Perception & Psychophysics
1984, 35 (2), 123-136

Vowels in consonantal context are perceived more linguistically than are isolated vowels: Evidence from an individual differences scaling study

BRAD RAKERD
University of Connecticut, Storrs, Connecticut
and Haskins Laboratories, New Haven, Connecticut

The purpose of this investigation was to determine whether the presence of neighboring consonants can exert a contextual influence on vowel perception and, if so, to characterize the influence. Two experiments were carried out toward that end. In both, subjects were asked to judge the linguistic similarity relationships that held among a set of American English vowels when those vowels occurred either: (1) in isolation, or (2) in /dVd/ consonantal context. In Experiment 1, the judgments were made in response to recordings of natural speech. In Experiment 2, they were made for subjects' memorial images of vowels as elicited by written stimuli. Individual differences scaling of the outcomes of the two experiments provided evidence that supported the following conclusions: (1) Consonantal context can significantly influence vowel perception; (2) for the /dVd/ context at least, the nature of the influence is to evoke more linguistic perceptual processing of vowels than occurs when they are presented in isolation; (3) the influence is more likely to be explained in terms of properties of the stimuli presented to perceivers than in terms of any sort of knowledge that perceivers bring to bear in perceptual processing; and (4) three features of linguistic description for vowels—advancement, height, and tenseness—have particular import for vowel perception and for vowel memory.

It has long been recognized that the acoustic correlates of a vowel can vary, sometimes to a substantial degree, depending on the identity of the consonants that precede and/or follow it (e.g., House & Fairbanks, 1953; Lindblom, 1963; Stevens & House, 1963). This variation has come to be understood in terms of the fact that a talker often coarticulates the neighboring segments of an utterance (that is, overlaps their respective productions) such that the acoustic signal is jointly influenced by those segments (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). How, then, do vowel perceivers adjust to these acoustic variations? One possibility is that in many or most cases they do not. If the variations are sufficiently minor, a perceiver could simply "ignore" them and achieve an acceptably high level of performance in identifying vowels. Alternatively, the perceiving of vowels might involve certain context-sensitive perceptual strategies, analogous to those that are generally thought to be required when listeners identify the consonants of an utterance (for reviews of the evidence regarding consonantal perception see, e.g., Liberman, 1982, Liberman et al., 1967; Liberman & Pisoni, 1977).

It becomes important to determine whether or not vowel perceivers are sensitive to consonantal context because (unlike most consonants) vowels can be freed from the influences of their neighboring segments and produced as isolated utterances. A good deal is known about the perception (and production) of these isolated vowels, and there is an important issue as to how to generalize from that knowledge to other cases. To the degree that listeners are, in fact, indifferent to the acoustic variations engendered by consonantal context, isolated productions might be taken to be the canonical vowel form and their acoustic signature to be the one that best exhibits the essential information for vowel perception. However, if the perceiving of vowels does, in general, involve context-sensitive strategies, then the isolated vowel form is but one of many variants and, arguably, one of the least representative variants because it occurs infrequently in natural speech. An answer to the question of whether isolated vowels are perceived dif-

This manuscript is based on portions of a doctoral dissertation that was submitted to the University of Connecticut. I am indebted to my major adviser—Robert Verbrugge—and the members of my committee—Alvin Liberman, Michael Turvey, and Benjamin Sachs—for their counsel on every phase of the dissertation project. Much of the research was conducted at the Haskins Laboratories and funded by grants awarded to that institution (NICHD Grant HD01994, BRS Grant RR05596). I gratefully acknowledge that support, as well as the support of the University of Connecticut Research Foundation. My present affiliation is with Michigan State University. Please address reprint requests and other correspondence to Brad Rakerd, Department of Psychology, Michigan State University, Psychology Research Building, East Lansing, MI 48824.

ferently from vowels in consonantal context therefore proves to be basic to vowel research.

Previous efforts to answer this question have generally been based on comparisons of the identifiability of vowels in and out of some consonantal frame. Evidence gathered with this method has, in a number of instances, been taken to favor the view that consonantal context can significantly affect vowel perception by exerting a positive influence on vowel identification (Gottfried & Strange, 1980; Strange, Edman, & Jenkins, 1979; Strange, Verbrugge, Shankweiler, & Edman, 1976). This finding remains a subject of debate, however. It has not been observed in all studies (Macchi, 1980; Pisoni, Canell, & Simnick, 1979), and it has been challenged on grounds of being largely an artifact of the method of assessment (Assmann, Nearey, & Hogan, 1982; Diehl, McCusker, & Chapman, 1981; but see Rakerd, Verbrugge, & Shankweiler, in press, and Strange & Gottfried, 1980).

The present study complements this work by addressing the question of a consonantal influence on vowel perception with evidence of a different kind from that offered in the past. To begin with, the data collected here were judgments of vowel similarity rather than absolute identification judgments; hence, they assessed the consonantal influence with a new perceptual measure. More importantly, the resulting data were analyzed with an individual differences scaling technique that highlighted aspects of the data structure that had not been considered previously. Those aspects are: (1) the dimensions of perception that had some shared significance for the set of subjects as a whole; and (2) the relative salience that those dimensions had for the individual subjects, depending on whether they judged vowels in or out of context. It will be seen that both aspects of the scaling solution were informative about the nature of vowel perception.

EXPERIMENT 1

The starting point for an individual differences scaling analysis of vowel perception is to collect, for each subject, a matrix of what Shepard (1962a, 1962b) has called "proximities" data. These are data indexing the network of perceptual relationships that hold among a set of vowels. A triadic comparisons procedure was employed to collect those data in the present experiment. That procedure was chosen because it had proven useful in previous vowel research. More than a decade ago, Pols and his associates (Pols, van der Kamp, & Plomp, 1969; Pols, Tromp, & Plomp, 1973) assessed the perceived vowel quality of spectrally constant speechlike sounds by requiring that subjects compare triplets of stimuli on a trial. Specifically, subjects were required to judge which two members of a triplet sounded most alike to them and which two least alike. They then proceeded to a new triplet and, over trials, judged all possible stimulus combinations. This procedure yielded reliable data that were interpretable, both from a linguistic standpoint and with respect to acoustic properties of the stimuli. Others (e.g., Singh & Woods, 1970; Terbeek & Harshman, 1971) have since employed the triadic comparisons method in vowel perception studies and obtained equally satisfactory results. It was used here to compare the perception of isolated vowels with that of vowels in a consonantal context (/dVd/).

Method

Subjects

Twenty-three subjects, randomly selected from a pool of individuals registered with the Haskins Laboratories in New Haven, Connecticut, were paid to participate in Experiment 1. All of them were native speakers of English, and none had any history of hearing difficulties. It was ensured that they had no prior knowledge of the purpose of this study or of the design of the experiment. Twelve of the subjects were assigned to the isolated-vowels condition of the experiment, and 11 were assigned to the consonantal-context condition.

Stimuli

The stimuli were natural productions of 10 American English vowels: /i, ɪ, ɛ, æ, ʌ, ɑ, ɔ, o, ʊ, u/. A single male talker, who spoke a general American dialect, recorded these vowels in each of two contexts: (1) in the trisyllabic frame /hədVdə/, where the second syllable (/dVd/) was stressed, and (2) in isolation. The /dVd/ consonantal frame was chosen because it imposed certain coarticulatory constraints on the talker. In order to produce initial and final /d/ consonants, the jaw must be closed and the tongue tip sealed against the back of the teeth. Articulation of the syllable vowel, which likewise requires an appropriate parameterization of the tongue and jaw, must therefore be coordinated with that of the consonants. Presumably owing to these coarticulatory constraints, there often is a substantial degree of acoustic modulation associated with /dVd/ syllables. The stressed target syllables were flanked by destressed syllables (/hə/ and /ə/) to ensure that the consonantal-context stimuli would not be meaningful words in English.

While seated in a sound-attenuated room, the talker produced several tokens of each vowel in each context. These productions were tape-recorded, low-pass filtered at 5 kHz, digitized at a sampling rate of 10 kHz, and stored in separate computer files. Two of the tokens of each vowel were used in the experiment. In all cases, these were the first two tokens produced unless some sort of articulatory anomaly such as vocal fry or "breathiness" was audible. When an anomaly was heard in one of the first two tokens, it was replaced by the third, and if that was anomalous, by the fourth, and so on. Acoustic analyses revealed that the stimuli selected by this procedure were acoustically "normal," in that their spectral and temporal characteristics were such as might be expected on the basis of data reported by previous investigators (e.g., Lehiste & Peterson, 1961; Peterson & Barney, 1952).

