woman who was frail and experimentally naive. Why, then, given his statement about the importance of normal control matching, did he use as his comparison standard a 31-year-old male, muscularly built, experiential phonetician, who was trained by the first of us in matters of respiratory physiology and who has a great deal of experience and sophistication in respiratory experiments?

In our letter, we stated that the normal control data of Abbs et al. (1983) "at best, need explanation but very likely are invalid" (p. 440). In his consideration of normal control data, Abbs takes the position that the data for the normal subject in the Abbs et al. publications are not invalid because they agree with other data on normal subjects generated in his laboratory by Hunker and Abbs (1982b, 1982c). However, it takes only a glance at the data provided by Hunker and Abbs (1982a, Figure 4, p. 32) to see the same types of aberrations that are present in the normal control data of Abbs et al. For example, lung volume is the same at the beginning and the end of the first syllable in the speech sample shown, suggesting that, as in other research from Abbs and his colleagues, the lung volume summing network was erroneously calibrated (see Hixon & Hoit, 1984, p. 440). And, contradictory to Abbs's statement that "in our studies of these naive subjects we never saw paradoxical RC and AB movements," the only data shown in the Hunker and Abbs work reveal repeated rib cage paradoxing.

We have discussed Abbs's attempt to argue that his normal control subject and ataxic dysarthric subject are different in

---

1. Until such time as the hypothesis in the Abbs et al. article is critically evaluated, we would contend that those proposing this hypothesis bear the burden of proving at least two unlikely underlying assumptions. The first assumption is that all representations of the somatotopic map of the impaired nervous system are equally and similarly impaired. Failing to demonstrate this empirically, likewise renders the documentation of differential impairment theoretically trivial.

2. Another criticism of the Hunker and Abbs work has been put forth by Weismer (1985) who pointed out that their data do not bear on the central issue of motor equivalence they sought to test. Unfortunately, Hunker and Abbs confused the distinction between volume compression and volume displacement in their experimental design.
mechanism, when, in fact, their displacement-time data are similar. Rather than address this and related points from our letter, Abbs chose to criticize data from our laboratory on 6 normal adult male subjects (Hixon, Goldman, & Mead, 1973). He contends that these data constitute an inappropriate comparison to his own data because they come from individuals he refers to as being "young, male, above average weight," knowledgeable in respiratory physiology, and experienced in respiratory experiments. It should be pointed out that we never suggested directly that the Hixon et al. data should be used as the standard. Abbs reached this conclusion himself. Having done so, he discusses the Hixon et al. data as if they constitute the whole of speech breathing kinematic data from normal subjects. This is very far from true. In fact, our laboratory alone has made public data on 57 normal adult subjects, both men and women, ranging in age from 18 to 72 years, with an extremely diverse range of physical characteristics, almost all being unknowledgeable of respiratory physiology, and almost all having no experience in respiratory experiments (Bright, Hixon, & Hoit, 1985; Hixon, 1982; Hixon et al., 1973; Hixon & Putnam, 1983; Hodge, Hixon, & Putnam, 1982; Hoit & Hixon, 1985; Hoit-Dalgaard, Lansing, Plassman, & Hixon, 1983; Putnam & Hixon, 1983; Watson & Hixon, 1985). Despite the wide variety of subject characteristics just mentioned, the data from these subjects show substantial homogeneity in respiratory kinematic patterns. This homogeneity discounts Abbs's contention that the data from the 6 subjects studied by Hixon et al. (1973) are unrepresentative of adult data in general. They are, in fact, highly representative of normal adult data and constitute a reasonable standard against which to make comparisons.

