Identification and discrimination of sweep formants

M. E. H. SCHOUTEN
University of Utrecht, Utrecht, The Netherlands

and

L. C. W. POLS
University of Amsterdam, Amsterdam, The Netherlands

Earlier identification experiments with sweep tones are repeated with rising and falling single formant (band) sweeps, with durations ranging from 15 to 40 msec and sweep rates from 0 to 40 oct/sec. Steady-state portions of 100-msec duration are then added to the sweeps. The general conclusions are that the tendency to perceive level and slightly rising tones as falling, which was such a prominent feature of the earlier results, disappears as the stimuli become more complex, and that sweep discrimination seems to be a function of the difference between the initial and the final frequency of a sweep.

The experiments to be reported on below form a part of a series of experiments designed to elucidate the role of rapid formant transitions in the perception of plosive and other consonants. It is assumed (see, e.g., Delattre, Liberman, & Cooper, 1955; Pols & Schouten, 1985) that the perceived identity of a plosive consonant is determined partly by the noise burst caused by the sudden release of the articulators, and partly by the rapid formant transitions caused by the rapid movements of the articulators from the consonant to the vowel. It is the perception of these latter cues that we want to investigate. We want to build a picture of plosive consonant perception by beginning with the most elementary components and gradually combining them into increasingly complex combinations. In other words, we want to move gradually from pure-tone psychoacoustics to speech perception.

The first step was taken by Schouten (1985), who used pure-tone sweeps with speech-like parameters: durations from 20 to 50 msec, sweep rates between 5 and 60 oct/sec, and center frequencies of 400, 1300, and 2700 Hz, reached at the time centers of the exponentially rising and falling sweep tones. Subjects had to identify sweep direction (rising or falling) and discriminate between sweep tones and steady tones at the center frequencies (400, 1300, and 2700 Hz) of the sweep tones. In addition, they were also asked to discriminate between rising and falling sweep tones that had the same sweep rate (5, 10, 20, 40, or 60 oct/sec). The initial expectation was that there would be two thresholds for sweep detection, positioned symmetrically around the zero octaves/second sweep rate of "level" tones: a threshold for rising and one for falling sweeps, resulting in a three-way division of the sweep

continuum. This expectation was based on the argument in Schouten (1980) that consonant perception was probably based on threshold detection of sweep direction.

No three-way division was found, however. It turned out that, on the whole, level tones and even lightly rising sweeps tended to be heard as falling (see also Pols & Schouten, 1987; Schouten, 1986); some subjects could not even distinguish between falling and rising tones and simply responded "up" to fast sweeps, and "down" to slow or zero-rate sweeps, irrespective of the actual direction. The overall tendency was in agreement with the findings of Pisoni (1976), who, using isolated formant transitions that had to be labeled as speech sounds, found two categories: clearly rising transitions and all other transitions, including steady formants and slightly rising transitions. It was also in agreement with asymmetries in the perceptual effects of rising and falling sweeps found by Gardner and Wilson (1979), Nabelek (1978), and Smoorenburg and Coninx (1980). In addition, the results seemed to agree with findings on the cochlear microphonics of frequency-varying signals by Shore and Cullen (1984).

The obvious next step, taken in the present study, is to repeat our experiments with formant-like stimuli. The results reported by Pisoni (1976), mentioned in the last paragraph, suggest that we may expect level and slightly rising formants to be perceived as falling; however, we do not yet want our subjects to label the stimuli as if they were speech sounds. For the time being, we will maintain the "up" and "down" categories, in order not to omit any steps in our move from psychophysics to speech perception.

The experiments in this paper, then, consist of identification and discrimination experiments, using stimulus parameters and paradigms quite similar to those described by Schouten (1985), except that now the stimuli consist of the output of a filter being swept through the harmonics

of a fundamental frequency. For reasons of economy, only a center frequency of 1300 Hz is used. The predictions, based on the results of the previous experiments, are:

1. There will be a fairly strong tendency to label level formants as "falling"; this effect will extend to slowly rising sweeps, such as those at 5 or even 10 oct/sec.

2. Discrimination between rising and level stimuli will be easier than discrimination between falling and level stimuli, since level stimuli tend to be perceived as falling.

Experiment 1 represents a replication, with slightly altered parameters, of the earlier sweep tone experiment with bandfilter sweeps. In Experiment 2, the same identification and discrimination tasks were carried out with a constant tone added to the band sweeps, joined to either the onsets or the offsets of the sweeps. Special measures needed to be taken in order to prevent the band sweep discrimination experiments from deteriorating into frequency or intensity discrimination experiments.