Procedure

Instructions. At the outset, it was explained to the subjects that the task would be to compare their perceptions of several different vowel sounds. It was also explained that they were to base the comparison on *linguistic* aspects of those sounds. The individual subjects were left to define their own criteria for the linguistic aspects. They were, however, given the following example as an aid: "If a child and an adult were both to say the vowel /i/ or the word /did/, you would surely hear some differences between the vowel

sounds. The child's vowels would doubtless be softer, higher in pitch, and so on. On the other hand, the /i/ vowels produced by the child and the adult would also have something in common, a quality or qualities that distinguished them from other vowels like the /ɛ/ in /dɛd/ or the /ɪ/ in /dɪd/. These are the qualities that you should attend to in this experiment."

Triadic comparisons. As indicated earlier, the specific task set for subjects was that of comparing triads of stimuli. On each experimental trial, three vowels were randomly selected for presentation from the set of 10 alternatives, with the constraint that the particular triad chosen had not occurred on any previous trial. A subject was allowed to listen to these three vowels in any order and any number of times with the goal of reporting which two of the three sounded *most* alike and which two *least* alike. Over the course of the experiment, listeners judged all possible triadic combinations of the 10 vowel alternatives (120 possibilities). Note that this meant that every vowel pair was, over trials, judged in relation to every other vowel in the set.

Data were accumulated over trials according to the following scoring procedure: vowel pairs judged most alike were assigned +1 scores, and those judged least alike, −1 scores. In this way, a matrix of data indexing the perceived relationships among the 10 vowel alternatives was obtained for each subject. As an example, the matrix for a subject who rated vowels in consonantal context is shown in Table 1.

One of the virtues of the triadic-comparisons procedure was that it placed minimal memorial demands on subjects, since only three stimuli had to be dealt with on each trial and these could be played and replayed in any order needed. A second virtue was that the self-paced nature of the procedure minimized the time pressure felt by subjects.

Familiarization with the equipment and procedures. A complete testing session took about 2 h. Roughly 30 min of that time was devoted to familiarizing subjects with the equipment and procedures used to present the stimuli and record the responses.

The subjects were tested individually. Each was fitted with headphones and seated in front of a computer terminal that was housed in a sound-attenuated room. Three of the keys on the terminal triggered presentation of the appropriate stimuli for each triadic trial. After a key was pressed, the corresponding stimulus was presented through headphones at a comfortable listening level. The most alike and least alike response choices were entered via a different set of keys on the terminal. Once these choices had been made, the system advanced to the next trial.

The equipment and testing procedure were demonstrated to the subjects over a series of training trials. There were between 15 and 25 such trials, depending on the individual. Each training trial comprised a different triadic combination of stimuli sampled from the set of 120 possibilities. For the first few such trials, the experimenter operated the equipment, directing the presentation of stimuli and entering response choices. After that, control was passed to the subjects and they paced themselves. They were invited to ask questions about all aspects of the procedure. The testing session was begun only after the subjects had demonstrated competence in operating the equipment and expressed confidence about understanding the perceptual task.

Analysis of the Data

To allow for a direct comparison between conditions, a single individual differences scaling analysis was carried out on all subject data from the two experimental conditions combined. The fundamental modeling assumption of individual differences scaling is that when judging the same set of stimulus items all subjects will make reference to the same perceptual dimensions. Subjects may differ from one another in terms of the relative weight (salience) that they attach to those dimensions, but they cannot differ in terms of the identity of the dimensions themselves (Carroll & Chang, 1970; Wish & Carroll, 1974).

Consistent with this assumption, the scaling solution has two components that, together, optimally account for the data structure of the individual subject matrices. The first component, called a *group space*, is a model of what the subjects have in common. The axes of the group space are the shared perceptual dimensions, and these index a set of appropriately positioned points representing the stimulus items.[1] The second component of the scaling solution is a *weight space*, which specifies the relative salience that the several dimensions of the group space have for each subject. More formally, a subject's weight for a particular dimension reflects the amount of variance in his/her data that can be accounted for in terms of that dimension. Together, the set of weights index a subject's location in the weight space.

A noteworthy property of individual differences scaling is that it dictates the orientation in which the scaling solution must be interpreted. When attempting to relate the solution to potentially relevant factors, an investigator is obliged to make reference to the shared perceptual dimensions. These dimensions enjoy a priority because they are the ones that account for the greatest percentage of variance in several subjects' data. The presence of this interpretive restriction sets individual differences scaling apart from other variants of multidimensional scaling[2] and strengthens claims that the dimensions of its scaling solution have some psychological reality (Carroll & Wish, 1974; Kruskal & Wish, 1978; Wish & Carroll, 1974).

It has often been pointed out (e.g., Kruskal & Wish, 1978; Wish & Carroll, 1974) that since multidimensional scaling methods are intended to aid in the description and understanding of data, evaluation of the "correctness" of various scaling decisions must be based on considerations that are substantive as well as statistical. The substantive considerations have to do with such factors as the interpretability and stability of results, the statistical with the goodness-of-fit between model and data. On the basis of these considerations, it was determined that the present data were most appropriately modeled (1) in three dimensions, and (2) at the ordinal scale of measurement.

Dimensionality of the space. With individual-differences scaling, a commonly used index to goodness-of-fit is the percentage of variance accounted for (VAF) in the several subjects' data (see, e.g., Carroll & Chang, 1970).[3] Increasing the number of dimen-

Table 1
Similarities Matrix for a Subject who Rated Vowels in Consonantal Context

	i	ɪ	ɛ	æ	ʌ	ɑ	ɔ	o	ʊ	u
i										
ɪ	6									
ɛ	−2	2								
æ	−4	−4	5							
ʌ	−2	−4	7	1						
ɑ	−8	−4	1	4	1					
ɔ	−5	−3	0	1	4	5				
o	−6	−5	−2	0	2	−3	3			
ʊ	−5	1	4	−1	5	−2	1	3		
u	−3	−2	2	−1	1	−1	−4	4	8	

sions will increase the VAF index, since the model has added degrees of freedom with which to fit the data.[4] The gains tend to diminish exponentially, however, and each new increment must be weighed against the substantive considerations mentioned earlier (interpretability and stability). Figure 1 displays the VAF function for the present data when modeled in two to five dimensions. The exponential nature of the function is clear. There is a relatively large increment in VAF for the shift from two to three dimensions, a much smaller one for the shift from three to four dimensions, and a negligible decrement (see footnotes 3 and 4) for the shift from four to five dimensions.

On the basis of these statistical data, at least, it appears that three or perhaps four dimensions would be the appropriate modeling choice. In the former case, 70% of the variance would be accounted for, in the latter, 72%.

The three-dimensional solution was chosen over the four-dimensional for two reasons. First, as will be seen shortly, all three of the dimensions are linguistically meaningful and therefore interpretable, and second, they are stable in that they emerged, as well, from analyses of individual subject data collected in a memory experiment (Experiment 2) and from separate analyses of the two perceptual conditions of this experiment.

Nonmetric scaling. The decision to model these data at the nonmetric (i.e., ordinal) scale of measurement was based on two considerations. The first was simply that the subjects' task was to make ordinal perceptual judgments. On each trial, it was required that they identify the most alike and least alike vowel pairs, but it was not required that they quantify the strengths of those pairwise relationships. It is true that by summing over trials some quantification was arrived at, but it was felt that the most conservative treatment of these data was to model their ranks.

The second reason for operating at the nonmetric scale has to do with the stability of the modeling outcome. The metric/nonmetric distinction made relatively little difference with respect to the present data, but it greatly affected the outcome of modeling the results of a memory experiment (Experiment 2) that was run, in part, to clarify the findings of the present perceptual experiment. This point will be discussed in greater detail later (see the section on Analysis of the Data of Experiment 2). For now, it is enough to note that a nonmetric scaling of these data revealed structure that was stably present for both of the conditions of Experiments 1 and 2.

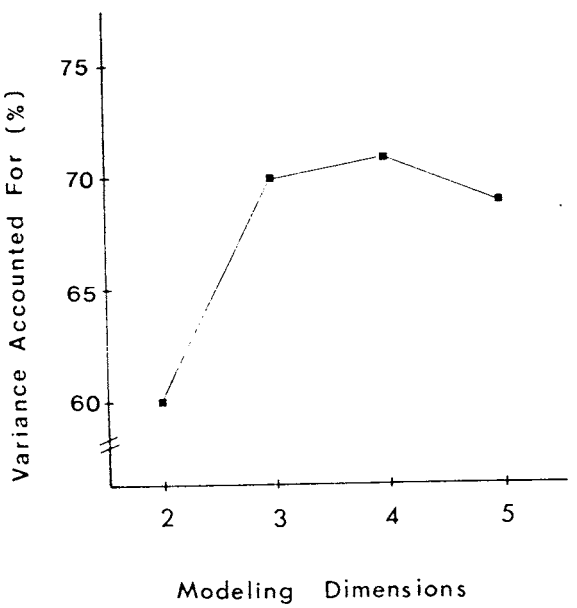

Figure 1. Percentage of variance in the perception data accounted for by modeling in from two to five dimensions.

