An auditory Stroop effect
with judgments of speaker gender

ELIZABETH J. GREEN and PAUL J. BARBER
Birkbeck College, London WC1E 7HX, England

A series of "auditory Stroop" experiments is described. These demonstrate an effect of stimulus words presented on speed of judgments of speaker gender and, conversely, an influence of speaker gender on judgments of words presented. In an experiment in which responses to speaker gender were semantically related to, but not identical with, stimulus words, the auditory Stroop effect was attenuated but remained in evidence. Potential parallels between this auditory paradigm and the visual Stroop color/word effect are explored, and it is suggested that the Stroop effects in the two modalities operate along broadly similar lines. The search for a common causal mechanism would therefore be justified.

The Stroop effect, the interference of the color-name aspect of a color/word stimulus with naming its incongruent ink color, is a well-documented phenomenon for the visual modality (Dyer, 1973; Jensen & Rohwer, 1966; Stroop, 1935).

Cohen and Martin (1975) and Hamers and Lambert (1972) have reported an analogous effect for auditory stimuli when subjects judged whether the pitch of words was high or low. Stimulus words could be either congruent or incongruent with their pitch (e.g., the word "high" or "low" presented in a high- or low-pitched voice). In these auditory Stroop tests, pitch therefore corresponded to the ink color of visual Stroop stimuli and the auditorily presented stimulus word to the written or printed color name.

The potential of the auditory Stroop paradigm for use as a research tool has already been demonstrated. Hamers and Lambert (1972) studied bilingualism in a task in which subjects responded to pitch in the same or in a different language from that in which stimulus words were presented, while Cohen and Martin (1975) used the paradigm to investigate hemispheric asymmetries. Cohen and Martin reported a greater degree of interference of the semantic aspect of the stimulus words with judgments of their pitch when they were presented to the left cerebral hemisphere rather than to the right, showing the auditory Stroop paradigm to be sufficiently sensitive to reflect the effects of cerebral lateralization.

Further investigations of the auditory Stroop effect would therefore be valuable. In particular, it would seem important to establish whether interference occurs when judgments of an aspect of word stimuli other than pitch are required. If auditory

Stroop interference were not present for other dimensions, an explanation of existing effects might lie in the phenomenon of pitch processing. Alternatively, if further auditory Stroop effects were found, one means of understanding their etiology would be to investigate whether they truly correspond to the visual case or whether similarities between the auditory and visual paradigms are merely superficial.

EXPERIMENT 1

This experiment therefore investigated whether auditory Stroop effects could be found in a task involving judgments of a speech dimension other than voice pitch. The experimental stimulus words were "man" and "girl," spoken by a male or a female speaker, and subjects were required to judge speaker gender. Congruent and incongruent word/speaker-gender combinations were generated as follows: The word "man" in the male voice and the word "girl" in the female voice were the so-called congruent. stimuli; the word "girl" in the male voice and the word "man" in the female voice represented the incongruent stimuli. "Girl" rather than "woman" was used because it was monosyllabic.

The question of whether the auditory Stroop effects reported so far in the literature are due to interference of stimulus words with pitch judgments when these elements are incongruent, to facilitation of judgments when they are congruent, or both, has still not been fully resolved, but the issue is of some importance if an analogy between the visual and auditory paradigms is to be pursued. The existence of facilitation effects in visual Stroop research has been considered crucial in evaluating certain theories of Stroop interference. The relevance of Morton's logogen model (Morton, 1964, 1979) to Stroop effects has been noted by Morton (1969) and

Address reprint requests to Paul Barber, Department of Psychology, Birkbeck College, Malet Street, London WC1E 7HX, England.

Warren (1972), for example, but Stirling (1979) has claimed that this model cannot adequately account for facilitation of responding in the visual Stroop task.

In their auditory experiments, Cohen and Martin (1975) used both pure tones and the neutral (with respect to pitch) word "bar" as control stimuli. Both these types of control stimulation produced response latencies intermediate in length, longer than those to congruent word/pitch combinations but shorter than those to incongruent combinations, suggesting that facilitation as well as interference contributed to the auditory Stroop effects. Hamers and Lambert (1972), on the other hand, with only tones as controls, found that responses to tones were faster than to word stimuli, whether congruent or incongruent, implying that facilitation did not occur in their experiments.

For these gender-judgment tasks, the words "mill" and "game" were chosen as controls. Words beginning with the same initial consonants as "man" and "girl" were selected because the apparatus incorporated a timing device sensitive to the level of sound at the onset of the stimulus words as the signal to begin timing responses. This could otherwise have been activated earlier, leading to longer measured response times by some consonants than by others.