**Differences in Transduction Techniques**

Much of Abbs's response is concerned with transduction technology for measuring chest wall movements. This section of his letter is devoted almost exclusively to his argument that the data from the normal control subject in Abbs et al. (1983) should not be viewed as invalid. The line of argument taken by Abbs is indirect and follows the tack that he should not be compelled to demonstrate that his normal control data are similar to those from other normal subjects because these other subjects have been studied with magnetometry rather than with inductive plethysmography (Respitrace). Surprisingly, Abbs seems not to be aware of the investigation from his laboratory by Bless, Hunker, and Weismer (1981), which compared magnetometer and Respitrace measurements during normal speech production of the normal control used as the normal control in the article by Abbs et al. (1983). This research demonstrated that magnetometer and Respitrace transduction systems yield essentially identical data. We also have compared the two transduction systems and have found that, when proper care is taken in the use of the two techniques, the differences in resultant speech production data are minor. It is important to note here that, although Abbs contends that Respitrace measurements are not affected by distortions in rib cage and abdominal compartments, his argument is based only on Sackner's (1980) theoretical conclusions, not on empirical conclusions. Several investigators with experience in using both magnetometers and Respitrace report postural artifacts with Respitrace to be equal to or greater than those encountered through the use of magnetometers (M. Goldman, personal communication, 1985; Hixon, 1985; S. Loring, personal communication, 1985; A. Putnam, personal communication, 1984; J. Smith, personal communication, 1985).

One of the transduction issues raised by Abbs pertains to what he considers to be potential time distortions between rib cage volume estimates from anteroposterior and transverse magnetometer measurements. Abbs's concern may be warranted for extremely effortful maneuvers of the chest wall (Melissinos, Goldman, Mead, Bruce, & Elliott, 1977; Sampson & De Troyer, 1982), but it can be documented to be unimportant in studies of speech production where volume accelerations of the chest wall are usually quite low in relation to those associated with effortful acts of breathing (Hixon, Mead, & Goldman, 1976). In all of the data published from our laboratory, whether from normal subjects or from subjects with various disorders, care has been taken to ensure that rib cage movement has involved only a single degree of freedom with respect to volume displacement. This has been done through simultaneous multiple-diameter measurements at various anteroposterior sites consistent with those studied by Konno and Mead (1967) as well as at various transverse locations (e.g., Hoit & Hixon, 1985). Under the occasional circumstance where there may be more than a single degree of freedom in the rib cage, such as in the case of infants, it has been routine in our laboratory to incorporate anteroposterior-transverse diameter cross-product computations in making estimates of rib cage volume change (Hixon, 1973). Magnetometry has a distinct advantage over Respitrace in matters of this nature because it determines the circumference of the chest wall and provides an output that is related to changes in both the A-P and lateral dimensions. "This is not so. The Respitrace does not measure circumference, nor does it measure diameters. Its transducer measures the average of the infinite number of cross sections through the circumference of its height. Thus, each transducer provides an output that is related to changes within an ideal cylinder described by its average cross-sectional area (Watson, 1970)."

In another of Abbs's concerns for differences in transduction techniques he brings up the issue of frequency response differences between magnetometer and Respitrace systems. Abbs contends that there is reason to believe that the many differences between the Abbs et al. (1983) normal control data and those from other normal subjects studied with magnetometers can be accounted for by differences in frequency response between the two transduction systems. Because the magnetometer data he questions are from this laboratory (Hixon et al., 1973), the issue of frequency response differences can be dealt with relatively directly. The currently used adult magnetometer system in our laboratory is custom-built and has an upper frequency response limit of approximately 15 Hz. This response is nearly identical to the upper frequency response of commercially available Respitrace systems (R. Strother, personal communication, 1985). Our magnetometer system is a fifth-generation device, each generation being a slightly improved version, in one form or another, of its predecessor. The earliest work from this laboratory (Hixon, Mead, & Goldman, 1970) has been replicated repeatedly with each generation magnetometer system added. Specifically, the results obtained with all generations of magnetometer devices used have been essentially identical, suggesting that the early work from this laboratory by Hixon et al. (1973) is fully borne out by additional observations with improved magnetometer systems over the years. In our opinion, there is no reason to believe that the differences between the normal control data of Abbs et al. and of Hixon et al. (1973) are related to any technical advantage of Respitrace over magnetometry.

Of final interest with regard to the transduction issues discussed by Abbs is electronic filtering. In our letter, we pointed out that the displacement-time tracings in the Abbs et al. (1983)
response that we "also suggest that these same signals contain frequent, rapid adjustments that reflect artifact." Putting these two points together, Abbs states that our conclusions in these two regards reflect an "illogical inconsistency." We disagree. There is no inconsistency between low-pass filtering and the appearance of frequent rapid adjustments in the output if the filter skirt is not precipitous. The rapid adjustments may be "seen," but they are not seen faithfully in amplitude. In relation to this matter, Abbs states that the signals presented by Abbs et al. (1983) "were not . . . filtered in any way." Taking Abbs's statement to be true, one is still confronted with three kinematic signals that make no sense when viewed in relation to their audio signal companion. We are at a loss to explain this discrepancy unless there is a possibility that a graphic assistant who prepared the data may have unwittingly misaligned the audio signal with the other three signals in constructing the various figures.