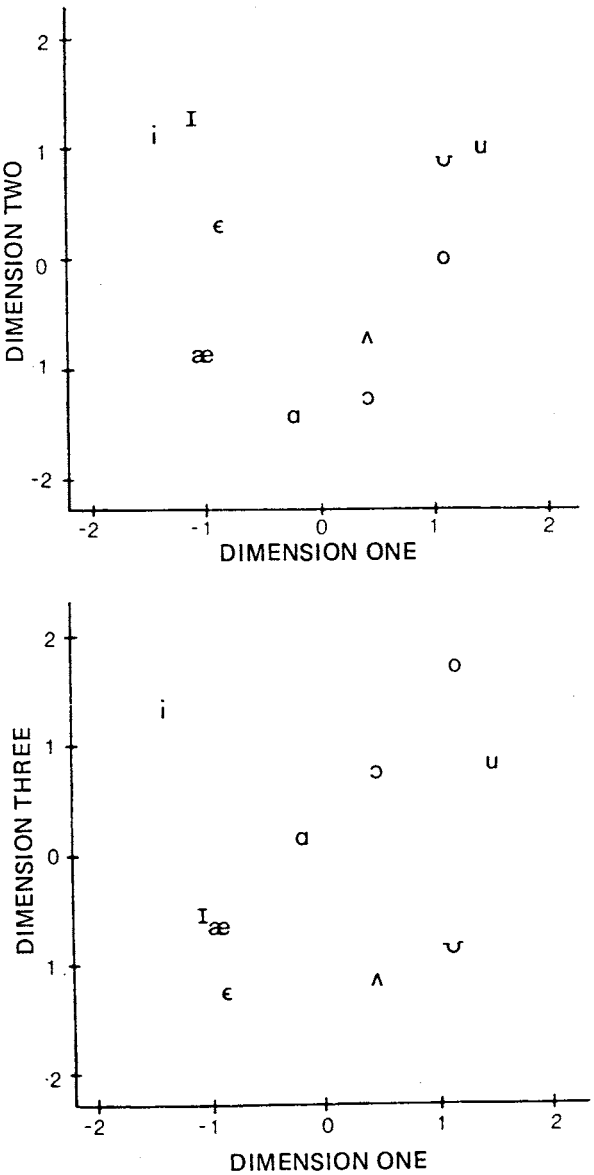

Figure 2. Group space for Experiment 1.

Results

Group Space

The group space for all subjects (isolated-vowels and consonantal-context conditions combined) is shown in Figure 2. Dimension 2 is plotted against dimension 1 in the top half of this figure, and dimension 3 against dimension 1 in the bottom half. These dimensions are orthogonal to one another and can be

interpreted independently. To do this, it is useful to visually "project" the points onto each axis and consider their ordering. In this way, it was determined that each of the dimensions of the group space corresponds closely to a traditional feature of linguistic description for vowels. Dimension 1, for instance, corresponds to a feature linguists have variously called advancement, front/back, and grave/acute. After Singh and Woods (1970), the term *advancement* will be used here. This feature distinguishes such vowels as /i, ɪ, ɛ, æ/ (seen to "project" onto the lower end of dimension 1) from such other vowels as /ʌ, ɑ, ɔ, o, ʊ, u/ (which "project" onto the upper end of the dimension). Dimension 2, in turn, corresponds to what has been called the height or compactness feature (*height* will be used here), and dimension 3 to tenseness or length (*tenseness* is the term that will be used).[5] (See, e.g., Hockett, 1958, Jakobson & Waugh, 1979, and Ladefoged, 1971, 1975, for comprehensive reviews of the vocabulary of vowel feature description.)

These features repeatedly surface when linguists try to document various aspects of linguistic behavior. To take just one example, speakers of English, though generally unaware of it, observe a grammatical rule for vowel usage that respects the tenseness feature: In English, words can end with "tense" vowels such as /i, o, u/ (there are such words as "he," "go," and "you"), but they cannot end in "lax" vowels such as /ɪ, ɛ, ʊ/ (there are no words ending in the vowel sounds heard in the *middle* of such words as "hit," "bet," and "book"). English speakers must be at least tacitly aware of this rule, since they respect it when creating new words for the language. There are countless other instances of linguistic behavior that is systematically related to not only the tenseness feature but to the advancement and height features as well (see, e.g., Jakobson & Waugh, 1979).

There is, as well, some evidence to support the claim that all of these features play a perceptual role (e.g., Hanson, 1967; Shepard, 1972; Singh & Woods, 1970). The present results are both consistent with such a claim and particularly compelling in this regard, given the nature of the scaling analysis that was employed here. Individual differences in the several subjects' data provided information that allowed for a nonarbitrary determination of the dimensions of the group space. Those dimensions shown in Figure 2 are the three that optimally accounted for the variance in the linguistic judgments made by subjects who participated in Experiment 1. The fact that each of those dimensions, in turn, corresponds closely to a linguistic feature strongly suggests that those features have some shared perceptual significance among speakers of English (see Rakerd, 1982, for an expanded consideration of this point).

Weight Space

The concern in this section will be to look at individual differences in the weighting of the dimensions of the group space, and, in particular, at differences between the isolated-vowels and consonantal-context subjects. The weight space for all subjects is shown in Figure 3. Weightings for dimension 2 are plotted against those for dimension 1 at the top of the figure; dimension 3 weightings are plotted against dimension 1 weightings at the bottom. Each "o" represents an individual from the isolated-vowels condition, and each "x" represents an individual from the consonantal-context condition.

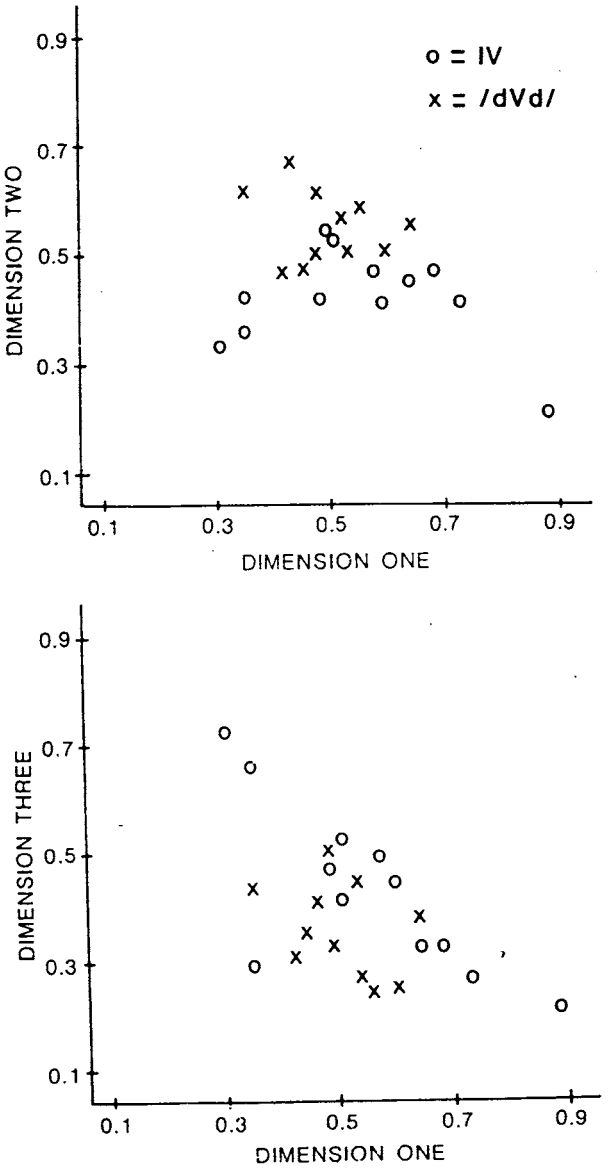

Figure 3. Weight space for Experiment 1.

The first thing to notice in this figure is that the "x"s are more tightly grouped than the "o"s in both the dimension 2 × dimension 1 and dimension 3 × dimension 1 planes. This indicates that there was less variability among consonantal-context subjects than there was among isolated-vowels subjects. In order to assess the statistical significance of this difference, a variance ratio (Snedecor's F) was calculated for the three-dimensional space as a whole. For each condition, the variance in three-space was determined as follows. First, the centroid, or average subject weight, was located in the space. Then the distance from this centroid was calculated for each subject according to the Pythagorean theorem.[6] And finally, the average squared distance was computed as the measure of variance. This measure is strictly analogous to the variance statistic (sigma squared), which is the average squared deviation about the mean for a set of numbers. The difference in variability between the two experimental conditions was, in fact, significant [$F(11,10) = 3.21$, $p < .05$].