EXPERIMENT 1
BAND SWEEPS AROUND 1300 HZ

Method
Stimuli. The stimuli were very similar to the ones used in the previous experiment on sweep tones (Schouten, 1985), although in light of the earlier results, the parameter values were slightly different. There were two directions (rising and falling), six sweep rates (0, 5, 10, 20, 30, and 40 oct/sec), and six sweep durations (15, 20, 25, 30, 35, and 40 msec). The stimuli were made by sweeping a digital filter with a fairly arbitrary bandwidth of 200 Hz through the harmonics of a 200-Hz pulse train; this was done exponentially—that is, linearly, in terms of octaves per second—in such a way that the center frequency of the band reached the geometric center frequency of 1300 Hz exactly halfway along the time course of a stimulus. As a result, the highest and lowest frequencies of the filter were 0.8 oct above and below 1300 Hz in a stimulus of 40 msec and 40 oct/sec. The amplitude of the harmonics decreased by 6 dB/oct.

Sampling frequency was 12.5 kHz, and the onsets and offsets were smoothed by means of a cosine window of 3 msec. With a fundamental frequency of 200 Hz, this meant that in a 15-msec stimulus less than two full periods of the waveform were faithfully reproduced. The filter parameters were updated at 1-msec intervals. The stimuli were low-pass filtered at 6 kHz and recorded on professional audio tape.

Identification Procedure. There were five different random orders of six identification blocks; duration was constant in each block, but all six sweep rates occurred four times, twice up and twice down, resulting in 24 randomly ordered stimuli per block. Thus, each subject responded 10 times (5 blocks × 2 occurrences) to each of the 2 (directions) × 6 (durations) × 6 (sweep rates) = 72 different stimuli.

A stimulus was presented every 2 sec at a comfortable listening level, over headphones in two sound-treated booths; 4 subjects were run simultaneously, 2 being seated comfortably in each booth. The response sheets contained separate columns for each block of 24 stimuli. The beginning of a new column was signaled by means of a loud 200-msec tone. A response was given by ticking one of two boxes, the box on the left containing an upward-moving diagonal line, and the one on the right a downward-moving one. A session took 45 min, including a brief training block containing 24 stimuli of 30-msec duration and brief pauses between blocks. Training was deliberately kept short, and no feedback was given, since

we wanted to know how our subjects would respond spontaneously, rather than what performance they would be capable of after training.

Discrimination Procedure. The stimuli were identical to the ones used in the identification experiment; only the presentation differed. We used a four-interval forced choice paradigm, in which each trial consisted of two pairs of stimuli, three of these stimuli being the same and one being different from the other three. The subjects had to indicate which of the two pairs contained the different stimulus. There were (6 durations × 3 conditions) 18 blocks of 40 trials each; within a block, all stimuli of all trials had the same duration. There were three conditions: (1) rising versus steady, (2) falling versus steady, and (3) falling versus rising. Within a 40-trial block, only one condition occurred. In Condition 3 (falling vs. rising), only equal sweep rates were compared. This resulted in 5 different trials per block, 1 for each of the five nonzero sweep rates; within a block, there were eight replications, which differed only in the position of the odd one out among the two stimuli in a trial: each of the two different stimuli was the odd one out half the time, and each occurred in all four positions in a trial.

The 18 blocks of trials were presented in one 50-min session, with short breaks between blocks and a longer one after the first 12 blocks. The subjects were asked to indicate which of the two pairs of stimuli contained the odd one out; they did this by circling a "1" or a "2" on a response sheet. There was an interval of 250 msec within each pair of stimuli, one of 500 msec between the two pairs on a trial, and one of 2 sec between two trials. The start of a new block (and of a new column on the response sheet) was indicated by means of four 100-msec steady tones.

Subjects. There were 15 subjects in both the identification and the discrimination experiments; 11 of these participated in both experiments. They received payment and were promised an extra bonus of 25 guilders if they came out on top in either experiment.

Results and Discussion
Identification. The results of the identification experiment are shown in Figure 1, for (from top to bottom) all 15 subjects, the 4 best subjects, and the 3 poorest subjects (selected on the basis of the total number of correctly identified rising and falling sweeps). We regard this sort of presentation as more informative than just averages and standard deviations over all subjects.

The decision as to whether a quartile was to contain 3 or 4 subjects was made on the basis of a clustering criterion: if at either end of the subject continuum 3 subjects clearly formed a separate group, this group was treated as a quartile; in all other cases, a quartile consisted of 4 subjects.