If both facilitation and interference were contributing to any auditory Stroop effect found, then responses to these stimuli should be faster than those to incongruent man/girl stimulus combinations and slower than those to congruent stimuli, paralleling the finding of Cohen and Martin (1975). The fact that "mill" and "game" began with the same initial phonemes as the man/girl stimuli, however, provided grounds for supposing that these controls might be less than wholly neutral. In the visual Stroop test, Dalrymple-Alford (1972) reported interference from non-color-related distractor words that began with the same initial letters as color names used in the experiment. In view of this finding, it was possible that a "pseudo-Stroop" effect might result in this auditory paradigm. If subjects associated the word "mill" with "man" and the word "game" with "girl," either through hearing the stimulus set of words or via the labels "man" and "girl" on the response keys, then faster mean responses to the word "mill" in the male voice and "game" in the female voice than to the word "mill" in the female voice and "game" in the male voice may result.

Stimuli were presented to subjects in one of two ways: Either all four stimulus types (congruent, incongruent, and the two control conditions) were intermingled in the stimulus list (random presentation) or the congruent and incongruent stimuli were presented in one block with the control stimuli in a separate block (blocked presentation). Besides increasing the number of possible orders of stimulus presentation subjects could receive, these two methods of presentation were devised as a precautionary measure to test the robustness of any effects found. Poulton (1973) has observed that a subject's response is determined not only by the present stimulus, but also by the available range of stimuli or responses, or both. It was possible, therefore, that subjects in the random presentation condition, confronted with a stimulus vocabulary of four words, might focus attention differently from blocked presentation subjects who were presented with stimulus vocabularies of only two words for each block; the word aspect of the stimulus might be more salient in the first case than in the second, for example. Few relevant design studies have been reported for the visual paradigm, so this represented an attempt to further an understanding of strategies involved in performing Stroop tasks generally.

Method

Design. A split-unit design was used, with stimulus-presentation sequence (blocked vs. random) as the between-subjects factor, and with stimulus type (Stroop vs. control words) and gender-word congruence (congruent vs. incongruent) as within-subjects factors. Stroop-type stimulation was provided by the words "man" and "girl," and control stimulation by the words "mill" and "game." The Stroop words "man" and "girl" spoken, respectively, in a male and female voice were "congruent" stimuli, and when spoken, respectively, in female and male voices were "incongruent" stimuli. Similarly, the control words "mill" and "game" in male and female voices, respectively, were "pseudocongruent" stimuli, and "pseudoincongruent" when in female and male voices, respectively. Table 1 summarizes the stimulus combinations used in Experiment 1. Half of the subjects were assigned to the blocked presentation condition, and half to the randomized condition. Assignment to groups was otherwise random.

For the blocked presentation condition, half the subjects, randomly chosen, received the Stroop block followed by the control block. The remainder received the stimuli in reverse order. For the random presentation condition, half the subjects received the two stimulus subsequences in the order in which they were recorded on the tape. The remainder received the second subsequence before the first.

Subjects. Thirty-two subjects were drawn from a panel of volunteers comprising students, ex-students, and technical/clerical staff of the University of London. Subjects were paid 60 pence for participation in the experiment. All subjects reported them-

Table 1
Words Used to Form the Voice-Gender by Stimulus Type
Combinations: Experiment 1

Stimulus Type		Voice Gender	
		Male	Female
Stroop	Congruent	man	girl
	Incongruent	girl	man
Control	Pseudocongruent	mill	game
	Pseudoincongruent	game	mill

selves to be free from any hearing deficit and used English as their first language. Their ages ranged from 18 to 45 years.

Stimulus tape recordings. A stimulus set of eight was used comprising the words "man," "girl," "mill," and "game" in a single male and a single female voice (as shown in Table 1). Six examples of each member of the stimulus set were then used to give sequences of 48 words/tape.

For both tapes, these experimental sequences were recorded in two parts (24 words each) separated on the tapes by a 30-sec interval. The interstimulus interval from onset of one word to onset of the next was otherwise 5 sec in all cases.

For random presentation, Tape 1 was used, for which each subsequence of 24 words made up half the total number of each stimulus type.

For blocked presentation, Tape 2 was used, with all Stroop stimuli occurring in the first subsequence followed by all control stimuli in the second.