It was observed, then, that in a perceptual task in which listeners were asked to relate a set of vowel sounds on the basis of their linguistic qualities, there was significantly greater agreement among individuals who heard vowels in a consonantal context than there was among those who heard isolated vowels. It can be inferred from this that the subjects employed somewhat different perceptual strategies in the two conditions. This, in turn, suggests the need for caution in generalizing from what is known about the perception of isolated vowels to the perception of vowels in context, a generalization that has been made in the past (e.g., Chiba & Kajiyama, 1958; Joos, 1948). Also, it would seem to warrant a methodological caveat for those who do vowel research in the future: namely, that they would do well to look at vowels in consonantal context. Given the importance of these implications, it was deemed appropriate to look at the stability of this result, and to ensure that it was not an artifact of the scaling procedure.

Regarding scaling, it is noteworthy that part of the variability in the weight space reflects differences in goodness of fit between the scaling model and the individual data. Subjects whose data were well fit by the model lie further from the origin of the space (in some direction) than do those whose data were poorly fit. It may be that the observed condition difference in variability was, in fact, a difference with respect to goodness of fit. One reason for believing this was not the case, however, is that, on average, the subject data in the two conditions were about equally well fit by the model. The average VAF for the isolated-vowels condition was 71%, and that for the consonantal-context condition was 69%, a difference that did not even approach significance.

Whether goodness of fit was significantly different for the two groups or not, it was certainly a source of variation that is of limited interest here. Therefore, the data was transformed to "factor out" its influence. A subjects' weight on a dimension is the square root of the percentage of variance accounted for by that dimension. The total variance accounted for (VAF) can thus be computed for any individual by squaring the weights and summing over all three dimensions. Between-subject differences in goodness of fit can, in turn, be compensated for by normalizing the data with respect to this VAF value. The most straightforward strategy for doing this is to divide a subjects' *squared* dimension weights by VAF and to take the square roots of the resultant dividends to be the adjusted weights.

These new values index subjects in the weight space shown in Figure 4. It can be seen to reflect statistical

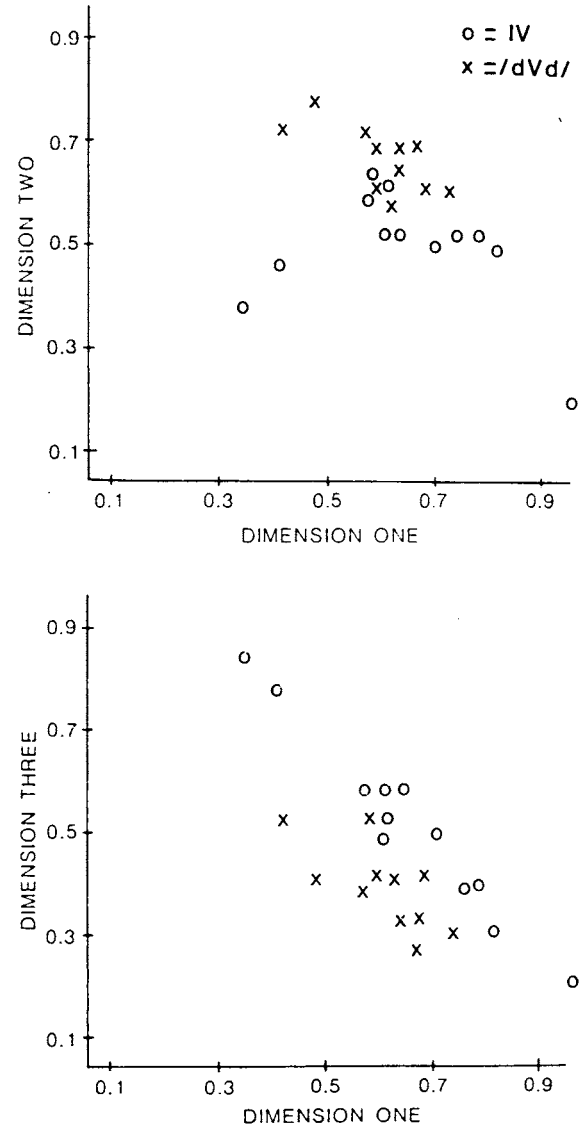

Figure 4. Adjusted weight space for Experiment 1.

compensation for goodness-of-fit differences between subjects, in that the original weights (shown in Figure 3) have been "compressed" along lines extending out from the origin of the space. Despite this compensation, the condition difference in subject variability remains significant [$F(11,10) = 3.18$, $p < .05$].

It proves to be the case, then, that even when individual differences in the goodness of fit of the model are factored out, there remains a significant difference between the two experimental conditions with respect to subject variability in the weight space.[7] This finding clearly supports the view that vowels in a consonantal context are perceived differently from vowels outside a consonantal context. It also hints at the nature of the difference, at least for the present experiment. The task set for subjects was to relate a number of different vowel sounds on the basis of their linguistic qualities. Subjects who heard vowels in a /dVd/ consonantal context exhibited significantly greater agreement as to what those linguistic relations were than did their counterparts who heard isolated vowels. Thus, it can be said that one of the effects of context was to stabilize linguistic judgments across subjects.

It is useful to take note of the nature of the stability; the consonantal-context subjects clustered toward *center* of the weight space, which indicates that they attached roughly equal weight to all three linguistically meaningful dimensions of the group space. Notice that this need not necessarily have been the case. Between-subject agreement would have been equally high for this condition had the clustering occurred out near one of the "corners" of the weight space, whereupon one or another of the perceptual dimensions could have been said to predominate. It turned out, however, that all three of the dimensions had substantial perceptual import for consonantal-context subjects.

The situation was markedly different for the isolated-vowels subjects. While several members of that group were positioned near the center of the weight space, most of them were in more "extreme" locations. The data for this latter group were largely accounted for in terms of perceptual sensitivity to just one or two of the linguistic dimensions. And it should be noted that the one or two dimensions that predominated were different for different individuals. Thus, it can be seen that isolated-vowels subjects were not constrained to perceive the stimuli in terms of the full set of linguistic dimensions. To the contrary, they attended to the dimensions in a piecewise manner, while the consonantal-context subjects integrated the dimensions in a more linguistically appropriate way.

Summary

The individual differences scaling analysis revealed two ways in which vowels in consonantal context can be said to have been perceived more linguistically than were isolated vowels. First of all, judgments about the linguistic qualities of vowel sounds were significantly more stable across subjects when the vowels were in context. And second, three linguistically meaningful dimensions of vowels were more integrated in perception when vowels were in context.

Discussion

How is this effect of consonantal context to be understood? It will be argued that, broadly speaking, there are two classes of accounts that might be brought to bear to explain it and that the results of Experiment 2 lend at least suggestive support to one over the other. The first class, which will be called *knowledge-based* accounts, turns on a subject's understanding of certain regularities in the occurrence of vowel categories. The second class, called *stimulus-based accounts*, turns on a subject's sensitivity to properties of the stimuli themselves. To illustrate the differing character of knowledge-based and stimulus-based accounts, several examples of each are provided below.

Knowledge-Based Accounts

One plausible example of a knowledge-based account is motivated by the fact that in English vowels most generally occur in some consonantal context. The context condition of the present experiment might therefore be expected to engage linguistic processing most effectively. Such an argument is encouraged by the observation that frequency of occurrence does positively affect performance on a number of other indices to language skill, such as the reaction time to identify a word as being in one's lexicon (Forster & Chambers, 1973). On the other hand, a wealth of linguistic phenomena resist explanation in such terms. Witness, in this regard, the fact that readers of Japanese name colors more rapidly when they are written in kana (a representation of the phonologic structure of the language) than in kanji (a logographic representation) even though the latter form is seen much more often (Feldman & Turvey, 1980).

A knowledge-based account of a rather different sort could draw on the fact that certain phonological rules for vowel usage are specific to consonantal context. One such rule that has already been mentioned is that "tense" vowels can occur at the ends of words but "lax" vowels cannot. A listener asked to make judgments about the linguistic quality of vowel sounds might therefore have a difficult time with an isolated "lax" vowel like /ɪ/, since no such isolated vowel is allowed in English. Singh and Woods (1970) advanced just such an argument to account for the fact that they found no evidence that tenseness had perceptual significance for listeners who rated the relative similarity of a set of isolated American En-

glish vowels. On the basis of the present findings, however, that failure might possibly be attributed to the fact that those investigators averaged their data over subjects prior to scaling. For certain isolated-vowels subjects in the present experiment, the tenseness dimension was particularly salient. For others, however, it had little or no salience. Averaging over all subjects, then, might "wash out" any statistical evidence of the significance of the tenseness feature.