Rib Cage-Abdominal-Diaphragmatic Movements and Mechanics

Perhaps the most interesting part of Abbs's response is that portion in which he tries to develop a line of argument that there are more than two possible means—rib cage and abdomen—by which changes in lung volume can occur. It is this line of argument that he proposes to account for the inspirations and expirations of the normal control subject and the ataxic dysarthric subject. And it is this line of argument that Abbs contends allows him not to be accountable for showing a lung volume signal that is the sum of the rib cage and abdomen signals in the figures published by Abbs et al. Abbs states that he and his colleagues "made no claims concerning the ability to perfectly derive total lung volume from BC and AB movements." This statement does not agree with statements in the Hunker et al. article from which the data of the Abbs et al. ataxic dysarthric subject were taken. Consider the following quotation.

The chest wall can be modeled as a two component system consisting of the rib cage and the diaphragm-abdomen unit (henceforth referred to as abdomen). Each component displaces a volume as it moves. The circumferential movements of the rib cage and abdomen are proportional to their respectively displaced (sic) volumes and, in sum, reflect changes in total lung volume. Hence, RIP transduction of circumferential displacements of the rib cage and abdomen provides (sic) direct estimates of the relative volume contribution of these individual parts, as well as estimates of lung volume. (Hunker et al., 1981, p. 2)

Abbs further states with regard to the Abbs et al. article that the only claim made concerning the ataxic dysarthric subject was "that the RC-AB movements in this patient were paradoxical." The critical reader is, thus, faced with an argument ad hominem. On the one hand, Abbs asks that we discount his rib cage and abdomen data because there are more than two possible means for changing lung volume. Whereas, on the other hand, he asks that we accept the same data to support paradoxing but in this case to ignore that we have been asked to discount the data on another basis.

Fortunately, both sides of this argument can be neglected because of Abbs's misinterpretations of the observations of other investigators. Abbs states that "a certain degree of AB movement cannot be interpreted quantitatively or qualitatively (as argued by . . . [Hixon & Holt], . . .) to reflect a volumetric contribution of the diaphragm to total lung volume changes." We did not, as Abbs suggests, argue this in our letter. In fact, we said nothing on this matter. We believe Abbs abstracted this contention from statements in his own response. Specifically, he quotes the following passage from Hixon et al. (1973).

The combined displacements of the rib cage wall and the abdominal wall together reflect the total lung volume change. Each of the two parts seems to move as a unit during breathing and, despite their shared common boundary at the costal margin, there is considerable functional separation between them. (pp. 79–80)

Abbs immediately follows this quotation in his response by stating that "inherent within this conception is the notion that one can infer the movements of the diaphragm from movements of the abdominal wall (i.e., these two parts of the respiratory system are coupled in an obligatory manner)." How he drew this conclusion from the Hixon et al. quotation cited is unclear. There is nothing in the Hixon et al. (pp. 79–80) statement that even remotely suggests that they believed the diaphragm and abdominal wall were coupled in an obligatory manner. There can be no doubt that the position of our laboratory on this matter is nothing at all like Abbs has portrayed it to be. This position is explicitly stated in articles from this laboratory as far back as a decade ago. Consider the following quotation.

We have no totally satisfactory means for determining precisely the volume displacement of the diaphragm. Under certain restricted conditions the diaphragm's volume displacement is accurately reflected in abdominal wall volume displacement. Previous work has used total lung volume change as an index of diaphragmatic volume displacement (Bouhuys et al., 1966; Hixon, Siebens, and Ewanowski, 1968), a convention we choose to follow here for lack of a better substitute given the current state of the art. (Hixon et al., 1976, p. 300)