Subsequences for both tapes were ordered pseudorandomly, with weak sequential constraints balancing numbers of repetitions and alternations of speaker gender and of word said over two consecutive stimuli. These were identical for Tapes 1 and 2 so that comparisons between blocked and random presentations could be made with a reduced possibility that differential sequential effects would contaminate the data.

Following the procedure adopted by Cohen and Martin (1975), familiarization and practice sequences were also recorded. For Tape 1, prior to the two halves of the experimental list, one example of each of the eight members of the stimulus set was recorded (familiarization sequence) followed by a randomly ordered practice list comprising two examples of each member of the stimulus set.

For Tape 2, familiarization sequences and practice lists were recorded before each half of the experimental list. Prior to the Stroop list, the familiarization sequence comprised the words "man" and "girl" in the male and female voices, and the randomly ordered practice list comprised two examples of each member of this stimulus subset. Similarly, before the control list, a familiarization sequence of the words "mill" and "game" in the male and female voice and a randomly ordered practice list of two examples of each of these stimuli were recorded.

Apparatus. Tapes were played to subjects on a Sony UHER mono tape recorder and presented binaurally through EAGLE SE-1 headphones.

During the experiment, the subject sat at a table on which were two platform response keys. He or she was asked to rest one hand lightly on each key throughout the experiment. Keys were labeled with the words MAN and GIRL, respectively, in block capital letters printed in black. Half the total number of subjects in each condition, randomly selected, had the right-hand key labeled MAN and the left labeled GIRL. For the remainder, this mapping was reversed. Responses by pressing one or the other of the keys, using a whole-hand movement, were requested. The onset of each stimulus word on tape triggered an electronic switch that started a millisecond timer. This was stopped when the subject pressed one or the other of the keys.

Procedure. The RTs to each stimulus were manually recorded by the experimenter, who was seated to the rear of the subject. During the experiment, the experimenter had a list of the correct gender responses for each trial, by means of which errors were monitored, but did not have a record of whether a given trial was congruent, incongruent, and so on.

Subjects were instructed to respond to the speaker gender of each stimulus word while attempting to ignore the word that was actually said in each case. They were told to be both as quick and as accurate as possible. Each subject was asked "just to listen" to the familiarization sequence and then received the practice sequence before the stimulus lists proper. For the random presentation condition a 1-min pause was allowed between the two subsequences of the tape, and for the blocked presentation condition a similar pause was allowed between blocks.

Results

For each subject, the mean RT was found for correct responses under each congruence × stimulus type combination. An analysis of variance, with one between-subjects factor and two within-subjects factors, was performed on these mean RTs.

The blocked vs. random stimulus presentation factor was not significant as a main effect or in interaction with other variables.

The effect of congruence was highly significant [$F(1,30) = 39.07$, $p < .001$], with congruent RTs faster than incongruent RTs. The interaction of congruence and stimulus type was also highly significant [$F(1,30) = 15.82$, $p < .001$], and the mean RTs for this interaction are shown in Table 2. To clarify the nature of this interaction, comparisons between means were made by post hoc Tukey ratio tests. Accordingly, it was found that congruent Stroop stimuli produced significantly more rapid responding than did incongruent Stroop words [$q(2,30) = 13.08$, $p < .01$], and pseudocongruent control stimuli produced significantly faster responding than did pseudoincongruent words [$q(2,30) = 4.95$, $p < .01$]. There was no significant difference between mean RTs to congruent Stroop and pseudocongruent control stimuli [$q(2,30) = 1.06$, n.s.], but incongruent Stroop words produced significantly slower responses than did pseudoincongruent controls [$q(2,30) = 6.89$, $p < .01$].

Discussion

A strong auditory Stroop effect with judgments of speaker gender is demonstrated by Experiment 1; mean RTs were significantly faster to congruent than to incongruent word/speaker gender combinations over blocked and random stimulus presentation conditions. Although the difference between mean responses to incongruent and to congruent stimulation (64 msec) is much smaller than the 120 to 250 msec reported by Cohen and Martin (1975), it was extremely reliable across subjects, with 31 of 32 subjects demonstrating the effect.