Investigators (Assman et al., 1982; Macchi, 1980) have also pointed to this phonological restriction on isolated vowel usage as a potential explanation for the recurring observation that vowels can be identified more accurately in consonantal context than out (e.g., Gottfried & Strange, 1980; Strange et al., 1979; Strange et al., 1976). This cannot explain the phenomenon in full, however, since Strange et al. (1979) have also observed a consonantal influence in CV syllables (e.g., "be") in which "lax" vowels are as phonologically impermissible as they would be in isolation.

Whatever the outcome of these individual debates, the tenor of this sort of knowledge-based account is clear: to the degree that listeners are sensitive (either consciously or unconsciously) to the fact that a phonological rule of English proscribes the occurrence of certain vowel sounds in isolation, those listeners' linguistic judgments may be affected.

Recently, it has been shown that a knowledge of how speech sounds are written may have an effect as well. For instance, listeners will more rapidly detect the rhyming quality of spoken words when those words are spelled alike (e.g., "fight"/"right") than when they are not ("you"/"two") (Seidenberg & Tanenhaus, 1979). It is perhaps relevant, then, to note that the subjects in the present experiment were literate and therefore had had a great deal of experience in reading and writing vowels. In at least one previous study (Diehl et al., 1981) it has been suggested that such experience can lead to a perceptual bias in favor of consonantal context.

Knowlege-based accounts of the consonantal influence have in common the fact that they look to a subject's long-term experience with stimuli of a particular type. Such accounts would have it, for example, that extended acquaintance with frequently occurring items, or with certain phonological or orthographic regularities regarding those items, explains the perceptual effect that was observed in Experiment 1. Thus, the "locus" of knowledge-based effects is at some remove from immediate stimulation. That is to say, these accounts have much more to do with the sorts of accumulated knowledge that might be brought to bear in processing stimulus information than they do with the information itself. Not so the more stimulus-based accounts that will now be considered.

Stimulus-Based Accounts

As examples of stimulus-based accounts, consider two that are motivated by the fact that (as is typically the case) the isolated-vowels stimuli tended to exhibit relative spectral constancy over their course (only the vowels /o/ and /u/ were noticeably diphthongized), while the vowels in /dVd/ context tended to be marked by more or less continuous formant frequency change.

One reason why formant change may have been the source of the enhanced linguistic processing for vowels is that its presence or absence may have differentially affected the duration of a vowel's representation in what have been called auditory and phonetic memory stores. By hypothesis, the former memory preserves a relatively unprocessed "neural analog" of the acoustic signal, and the latter preserves features of the input that are specifically relevant to speech. Fujisaki and Kawashima (1969, 1970; see also Pisoni, 1973) have pointed to the differential presence of vowels and consonants in these memories as a potential psychoacoustic basis for the observation that the former (particularly isolated vowels) tend to be less categorically perceived than the latter (Fry, Abramson, Eimas, & Liberman, 1962; Pisoni, 1973). The present argument would simply extend this reasoning to a perceptual difference that holds between two classes of vowels, those in and out of consonantal context.

An alternative reason why formant change might be expected to engage linguistic processing particularly effectively is that properties of speech signal are, most generally, dynamic in character (Liberman, 1982). This is a consequence of the fact that the several segments of an utterance tend to impose competing demands on the articulators, making it necessary for talkers to interleave their productions in a series of rapid articulatory gestures. By the laws of physical acoustics, these gestures result in an assortment of dynamic modulations of the signal. Owing to this fact, a speech perceiver might be expected to be particularly attuned to any sort of acoustic change.

Although other stimulus-based accounts might be advanced, it should be apparent from just these that the focus of all such accounts will be on some acoustic property or properties of the stimulus set.

How to Distinguish Between the Two Classes of Accounts

An essential difference between knowledge-based and stimulus-based accounts has to do with the degree to which they reflect a sensitivity to properties of the stimulation. Since knowledge-based factors turn on long-term experience with stimuli of a type, they should be relatively little affected by the immediate experience obtained through any particular encounter with a stimulus. Stimulus-based factors,

by contrast, are expressly defined in terms of stimulus properties. It should be possible, then, to gain some evidence as to the "locus" of the consonantal influence observed here by looking at a case in which the relative contribution of immediate stimulation is reduced. If knowledge-based factors were critical to the present result, they might be expected to be manifest there as well. If, instead, stimulus-based factors were most important here, the consonantal influence should be diminished. Experiment 2 provides a case relevant to these predictions.

EXPERIMENT 2

In this experiment, subjects' memories for vowel sounds were examined with a procedure analogous to that employed in Experiment 1. The subjects were asked to *imagine* vowels—as occurring in isolation or in /dVd/ context—and to make judgments about the linguistic relationships among the images. It was expected, first of all, that an analysis of these judgments might help to clarify the results of Experiment 1. There it was found that the presence of consonantal context had the effect of evoking somewhat more linguistic perceptual processing of vowels than occurred in its absence, and it was concluded that while a number of different accounts of this effect could be put forth, broadly speaking, these turned either on various properties of the stimuli themselves or on properties having more to do with the occurrence and recurrence of vowels as meaningful categories in English. If these latter, more knowledge-based factors are the critical ones, then it might be expected that the presence or absence of consonantal context would affect the outcome of this memory experiment no less than it did that of the perception experiment, because vowel usage rules, orthographic regularities of vowel transcription, and so on, remain in force here. On the other hand, to the degree that stimulus properties are critical to the effect, the condition difference should be reduced here (or perhaps eliminated), since vowel memory is at some remove from the acoustic stimulation.

The results of this second experiment could also prove useful in a second way: they could point to the dimensions of organization for subjects' long-term memorial representations of vowels. A question arises, for instance, about the number of such dimensions. Are there three, as in Experiment 1, and if so, are these the same three linguistically meaningful dimensions that were found to have perceptual import? And it may be asked whether the same dimensions are utilized in the same way by different subjects.

Method

Stimuli

The stimuli for Experiment 2 were written analogs of the spoken stimuli used in Experiment 1. That is to say, they were orthographic representations of both isolated vowels and vowels in a trisyllabic frame in which the medial syllable was stressed /dVd/. The stimulus set comprised the same 10 vowels that were used in the perception experiment. Table 2 presents a summary of all stimuli.

In English orthography there are numerous ambiguities with respect to the spelling of vowel sounds. The letters "oo", for example, stand for the vowel /u/ in the word "tool" and for /ʊ/ in the word "book." There are indications that these spelling ambiguities can affect listeners' perceptions of vowels (Assmann et al., 1982), and, although the present experiment was not strictly perceptual, it was thought advisable to devise vowel spellings that were unique to each sound. These are presented in the second and third columns of Table 2. In all cases, the vowels were spelled with two-letter sequences. These sequences were presented alone for isolated vowels and embedded in the frame AD _ _ DA for the trisyllables. In the latter case, the subjects were told to read each stimulus as a three-syllable nonsense word, the first and last syllables of which consisted of unstressed schwa (/ə/) vowels and the middle syllable of which was stressed /dVd/.

The subjects were familiarized with the new orthography with the aid of a training sequence. This sequence paired each written vowel form with three English monosyllabic words containing the vowel sound that the form was meant to represent (see Table 2). These words were selected so as to be similar to, but distinct from, both the isolated and /dVd/ contexts employed in the experimental test.

The test series consisted of triads of stimuli presented in three adjacent columns. All possible triadic combinations of the vowels were included in each series. The order of occurrence of triads was randomized, as was the assignment of the words of each triad to the columns.

Procedure

Instructions. As nearly as possible, the instructions for Experiment 2 paralleled those for Experiment 1. It was explained to the subjects that their task would be to imagine a number of different vowel sounds and to make linguistic comparisons of the images. A sense of what it would mean to make linguistic comparisons was again provided by the example of distinguishing an adult's and a child's productions of the vowel /i/ from their productions of vowels such as /ɛ/ and /ɪ/ (see the instructions section of Experiment 1).