The representation in Figure 1 has been made in terms of the percentage of "down" responses. This was done for reasons of clarity, and it means that only the left half of each panel directly shows correct scores. A four-way analysis of variance (subjects × direction × duration × sweep rate, with only subjects treated as a random factor) was performed on the *correct score percentages*. Separate analyses were performed on the data appearing in each of the three panels of Figure 1.

As the bottom panel of Figure 1 clearly shows, none of the three nonrandom factors had any significant effect on the scores of the subjects of the lower quartile; only the factor *subjects* produced a significant effect ($p < .01$).

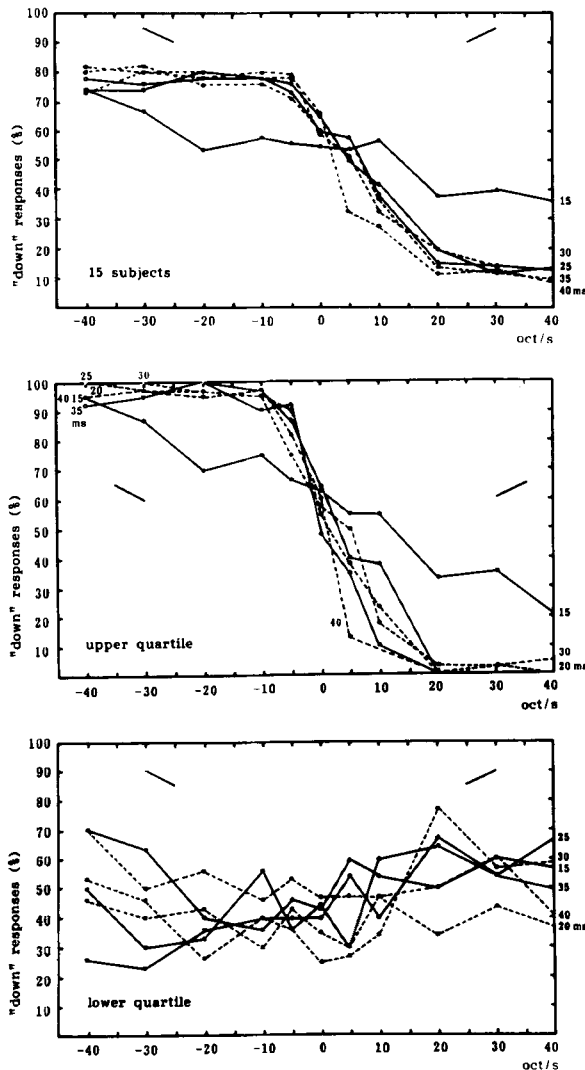In none of the cases did direction have any effect on the correct score percentages. Duration and sweep rate,

Figure 1. Average identification percentages of "down" responses in Experiment 1. Values along the abscissa indicate rates of falling (negative) and rising (positive) frequency in octaves/second. The parameter within the panels is duration in milliseconds. Stimulus direction is indicated in each half of a panel by means of a rising and falling line. From top to bottom, the panels represent the responses by all 15 subjects, the upper quartile, and the lower quartile.

on the other hand, did have highly significant effects (p < .0001) among the 15 subjects (top panel) and the subjects in the upper quartile (middle panel). With the difference between the quartiles as great as it is here, it is not surprising to find that nearly 35% of the variance in the data of all 15 subjects is due to the factor *subjects*.

Let us now look at the lower quartile (Figure 1, bottom). Their responses appear to be completely random. This had not been the case with the lower quartile of subjects identifying sweep tones in Schouten (1985): they had reported perceiving fast sweeps as rising, and slow and zero sweeps as falling. The lower quartile in the present experiment, however, seem not to have heard any cues

that could be interpreted in a consistent manner; perhaps their perception was completely dominated by the unchanging fundamental.

The other two panels have a great deal in common. It is clear, for example, that there is a sudden jump in identifiability between 15 and 20 msec, and that durations greater than 20 msec have very little additional effect. If we disregard the otherwise rather uninformative 15-msec lines, we see that in both panels the curves are not symmetrical around 0 oct/sec, but around a positive sweep rate value. Put differently, the response appears to be asymptotic between −40 and −5 oct/sec and again between +20 and +40 oct/sec. This means that there is a tendency to perceive zero and slowly rising sweeps as falling, although this tendency is not nearly as pronounced as it was in the tonal sweeps in the earlier experiment. In addition, the asymptotes for the 15 subjects (top panel) are not symmetrical around 50%; they are just below 80% correct for falling sweeps and just below 90% correct for rising ones. The reason for this vertical shift is not clear.