A smaller (25 msec) and slightly less reliable, but still statistically significant, "pseudo-Stroop effect" was observed in the form of faster mean responses to pseudocongruent than to pseudoincongruent control stimuli. This suggests that Dalrymple-Alford's

Table 2
Mean Reaction Times (in Milliseconds) for Experiments 1-4

Experiment	Stimulus Type			
	Congruent	Incongruent	Pseudo-congruent	Pseudo-incongruent
1	401	465	406	431
2	353	421	367	399
3	385	464	413	424
4	454	480	445	480

(1972) finding of interference caused by distractor words having the same initial letters as potential responses in the visual Stroop test also applies in the auditory case. It also implies that conclusions about the relative contribution of facilitation and interference to the auditory Stroop effect based on data from these particular control words must be drawn with caution. Although the expected interaction between congruence/incongruence and Stroop/control stimulus type did occur, post hoc testing revealed that the incongruent Stroop stimuli resulted in significantly slower responding than did the pseudo-incongruent controls, whereas the difference between mean responses to congruent and to pseudocongruent stimuli was not significant. The evidence for the existence of facilitation effects in the data is therefore weaker than that for interference, despite the fact that congruent Stroop stimulation did produce faster responding than did all control stimulation. However, the possibility cannot be ruled out that facilitation of pseudocongruent responding was instrumental in producing the pseudo-Stroop effect and that a facilitation in the Stroop effect proper was thereby concealed.

The blocked vs. random presentation condition had no noticeable influence on the outcome of Experiment 1, suggesting that this variable did not effectively induce subjects to use different processing strategies. The failure of stimulus set size to alter an auditory Stroop effect is consistent with the evidence on the visual Stroop test. Golden (1974) reported no significant differences in performance whether three, four, or five colors and color names were used, while, more recently, Simon and Sudalamaithu (1979) have demonstrated a visual Stroop effect using only two colors and color names.

EXPERIMENT 2

A replication study based on the paradigm of Experiment 1 was run. This was done because of the small number of stimulus sequences used in the first experiment (two tape recordings affording four different orders of stimulus presentation). To eliminate the possibility that the Stroop and pseudo-Stroop effects reported in Experiment 1 resulted from the particular characteristics of the speaker voices used or the control words selected, a new stimulus list was constructed that incorporated two new speaker voices and different control words.

Short-range practice effects were examined by requiring subjects to attend two sessions. It was not anticipated that this would produce radical changes in the auditory Stroop effect in view of the small effects of practice observed in the visual color-word test (Jensen & Rohwer, 1966). Indeed, Smith and Nyman (1962) found most practice effects occurred in the first few trials and reported that performance was asymptotic after about five trials.

Method

The task, apparatus, and general procedure for Experiment 1 were used. Subjects received two presentations of the same stimulus material, separated by an interval of 30 min.

Design. In addition to the two within-subjects factors from Experiment 1, a within-subjects factor of sessions (Session 1 vs. Session 2) was included.

Subjects. Eight subjects were drawn from the panel of volunteers described for Experiment 1.

Stimulus tape recording. A new tape recording, Tape 3, was made. This used a new order of stimuli with the same sequential constraints as applied to Tapes 1 and 2. Two new speaker voices were used, and different control words, "mile" and "gold," were employed in place of "mill" and "game," respectively.

Procedure. This was the same as for the randomized presentation condition of Experiment 1, with one additional balancing feature arising from the sessions factor. At the second session, the four subjects who, for the first session, had been randomly chosen to receive the two subsequences of the stimulus list in the order in which these had been recorded on the tape received the second subsequence first, followed by the first subsequence. For the remaining four subjects, the presentation orders over sessions were reversed.

Results

The data were treated as for Experiment 1, and an analysis of variance with three within-subjects factors was performed on the mean RTs.

The session was not significant as a main effect or in interaction with other variables. As for Experiment 1, the effect of congruence was highly significant [$F(1,7) = 39.64$, p $<$ .001] with congruent RTs faster than incongruent RTs. The interaction of congruence and stimulus type was also significant [$F(1,7) = 7.36$, p $<$ .05]. There was a significant difference between congruent and incongruent Stroop RTs [$q(2,7) = 10.46$, p $<$ .01] and a smaller but still significant difference between the pseudocongruent and pseudo-incongruent RTs [$q(2,7) = 4.92$, p $<$ .05].

Discussion

The results of Experiment 2 closely replicate the findings of Experiment 1, testifying to the robustness and reliability of the auditory Stroop effect with voice-gender stimuli. The basic Stroop effect was 68 msec (64 msec in Experiment 1) and the pseudo-Stroop effect was 32 msec (25 msec in Experiment 1).

The possibility that performance in this task changes with extended practice cannot be ruled out, but short-range effects seem to be negligible. This evidence is consistent with that from the visual paradigm that performance on the Stroop task is asymptotic after only a few trials.