The triadic comparisons testing procedure was explained in detail. Subjects were told that they would be given a test series (either the isolated-vowels series or the consonantal-context series) and a cover sheet. The cover sheet had a small slit cut in it to allow the viewing of only one test line (a trial) at a time. The procedure was: (1) to move the cover sheet down the test page, thereby exposing the three stimuli of a trial; (2) to make a triadic comparison among the images of the three vowels represented on the line; (3) to write down the column numbers of the most alike and least alike vowel

Table 2
Stimuli for Experiment 2

	"Spellings"		English Exemplars		
Vowel	IV	/dVd/	1	2	3
i	EE	ADEEDA	eat	heel	brief
ɪ	IH	ADIHDA	it	him	brim
ɛ	EH	ADEHDA	egg	hen	bread
æ	AE	ADAEDA	at	ham	brash
ʌ	UH	ADUHDA	up	hull	brush
ɑ	AH	ADAHDA	odd	hop	bronze
ɔ	AW	ADAWDA	ought	haul	brawn
o	OH	ADOHDA	oat	home	broach
ʊ	UU	ADUUDA	oomph	hood	brook
u	OO	ADOODA	ooze	hoop	broom

Note—IV = isolated vowels; /dVd/ = vowels in consonantal context.

pairs; and (4) to proceed to the next trial. It was emphasized to subjects that they were under no time pressure. To the contrary, they were instructed to proceed at whatever pace they found comfortable, with the constraint that they not look back at any trial once it had been completed.

Orthographic training and administration of a pretest. Prior to the test, the subjects were given extensive work with the orthographic training sequence. The experimenter first read the sequence aloud, pointing out potential errors to be avoided. Next, the subjects were allowed to ask questions about the spellings, and then the sequence was read aloud a second time. Finally, the subjects were told to study the sequence on their own for as long as was needed to commit the spellings to memory.

At the end of the individual study sessions and before the actual test series were presented, the subjects were asked to complete a pretest designed to assess competency with the new orthography. The pretest was straightforward: The subjects were presented a randomized list of written vowel stimuli and asked to give three examples of English words that contained the vowels indicated. The examples they gave had to be different from those used in the training sequence. The subjects' test results were omitted from the data analysis if they made more than one error on the pretest.

Subjects

Thirty-three undergraduates, enrolled in an introductory psychology course at the University of Connecticut, participated in the experiment for course credit. These individuals were native English speakers. They had no prior knowledge of either the purpose of the experiment or its design. On the basis of their performance on the pretest, six subjects were eliminated. Twelve of the remaining 27 subjects were in the isolated-vowels condition, 15 were in the consonantal-context condition.

Analysis of the Data

Since Experiment 2 was designed to test subjects' memories for vowels in a way that paralleled the perception test of Experiment 1, there was reason to expect that the most appropriate scaling solution for the present data might be the three-dimensional one that appeared earlier. After various modeling alternatives were examined, it was concluded that, with one methodological exception to be noted below, this was in fact the case.

Dimensionality of the space. The percentage of variance accounted for was computed as a function of the number of modeling dimensions. This function had roughly the same shape as its counterpart in Experiment 1, an observation consistent with the expectation that a three-dimensional modeling of these data might prove as appropriate here as it had in Experiment 1. The VAF comparison with Experiment 1 also showed that at each dimension level these memory data were somewhat less well fit by the model than were the perception data, which is to say that the data were somewhat "noisier" here. This is not surprising. In the perception test, the items presented to subjects were highly familiar (spoken English vowels) and were, in fact, the perceptual objects of interest. Here, by contrast, the items presented were rather unfamiliar vowel spellings, which only mediated contact with the memory images that were the true objects of study.

Nonmetric scaling. It has been pointed out that the stability of a modeling outcome must be considered when making decisions about scaling (see Wish & Carroll, 1974, for a discussion of this point). With respect to the present study, this consideration bore more directly on the decision to perform nonmetric individual-differences scaling, as against a more commonly used metric procedure such as INDSCAL (Carroll & Chang, 1970). For certain of the analyses of Experiment 1 in particular, the metric/nonmetric modeling distinction made little or no difference in the outcome.[8] However, this could not be said to be the case in Experiment 2; modeling these data at the metric scale resulted in an uninterpretable group space. At the nonmetric scale, on the other hand, the group space was not only interpretable, but also quite evidently relatable to the group space of Experiment 1.

This perception/memory difference in what might be called "measurement level" may be interesting in its own right. It suggests that the memory space for vowels is a sort of nonlinearly transformed version of the perceptual space. Interval relationships among the vowels hold in perceptual space but not in memory. On the other hand, the relatively noisy character of the memory data has already been noted, and the "measurement level" difference between the perception and memory experiments may simply reflect task variables. Whatever the true state of affairs, the approach taken in this study has been to model all data at the more conservative nonmetric level.

Starting configuration. It proved to be the case that the three dimensions of the group space could not be interpreted as originally modeled. They correspond neither to the linguistic features of advancement, height, and tenseness, as they had in Experiment 1, nor to other recognized features of articulatory or acoustic descriptions for vowels. This was equally the case for the two- and four-dimensional group spaces.

The scaling procedure was therefore rerun in three dimensions with the group space of Experiment 1 taken as a *starting configuration*.[9] This was, in effect, a test of the appropriateness of that earlier group space as a model for memory data structure. It did prove to be an appropriate model, as evidenced by the fact that it fit the memory data nearly as well as had the original, uninterpretable, three-dimensional solution (mean VAF was 59% with the starting configuration, 61% without it).

Results

Group Space

The group space for all subjects who participated in the two conditions of the memory experiment is shown in Figure 5. It is quite evidently similar to the group space for Experiment 1 (Figure 2). In the dimension 2 × dimension 1 plane, only the vowel /ʌ/ has shifted its position substantially: it is "higher" and more "fronted" in the present analysis. In the dimension 3 × dimension 1 plane, the only vowels that moved noticeably are /æ/ and /a/. The former can be seen to have taken on a dimension 3 value that is somewhat more "tense," and the latter one that is more "lax." These shifts do not substantially alter the overall configuration, however. On the whole, then, it can be said that this combined group space for the memory experiment does not differ substantially from that for the perception experiment. In both cases, the nonarbitrary axes of the space correspond to the linguistic features of advancement, height, and tenseness.

Weight Space

Since the group spaces are similar for Experiments 1 and 2, it is interesting to see how the weight spaces compare. Indeed, a primary motivation for carrying out Experiment 2 was to determine whether the condition difference in dimension weightings seen for perception would be manifest in memory as well. A look at Figure 6, which displays the combined weight space for Experiment 2, indicates that it was not.

In Experiment 1, subjects in the consonantal-context condition were consistent with one another in

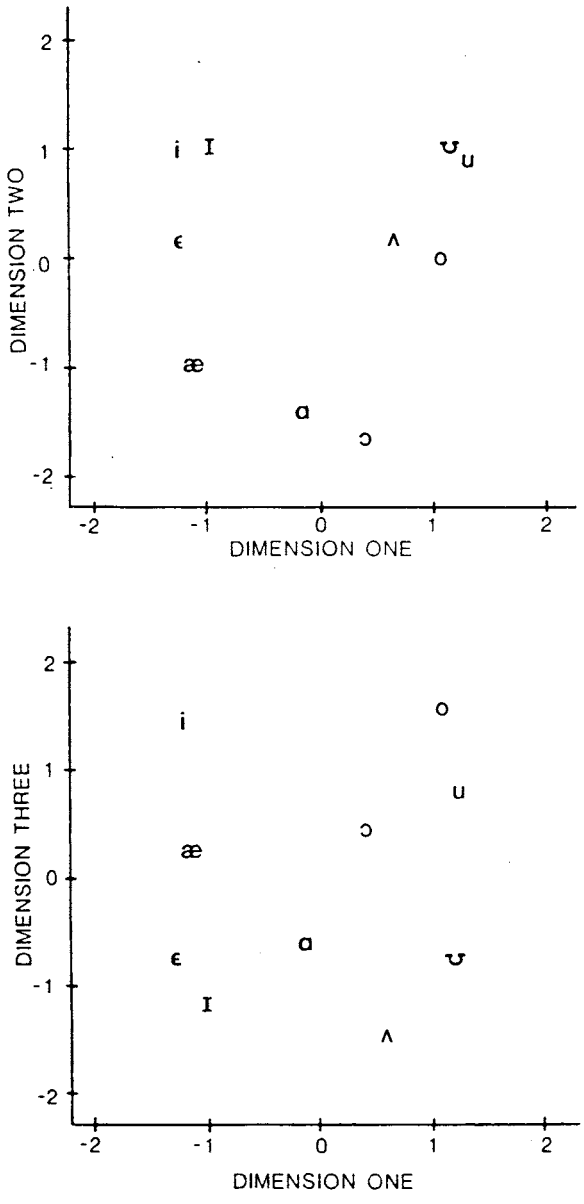

Figure 5. Group space for Experiment 2.

attaching substantial weight to all three linguistically meaningful dimensions of the group space, while isolated-vowels subjects were quite variable, with different individuals weighting different dimensions disproportionately. Here, by contrast, subjects in *both* conditions behaved in a fairly comparable way: they clustered toward the center of the weight space (roughly as did the consonantal-context subjects of Experiment 1).[10] It turned out that isolated-vowels subjects were, if anything, less variable in exhibiting this pattern than were their consonantal-environment counterparts—the opposite result from that observed in Experiment 1. (This trend was not significant in the original weight space shown in Figure 6 [$F(14,11) = 2.27$], but was in a weight space adjusted to compensate for goodness-of-fit differences among subjects (cf. Experiment 1) [$F(14,11) = 4.30$, $p < .01$].)