What Abbs erroneously concludes in his letter with regard to rib cage-abdominal-diaphragmatic movements and mechanics can be traced to his apparently having not understood the meaning of the work of Mead and Loring (1982), in particular, and of Loring and Mead (1982), Macklem, Macklem, and De Troyer (1983), Mead (1974), Goldman (1974), and Grassino (1974), in general. Abbs selected quotations from the work of Mead and Loring that are not germane to the arguments he attempted to develop and then used these quotations out of context to construct a meaning consistent with his misunderstanding. Contrary to Abbs's contention, the work of Mead and Loring does not contradict statements from this laboratory with regard to diaphragm-abdomen volume independence. The Mead and Loring article, in fact, provides theoretical and empirical support for the earlier statements of Hixon et al. (1976) (S. Loring, personal communication, 1985). The quotation Abbs presented from Mead and Loring (1982, p. 751) is simply a reaffirmation of what has been known about the diaphragm-abdominal pathway issue for about 15 years but had not been so elegantly quantified until the Mead and Loring publication. It seems clear that Abbs made an unwarranted inferential leap between his misconception of diaphragm-abdominal volume displacements and the notion he argues that there are means that include more than the rib cage and abdomen for changing lung volume. This leap is tied to another apparent misconception of what is meant by the following statement in Mead and Loring.

Rib cage and abdominal volume displacements measured by the method of Konno and Mead are not quite equal to the volumes swept by the surfaces of the anatomic rib cage and abdomen. The reason for this discrepancy is that the cephalic and lateral portions of the abdominal wall move with the rib cage. The Konno-Mead estimates of rib cage volume displacements include the displacements due to the expansion of the cephalic anterior and lateral wall "tented" by the rib cage, and hence they somewhat overestimate displacements of the anatomic rib cage and underestimate
displacements of the anatomic abdomen. (Mead & Loring, 1982, p. 751)

As confirmed to us by S. Loring (personal communication, 1985), this statement has nothing to do with whether or not the summation of the surface displacements of the rib cage and abdomen will equal lung volume change. They must equal lung volume change because the extrapulmonary contents of the torso are essentially incompressible (Hixon, 1972). Perhaps why Abbs misled himself in this regard is that in abstracting the quote from Mead and Loring (1982, p. 751), he took its altered form as a statement of what the original authors meant. Abbs's version of Mead and Loring's statement reads as follows.

As noted specifically by Mead and Loring (1982), the Konno-Mead technique...overestimates contributions to lung volume by the rib cage and underestimates contributions by the abdomen (cf. Mead & Loring, 1982, p. 751).

Note that Abbs deleted the adjective "anatomic" before rib cage and abdomen and failed to mention the "tent" portion of the abdomen. What Mead and Loring mean in the "anatomic" part of their quote (S. Loring, personal communication, 1985) is similar to what Bergofsky (1964) meant more than 20 years ago when he noted that the physiologic rib cage and abdomen are difficult to precisely identify during various functional activities because the two parts of the chest wall share a common, but moving, boundary. What Mead and Loring mean in the "tent" portion of their quote is that there is a zone of apposition between the diaphragm and rib cage because the diaphragm extends upward along the inner wall of the lower portion of the rib cage (S. Loring, personal communication, 1985). This is an alternate way of saying what Agostoni, Mognoni, Torri, and Saracino (1965) also said 20 years ago. Abbs has seemingly confused anatomic rib cage and abdomen with physiologic rib cage and abdomen. Because of this confusion, he has erroneously concluded that there are more than two possible means for changing lung volume and has, therefore, argued for the validity of data that show both his normal control and ataxic dysarthric subject to demonstrate inspirations and breath holdings while they are talking on expiration.

Conclusions

The speech-language pathology literature remains without a documentation of differential subsystem impairment, differential motor system impairment, and decomposition of compound respiratory movement in an individual with ataxic dysarthria. Although such a clinical entity may exist, it has not yet been demonstrated empirically.

Thomas J. Hixon
Jeannette Hoit
University of Arizona, Tucson

ACKNOWLEDGMENTS

The preparation of this letter was supported in part by a grant from the National Institute of Neurological and Communicative Disorders and Stroke (NS-21574). We thank Peter Watson for his assistance in the preparation of the manuscript and Janet Hawley-Gunckel for her conceptual contribution to the endeavor.

REFERENCES


Hoit-Dalgaard, J., Lansing, R., Plassman, B., & Hixon, T.


Received May 30, 1985
Accepted May 31, 1985