EXPERIMENT 3

This experiment aimed to establish whether it was possible to obtain "reverse Stroop" interference in the auditory modality. A paradigm identical to that of Experiment 1 was used, except that subjects were required to respond to the stimulus word spoken and not to the speaker gender. Reverse Stroop in-

terference would occur if speaker gender exerted an influence upon the time taken to make a decision about which word was spoken.

Reverse Stroop interference beyond minimal levels has typically proved difficult to obtain in the visual test in which subjects are asked to read the words of conflicting color/word stimuli while ignoring the ink color in which they are written (Dyer, 1973). Exceptions to this are cases in which the word aspect of the stimulus has been attenuated in some way, for example, by reducing its legibility (Gumenik & Glass, 1970).

However, the direction of visual Stroop interference appears to depend on the time relations involved in the processing of the ink color and the word aspect of the stimulus. Although words are normally read faster than colors can be named (Fraisse, 1969), colors can be physically matched faster than words—a fact that Murray, Mastronadi, and Duncan (1972) showed to result in reverse Stroop interference in a visual task in which a matching response rather than a naming response was required. The presence or absence of reverse auditory Stroop interference may therefore depend upon whether mapping an auditorily presented word onto the appropriate response key can be accomplished more rapidly than judging the gender of the speaker of that word, a question to be determined empirically.

Method

The same apparatus and procedure were used as for the blocked presentation condition in Experiment 1, with the exception that the subjects were instructed to respond to the word said instead of speaker gender.

Design. A split-unit design was used, with one between-subjects variable of stimulus type (Stroop vs. control stimuli) and one within-subjects variable of stimulus congruence (congruent vs. incongruent).

Subjects. There were 16 subjects from the pool described for Experiment 1.

Procedure. The general procedure was the same as for Experiment 1. The response keys were labeled with the words MAN and GIRL, respectively, for the Stroop stimulation group and with the words MILL and GAME, respectively, for the control group.

Tape 2 was used to present stimuli, the Stroop group receiving the Stroop stimulus block only and the control group receiving the control stimulus block only. Subjects were instructed to respond to the word said as quickly and accurately as possible and to attempt to ignore the gender of the speaker of each word.

Results

The data were treated the same as for Experiment 1, and an analysis of variance with one between-subjects factor was performed on the mean RTs.

The effect of congruence was highly significant [$F(1,14) = 20.69$, $p < .001$], and the interaction of congruence with Stroop vs. control stimulation was also significant [$F(1,14) = 11.39$, $p < .01$].

The effect of Stroop vs. control stimulation was negligible [$F(1,14) = .01$, $p = .91$]. Post hoc Tukey

ratio testing revealed that congruent Stroop words produced significantly faster RTs than did incongruent Stroop words [$q(2,14) = 11.29$, $p < .01$], but mean pseudocongruent and pseudoincongruent RTs were not significantly different [$q(2,14) = 1.572$, n.s.].

Discussion

Reverse auditory Stroop interference of the same order as that of the auditory Stroop effect of Experiment 1 (79 msec compared with 64 msec) was demonstrated. The evidence for a mirroring of this reverse auditory Stroop effect in the control stimuli was slender, a difference of only 11 msec between mean responses to pseudoincongruent and pseudocongruent stimuli being evident. The significant interaction between stimulus congruence and stimulus type suggests that both interference and facilitation effects might be instrumental in producing this reverse Stroop interference, although the same reservations about such a conclusion that applied in the case of Experiment 1 (concerning the control stimuli used) must be taken into account. Overall RTs to word spoken in this experiment were no more rapid than responses to speaker gender in Experiment 1 (see Table 2). This contrasts with the visual case in which reading of the color name of the Stroop stimuli is almost invariably accomplished more rapidly than naming the colors (Seymour, 1979; Stroop, 1935). Nothing like the dominance of reading over naming (Fraisse, 1969), which is almost certainly a factor in the unilateral nature of visual Stroop interference, was observed in this auditory case. A balance in performance may be critical in determining whether Stroop and reverse effects occur. Experimental devices that bring about a closer correspondence between component processing times (e.g., Gumenik & Glass, 1970) offer a way of examining this issue. Support for this notion may also be drawn from Stroop's (1935) finding that visual reverse Stroop interference can be observed in subjects for whom color identification is highly practiced.

EXPERIMENT 4

Although the two paradigms are not identical, data from Experiments 1 and 2 suggest that the auditory Stroop effect resembles the visual one beyond the superficial similarities already noted.