Clearly, the pattern of dimension weightings obtained for memory judgments made at some remove from the acoustic stimulation is substantially different from that obtained in perception. This strongly suggests that stimulus-based factors were critical to the perceptual influence of consonants observed in Experiment 1.

Discussion

In Experiment 1, it was concluded that the /dVd/ context had the effect of evoking more linguistic perceptual processing of vowels than occurred in isolation. There are a number of knowledge-based accounts of why this might have been the case, including the fact that vowels more frequently occur in consonantal context than they do out, that certain phonological rules are specific to consonantal context in English, and that regularities (and irregularities) in English orthographic representations of vowels may differ with context. Since these knowledge-based factors reflect a history of experience with vowels as meaningful categories in English, it might be expected that they would have an influence in this vowel memory experiment as well. However, the variance analyses of the subject weights indicate that they did not. It can be at least tentatively concluded, therefore, that the consonantal influence in perception had more to do with stimulus-based factors than with knowledge-based factors.

In Experiment 1, a close correspondence was observed between three features of linguistic description for vowels (advancement, height, and tenseness) and the three dimensions of the group space. The fact that individual-differences scaling showed these to be the dimensions that optimally accounted for variance in the several subjects' data was taken as particularly strong evidence that those linguistic features have some significance for the perception of vowels. A related point can now be made with respect to vowel memory, although it must be somewhat tempered by the reservation that the present analysis was initiated by a starting configuration. The orientation of axes for the resulting group space was nevertheless dictated by the character of individual subject data, and the observed correspondence between the linguistic features and the dimensions of this space strongly suggests that listeners' memories for vowels are, at least in some measure, organized in a way that respects those features. Thus, there appears to be a consistent recurrence of the features in perception, memory, and in linguistic behavioral data such as those having to do with grammatical rules for vowel usage.

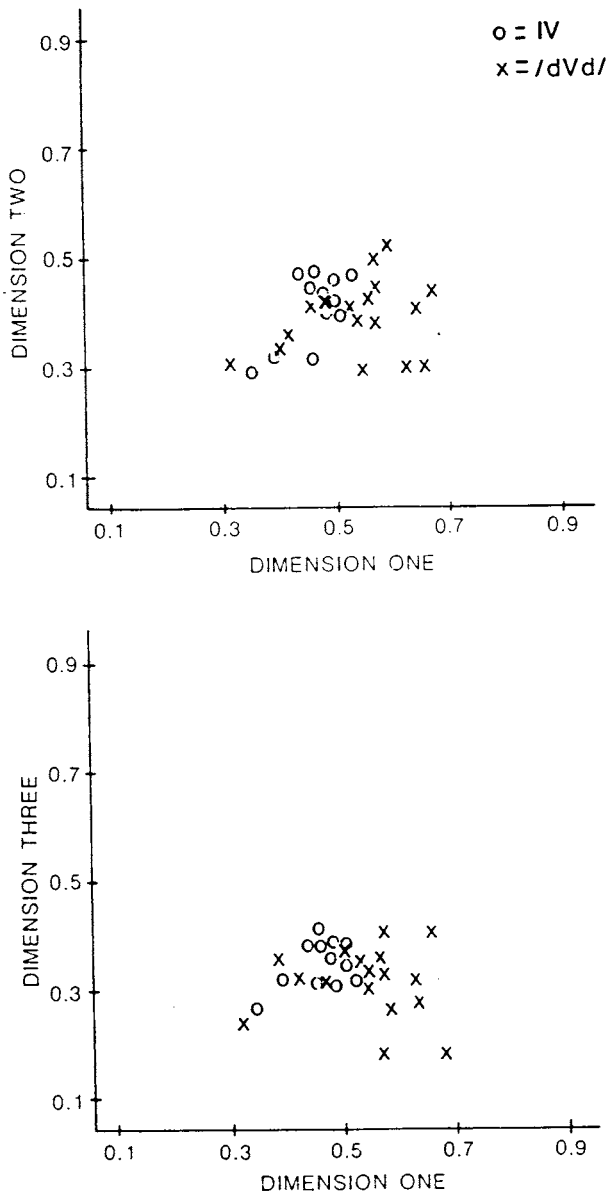

Figure 6. Weight space for Experiment 2.

It is important to recognize that altogether different results might have been obtained. First of all, the several subjects participating in this memory experiment might have exhibited no consistent pattern of responding at all, in which case the model would have failed to account for a reasonable percentage of the variance in the data and the dimensions of the group space would have been uninterpretable. Alternatively, to the degree that subjects behaved consistently, they might have done so in a way that made little or no sense from a linguistic standpoint. Since the stimuli of this experiment were presented by eye, subjects might, for example, have made their judgments on the basis of visual features of the input, but they did not.

SUMMARY AND CONCLUSIONS

This study was motivated by an interest in the question of whether vowel perception is greatly influenced by the consonantal context in which a vowel occurs. A good deal is known about the perception (and production) of isolated vowels, and an answer to this question of consonantal influence will determine how researchers generalize from that knowledge base. To the degree that the influence on perception is minor, the isolated vowel form might reasonably be viewed as canonical (since it is unencumbered by any context effects at all) and its acoustic signature might be taken to be composed of the essential information for vowel perception. On the other hand, if consonantal context is found to significantly affect the perception of vowels, then the isolated form can only be considered to be one variant of the vowel and, given the infrequency of its occurrence in natural speech, an arguably unrepresentative variant. Caution would therefore be required in generalizing from what is known about it.

The results of the present study clearly support the latter position. Vowels were here found to be perceived significantly differently in consonantal context than they were in isolation. One aspect of that difference was that listeners exhibited greater agreement with one another about the linguistic relationships that held among a set of vowels when those vowels were in /dVd/ context than when they were in isolation. A second aspect was that with isolated vowels listeners attended in a piecewise manner to three different vowel dimensions, while with vowels in context they integrated those dimensions in a way that was more consistent with other aspects of linguistic behavior.

These findings have been interpreted as indicating that /dVd/ context had the effect of eliciting more linguistic perceptual processing of vowels than occurred when they were presented in isolation. To the degree that this interpretation is appropriate, it follows that those who do linguistic research on vowels in the future would do well to examine them in some consonantal context.

REFERENCES

Assmann, P. F., Nearey, T. M., & Hogan, J. T. (1982). Vowel identification: Orthographic, perceptual, and acoustic aspects. *Journal of the Acoustical Society of America*, 71, 975-989.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35, 283-319.

Carroll, J. D., & Wish, M. (1974). Multidimensional perceptual models and measurement methods. In E. C. Carterette & M. P.

Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press.

CHIBA, T., & KAJIYAMA, M. (1958). *The vowel and its structure.* Tokyo: Phonetic Society of Japan.

DIEHL, R. L., MCCUSKER, S. B., & CHAPMAN, L. S. (1981). Perceiving vowels in isolation and in consonantal context. *Journal of the Acoustical Society of America*, **68**, 239-248.

FELDMAN, L. B., & TURVEY, M. T. (1980). Words written in Kana are named faster than the same words written in Kanji. *Language and Speech*, **23**, 141-147.

FORSTER, K., & CHAMBERS, S. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, **12**, 627-635.

FRY, D. B., ABRAMSON, A. S., EIMAS, P. D., & LIBERMAN, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, **5**, 171-189.

FUJISAKI, H., & KAWASHIMA, T. (1969). *On the modes and mechanisms of speech perception* (Annual Report of the Engineering Research Institute, **28**, pp. 67-73). Tokyo: University of Tokyo.

FUJISAKI, J., & KAWASHIMA, T. (1970). *Some experiments on speech perception and a model for the perceptual mechanism.* (Annual Report of the Engineering Research Institute, **29**, pp. 207-214). Tokyo: University of Tokyo.

GOTTFRIED, T. J., & STRANGE, W. S. (1980). Identification of coarticulated vowels. *Journal of the Acoustical Society of America*, **68**, 1626-1635.

GUTTMAN, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, **33**, 469-506.

HANSON, G. (1967). Dimensions in speech sound perception: An experimental study of vowel perception. *Ericsson Technics*, **23**, 3-175.

HOCKETT, C. F. (1958). *A course in modern linguistics.* New York: Macmillan, 1958.

HOUSE, A. S., & FAIRBANKS, G. (1953). The influence consonant environment has upon the secondary acoustic characteristics of vowels. *Journal of the Acoustical Society of America*, **25**, 105-113.