Contrary to such a view, Dyer (1973) suggested that the auditory Stroop effects reported by Hamers and Lambert (1972) may simply have been an artifact of subjects' tendency to repeat or shadow an auditorily presented word. In the case of congruent stimuli, shadowing the stimulus word would result in the production of a correct response and spuriously rapid RTs could thus result. In the case of incongruent stimuli, shadowing the stimulus word

would give rise to an incorrect response, which would be recorded as an error and would therefore fail to be included in the analysis.

Dyer's suggestion was tested in the present experiment by asking subjects to make responses to the same auditory Stroop stimuli used in Experiments 1, 2, and 3, but in terms of the *names* of the speakers who made the stimulus tape recordings rather than by use of the words MAN and GIRL, which were also stimulus distractor words. The task was described to subjects as distinguishing between "Dave's voice" and "Joan's voice," and response keys were labeled "Dave" and "Joan," respectively.

Proper names were deliberately used for this task in order to test a further claim by Dyer (1973) that auditory Stroop interference could not be generated using two named speakers saying their own name or that of their counterpart (Dyer & Severance, cited in Dyer, 1973). This failure to find any effect may have been due to the fact that associations between the voices and the proper names of the speakers were insufficiently overlearned by subjects, so the present experiment sought to demonstrate that it was not the use of proper names per se which was responsible.

The task thus involved the use of responses semantically related to, but not identical with, the Stroop stimulus words. Similar tasks have been employed in modifications of the visual Stroop test: Harrison and Boese (1976), for example, established that visual Stroop interference did occur when a response semantically related to the ink color of a Stroop color/word stimulus was required (e.g., "SKY" in response to the blue ink of a distractor word "red"), although this interference was smaller than when the responses were color names or when the distractor words were also color associates.

If, therefore, auditory Stroop interference follows the same pattern that visual does, interference should still occur with the Dave/Joan responses demonstrating that the auditory effect results, at least partly, from semantic processes. Like the visual effect, however, auditory interference may be attenuated in this paradigm by the loss of that part of the total interference that is normally related purely to the response stage of the process. In the control stimuli, a pseudo-Stroop effect like that of Experiment 1 might be found if this resulted, partially at least, from the similarity of the initial phonemes of the control words "mill" and 'game" with the Stroop *stimulus* words "man" and "girl" rather than with the response labels.

Method

The apparatus and procedure were the same as for the random presentation condition in Experiment 1 except that subjects were instructed to respond to "Dave's voice" and to "Joan's voice" with the response keys labeled accordingly.

Results

The data were treated the same as for all previous experiments, and an analysis of variance for two within-subjects factors was performed on the mean RTs.

The effect of congruence was highly significant $[F(1,15) = 25.05, p < .001]$. Neither the effect of stimulus type nor its interaction with congruence approached significance (see Table 2).

Data from this experiment were compared with those from the random presentation condition of Experiment 1, and an analysis of variance for one between-subjects and two within-subjects variables was performed to compare the two sets of data. The between-subjects factor was response labeling (MAN/GIRL vs. DAVE/JOAN), with the within-subjects factor the same as for the previous analysis. The effect of response labeling was not significant $[F(1,30) = .76, p > .05]$. The effect of congruence was highly significant $[F(1,30) = 52.21, p < .001]$, and the three-way interaction of response labeling × stimulus type × stimulus congruence was significant $[F(1,30) = 5.77, p < .025]$. No other main effect or interaction was significant.

Discussion

Auditory Stroop interference did still occur when responses were labeled with words semantically related to gender that were not identical to, or physically similar to, any of the stimulus words. The magnitude of the effect was smaller than in the data from Experiment 1 (26 msec compared with 64 msec); however, the pseudo-Stroop effect also remained in evidence and was not attenuated (35 msec compared with 25 msec in Experiment 1), suggesting that stimulus-related, rather than response-related, processes are responsible for its existence.

The diminution in the size of the auditory Stroop, but not the pseudo-Stroop, effect apparently accounted for the loss of the interaction of stimulus type with stimulus congruence, which was present in the Experiment 1 data but not in those of Experiment 4 (see Table 2). The presence of the three-way interaction in the comparison of Experiments 1 and 4 (response labeling/stimulus type/stimulus congruence) also appears to have resulted from this change in the size of the Stroop but not of the pseudo-Stroop effect with the relabeled DAVE/JOAN responses. In particular, congruent Stroop responding in Experiment 4 appears to have been slowed down

relative to the Experiment 1 congruent Stroop data. There is therefore no evidence for facilitation of congruent Stroop responding in the Experiment 4 data. It is possible that, if facilitation effects do occur in some auditory Stroop paradigms, they result from peripheral, response-related processes, such as the tendency to repetition or shadowing postulated by Dyer (1973). Interference itself, on the other hand, appears to be at least partly the result of semantic processes.