JAKOBSON, R., & WAUGH, L. R. (1979). *The sound shape of language.* Bloomington: Indiana University Press.

JOOS, M. (1948). Acoustic phonetics. *Language*, 24, Suppl., 1-136.

KRUSKAL, J. B., & WISH, M. (1978). *Multidimensional scaling.* Beverly Hills: Sage.

LADEFOGED, P. (1971). *Preliminaries to linguistic phonetics.* Chicago: University of Chicago Press.

LADEFOGED, P. (1975). *A course in phonetics.* New York: Harcourt Brace Javonovich.

LEHISTE, I., & PETERSON, G. E. (1961). Transitions, glides, and diphthongs. *Journal of the Acoustical Society of America*, **33**, 268-277.

LIBERMAN, A. M., (1982). On finding that speech is special. *American Psychologist*, **37**, 148-167.

LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. (167). Perception of the speech code. *Psychological Review*, **74**, 431-461.

LIBERMAN, A. M., & PISONI, D. B. (1977). Evidence for a special speech-perceiving subsystem in the human. In T. H. Bullock (Ed.), *Recognition of complex acoustic signals.* Berlin: Springer-Verlag.

LINDBLOM, B. E. F. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, **35**, 1773-1781.

MACCHI, M. J. (1980). Identification of vowels spoken in isolation and in consonantal context. *Journal of the Acoustical Society of America*, **68**, 1636-1642.

PETERSON, G. E., & BARNEY, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, **24**, 175-184.

PISONI, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, **13**, 253-260.

PISONI, D. B., CANELL, T. D., & SIMNICK, S. S. (1979). Does a listener need to recover the dynamic vocal tract gestures of a talker to recognize his vowels? *Journal of the Acoustical Society of America*, **65**, S6. (Abstract)

POLS, L. C. W., TROMP, H. R. C., & PLOMP, R. (1973). Frequency analysis of Dutch vowels from 50 male speakers. *Journal of the Acoustical Society of America*, **53**, 1093-1101.

POLS, L. C. W., VAN DER KAMP, L. J. TH., & PLOMP, R. (1969). Perceptual and physical space of vowel sounds. *Journal of the Acoustical Society of America*, **46**, 458-467.

RAKERD, B. (1982). *Vowels in consonantal context are perceived more linguistically than isolated vowels: Evidence from an individual differences scaling study.* Unpublished doctoral dissertation, University of Connecticut.

RAKERD, B., VERBRUGGE, R. R., & SHANKWEILER, D. P. (in press). Monitoring for vowels in isolation and in a consonantal context. *Journal of the Acoustical Society of America.*

SEIDENBERG, M. S., & TANENHAUS, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, **5**, 546-554.

SHEPARD, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, **27**, 125-140.

SHEPARD, R. N. (1926b). The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika*, **27**, 219-246.

SHEPARD, R. N. (1972). Psychological representation of speech sounds. In E. D. David & D. P. Denes (Eds.), *Human communication: A unified view.* New York: McGraw-Hill.

SINGH, S., & WOODS, D. R. (1970). Perceptual structure of 12 American English vowels. *Journal of the Acoustical Society of America*, **49**, 1861-1866.

STEVENS, K. N., & HOUSE, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*, **6**, 111-128.

STRANGE, W., EDMAN, T. R., & JENKINS, J. J. (1979). Acoustic and phonological factors in vowel identification. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 643-656.

STRANGE, W., & GOTTFRIED, T. L. (1980). Task variables in the study of vowel perception. *Journal of the Acoustical Society of America*, **68**, 1622-1625.

STRANGE, W., VERBRUGGE, R. R., SHANKWEILER, D. P., & EDMAN, T. R. (1976). Consonant environment specifies vowel identity. *Journal of the Acoustical Society of America*, **60**, 213-224.

TAKANE, Y., YOUNG, F. W., & DE LEEUW, J. (1977). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika*, **42**, 7-67.

TERBEEK, D., & HARSHMAN, D. (1971). Cross-language differences in the perception of natural vowel sounds. *UCLA Working Papers in Phonetics*, **19**, 26-38.

WISH, M., & CARROLL, J. D. (1974). Applications of individual differences scaling to studies of human perception and judgment. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press.

NOTES

1. The "appropriateness" of point positioning has to do with the distances *between* the points. Those distances should, as nearly as possible, be ordered in a manner that reflects order in the perceptual data (see Footnote 9).

2. With other scaling methods, such as those designed for the analysis of single matrices of data (e.g., Guttman, 1968; Shepard, 1962a, 1962b); it is necessary to perform a post hoc rotation of the scaling solution in order to bring it into any sort of interpretable orientation. The particular rotation performed is necessarily

shaped by an investigator's intuitions about the appropriate dimensions of interpretation and is, correspondingly, vulnerable to the challenge that some other dimensions would have been equally (or more) appropriate had some other rotation been carried out. It is just this post hoc rotation that is precluded with individual differences scaling (Carroll & Chang, 1970; Wish & Carroll, 1974).

3. Since this is the most commonly discussed index to fit, it is the one that will be considered here. Nevertheless, it should be noted that the data were, in fact, scaled with a procedure (ALSCAL, designed by Takane, Young, & de Leeuw, 1977) that, for computational reasons, optimizes a related, but slightly different, index called SSTRESS. This undoubtedly accounts for the slight decrement in the VAF function seen at five dimensions (there was no such decrement in the SSTRESS function). Solutions obtained by optimizing SSTRESS are extremely similar to those obtained with alternative individual differences scaling methods (Takane et al., 1977).

4. VAF might not increase with an increase in dimensionality if a scaling algorithm was halted after a fixed number of iterations or, more commonly, if it was halted due to the encounter of a "local minimum" in the optimizing function (also see Footnote 3).

5. In English, the tenseness and length feature labels are not quite so interchangeable as are, say, height and compactness. The "tense" vowels are generally "long" vowels as well, but there is one notable exception: The vowel /æ/. This vowel is phonologically "long," yet a usage rule treats it as "lax" in that it cannot appear in open position. With respect to the group space, /æ/ is likewise grouped with the "lax" vowels along dimension 3, which makes the choice of the tenseness label particularly appropriate for this dimension.

6. The Pythagorean theorem holds that the distance between two points in a three-dimensional space will be equal to the square root of the sum of the squared distances between those points' coordinates along the three reference axes. Hence, the distance between Subject 1 (indexed by the coordinates x_1, y_1, z_1) and the centroid for all subjects in the same condition (indexed by x_c, y_c, z_c) was computed with the equation:

$$\text{distance} = [(x_1 - x_c)^2 + (y_1 - y_c)^2 + (z_1 - z_c)^2]^{1/2}.$$

7. It is also noteworthy that when the two conditions of the experiment were modeled *separately*, this significant difference in subject variability remained (see Rakerd, 1982, for the analysis).

8. In Experiment 1, the group and weight spaces for the metric analysis of the combined data were virtually identical to those for the nonmetric analysis shown in Figures 2 and 3. When the conditions were modeled separately (Rakerd, 1982), the metric and nonmetric solutions did differ, at least in detail.

9. To get a sense of what employing a starting configuration entails, it is important to understand how the scaling procedure operates more generally. An optimal fit to a set of data is achieved by successively adjusting the stimulus configuration over a series of iterations. The scaling procedure halts when the improvement achieved on any given iteration is less than some specified tolerance value. The adjustment that is made to the configuration amounts to moving the individual stimuli around in the group space in a way that is sensitive to the modeling shortcomings of the existing configuration. If, for example, the vowels /i,ɪ,ɛ/ were currently positioned in the space such that the distances among them were ordered as follows:

/i/-----/ɪ/----------/ɛ/,

and yet most subjects ranked vowel-pair similarity such that /i-ɪ/ was judged *less* similar than /ɪ-ɛ/, then on the following iteration of the procedure there would be a shift in the positioning of /ɪ/ to correct the mismatch between model and data, that is,

/i/----------/ɪ/-----/ɛ/.

Since the scaling procedure is capable of making such adjustments, it is possible to start with a truly random stimulus configuration and gradually, over trials, to move to one that fits a data set quite well. It is equally possible, however, to start with a configuration that, for a priori reasons, might be expected to fit the data closely from the outset. To the degree that it does, the procedure will make only minor improvements and will halt in a relatively small number of iterations (because those improvements will be less than the halting tolerance value). Because of this feature, it is possible, in effect, to test out the appropriateness of a particular starting configuration for the individual differences modeling of any set of data.

10. Since scaling accounted for a relatively smaller percentage of the variance in the memory data that it accounted for in the perception data (Experiment 1), all weights are, on average, smaller here (i.e., closer to the origin of the weight space).

(Manuscript received February 3, 1983;
revision accepted for publication November 2, 1983.)