The results from this auditory experiment are consistent with those of Harrison and Boese (1976) in the visual modality, demonstrating a reduction in interference but not the elimination of the effect when responses semantically related to, but not identical with, stimulus distractor words are required of subjects.

GENERAL DISCUSSION

These experiments demonstrate the existence of an auditory Stroop effect for judgments of speaker gender. Auditory Stroop interference is not, therefore, a phenomenon confined to judgments of voice pitch, and its broader theoretical implications should be considered.

Our findings suggest that the correspondence with the visual Stroop phenomenon is more than superficial. The auditory effect remains present, but is attenuated, when responses semantically related to members of the stimulus set of words are used (Experiment 4), and the presence of a "pseudo-Stroop effect" also suggests a possible parallel with the visual Stroop literature. It may be possible to exploit these similarities between the auditory and visual paradigms to investigate differences in processing strategies due to visual vs. auditory presentation of stimulus material. For example, Martin (1978), using visual Stroop experiments, has provided evidence that interference from the word aspect of visual Stroop stimuli results from phonological as well as graphemic encoding. In the auditory paradigm, in which a phonological representation of the stimulus is immediately available, it is possible that there exists no interference due to graphemic encoding. If so, then separation of the effects of phonological plus graphemic encoding from those of phonological encoding alone may, in principle, be possible.

One point of difference between the visual and auditory paradigms was the presence of a large "reverse Stroop" effect in the auditory modality (Experiment 3). Consideration of this difference may aid an understanding of the nature of auditory Stroop interference or of Stroop-like effects (visual and auditory) generally. Nothing like the dominance of reading over naming in the visual paradigm (Fraisse, 1969) appears to be operating in the au-

ditory case. Similarly, it is difficult, in view of the symmetry of the auditory Stroop and reverse auditory Stroop results, to analyze the auditory phenomenon in terms of the interference of the more S-R compatible aspects of a task with those that are less S-R compatible, an interpretation that Beller (1975) has advanced for the visual case.

The concept of dominance does, however, provide a starting point for accounting for the pattern of findings regarding Stroop and reverse-Stroop effects. It is assumed that the secondary, irrelevant activity exerts its interference effect for an interval centered toward its own completion. If the secondary activity is slow enough, then the critical interval will be sufficiently late for no interference with the primary activity to occur because the primary activity will itself have been completed. If, however, the secondary activity is fast enough for the critical interval to be entirely contained within the time taken for the primary activity to reach completion, then maximum interference will occur. This at once explains why reverse Stroop effects are hard to obtain in the color-word task (precisely because there is a speed advantage of reading over naming) and why there is parity between Stroop and reverse Stroop effects in the auditory gender-word task (because maximum interference arises through the roughly equal durations of primary and secondary activities.)

Another view of visual Stroop interference that would also be consistent with the auditory case is that in terms of the integrality of certain stimulus dimensions; integral dimensions of a stimulus cannot be processed independently of one another, and this fact gives rise to a number of experimental effects (e.g., Garner, 1974, 1976, 1978; Lockhead, 1966). One such effect is the interference caused by redundant, random variation of one of a pair of integral dimensions on judgments of the other. Since such dimensions can exist in an asymmetrically integral relation to one another, redundant, random variation in dimension a may interfere with judgments of dimension b, while the reverse pattern (interference of b with judgments of a) will not occur. Typical visual Stroop effects resemble this pattern. Auditory Stroop phenomena, on the other hand, resemble those of true stimulus integrality—redundant, random variation in either dimension causing interference with judgments of the other. The potential relevance of stimulus integrality to a variety of visual interference effects has already been suggested by Clarke and Brownell (1976), and its relevance to the Stroop color-word phenomenon, in particular, has been suggested by Posner and Snyder (1975). The validity of using the concept of stimulus integrality in this context has, however, been called into question by Garner (1976) himself,

who argued that Stroop effects were a special case of interference, distinct from the effects discussed by him and his colleagues. Similarly, Seymour (1979) has claimed that effects due to stimulus integrality result from pictorial encoding, whereas Stroop interference is a product of the semantic code and therefore qualitatively distinct. It is not, however, impossible that the effects of both semantic and pictorial encoding may coexist in a single task, and there therefore seems no a priori reason why stimulus integrality effects should not operate in the Stroop and auditory Stroop phenomena.

In summary, the evidence for a correspondence between the visual and auditory Stroop effects advanced in this paper suggests that the search for a common causal mechanism would be justified. A reconsideration of the potential relevance of the concept of stimulus integrality to Strooplike phenomena may, therefore, be in order. This is not to argue that such an account, with its emphasis on the characteristics of a given stimulus rather than on the capacity limitations of the perceiver, should be used to the exclusion of all others. Such an interpretation may, however, be one of a number of fruitful ones in a consideration of Stroop and auditory Stroop research.

REFERENCES

BELLER, H. K. Naming, reading and executing directions. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, 1, 154-160.
CLARKE, H. H., & BROWNELL, H. H. Position, direction and their perceptual integrality. *Perception & Psychophysics*, 1976, 19, 328-334.
COHEN, G., & MARTIN, M. Hemisphere differences in an auditory Stroop test. *Perception & Psychophysics*, 1975, 17, 79-83.
DALRYMPLE-ALFORD, E. C. Sound similarity and color-word interference in the Stroop test. *Psychonomic Science*, 1972, 28, 209-210.
DYER, F. N. The Stroop phenomenon and its use in the study of perceptual, cognitive and response processes. *Memory & Cognition*, 1973, 1, 106-120.
FRAISSE, P. Why is naming longer than reading? *Acta Psychologica*, 1969, 30, 96-103.
GARNER, W. R. *The processing of information and structure*. Potomac, Md: Erlbaum, 1974.
GARNER, W. R. Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, 1976, 8, 93-123.
GARNER, W. R. Selective attention to attributes and to stimuli. *Journal of Experimental Psychology: General*, 1978, 107, 287-308.

GOLDEN, C. J. Effect of differing numbers of colors on the Stroop color and word test. *Perceptual and Motor Skills*, 1974, 39, 550.
GUMENIK, W. E., & GLASS, R. Effects of reducing the readability of the words in the Stroop color-word test. *Psychonomic Science*, 1970, 20, 247-248.
HAMERS, J. F., & LAMBERT, W. E. Bilingual interdependencies in auditory perception. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 303-310.
HARRISON, N. S., & BOESE, E. The locus of semantic interference in the Stroop color-naming task. *Perception & Psychophysics*, 1976, 20, 408-412.
JENSEN, A. R., & ROHWER, W. D. The Stroop color-word test: A review. *Acta Psychologica*, 1966, 25, 36-93.
LOCKHEAD, G. R. Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, 1966, 72, 95-104.
MARTIN, M. Speech recoding in silent reading. *Memory & Cognition*, 1978, 6, 108-114.
MORTON, J. A preliminary functional model for language behavior. *International Audiology*, 1964, 3, 216-225.
MORTON, J. Categories of interference: Verbal mediation and conflict in card sorting. *British Journal of Psychology*, 1969, 60, 329-346.
MORTON, J. Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language* (Vol. 1) *Proceedings of the conference held at the Institute for Perception Research IPO Eindhoven, The Netherlands, September 1977*. New York: Plenum Press, 1979.
MURRAY, D. J., MASTRONADI, J., & DUNCAN, S. Selective attention to "physical" versus "verbal" aspects of colored words. *Psychonomic Science*, 1972, 26, 305-307.
POSNER, M., & SNYDER, C. R. R. Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, N.J: Erlbaum, 1975.
POULTON, C. C. Unwanted range effects from using within-subject experimental designs. *Psychological Bulletin*, 1973, 80, 113-121.
SEYMOUR, P. H. K. *Human visual cognition*. Middlesex: Collier Macmillan, 1979.
SIMON, J. R., & SUDALAMAITHU, P. Effects of S-R mapping and response modality on performance in a Stroop task. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, 5, 176-187.
SMITH, G. J. W., & NYMAN, G. E. The serial color-word test: A summary of results. *Psychological Research Bulletin*, Lund University, Sweden, 1962, 2, No. 6.
STIRLING, N. Stroop interference: An input and output phenomenon. *Quarterly Journal of Experimental Psychology*, 1979, 31, 121-132.
STROOP, J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 1935, 18, 43-62.
WARREN, R. E. Stimulus encoding and memory. *Journal of Experimental Psychology*, 1972, 94, 90-100.