The predicted tendency for steady and slowly rising single formants to be perceived as falling is present, but it is considerably smaller than in the case of steady and slowly rising pure tones. At present, there is no clear explanation for this difference. The tentative physiological explanation put forward in a footnote in the previous paper and based on Shore and Cullen (1984) was that at any point in time the excitation pattern produced by a falling tone resembles that of a level tone in that there is always a local maximum, whereas rising tones produce an excitation pattern that is much more spread out. If this explanation is valid, it could be extended to the present formant sweeps: the excitation produced by a falling *formant* will be wider than that of a falling *tone*, so the difference between falling and level formants on the one hand and rising formants on the other should be much less clear than the same difference between pure tones. This would lead us to expect that, in the discrimination experiment to be described below, the relative advantage of a condition in which relatively dissimilar level and rising stimuli have to be compared (over a condition with relatively similar level and falling stimuli) will be smaller than it was with sweep tones.

Discrimination. The presentation of results again includes quartiles, this time based on the total number of correct discriminations: the lower quartile (bottom panel of Figure 2) consisted of 3 subjects (the same 3 as in Experiment 1), the upper quartile (middle panel) of 4 subjects. Each panel consists of three drawings, one for each of the three conditions, which are indicated by combinations of small rising, falling, or level lines below the discrimination curves for the various sweep durations. The random scores produced by the lower quartile in the identification experiment are not repeated in the discrimination experiment, although they were the same 3 subjects. Evidently, it is easier to discriminate formant sweeps from each other and from level formants than it is to identify the direction of the sweeps.
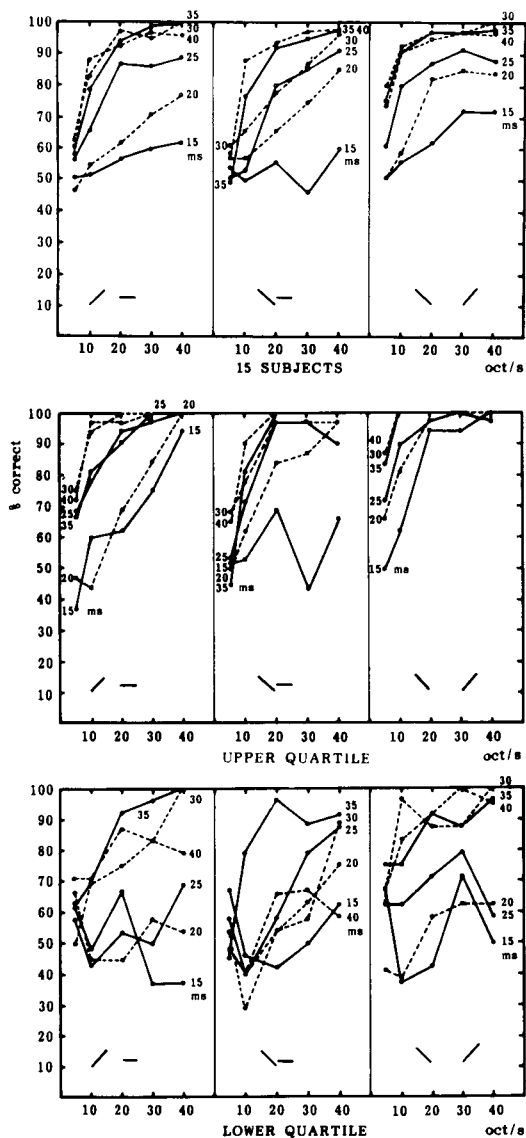
Figure 2. Correct discrimination scores in Experiment 1, with (from top to bottom) the scores for all 15 subjects, for the upper quartile, and for the lower quartile. The three conditions are (from left to right): (1) rising versus level, (2) falling versus level, and (3) falling versus rising.

A five-way analysis of variance—subjects × conditions (rising vs. steady, falling vs. steady, falling vs. rising) × position of the odd one out × duration × sweep rate, with only subjects treated as a random factor—revealed that position of the odd one out had a marginal effect ($p < .05$) on the discrimination scores of all 15 subjects and of the upper quartile, and that the factor *subjects* had a significant effect on the 15 subjects ($p < .0001$) and a marginal effect ($p < .05$) on the lower quartile. The effect of duration was significant in all cases, but sweep rate reached the 1% level only in the case of the 15 subjects and the upper quartile.

The most important factor (in view of our predictions), conditions, did not have any effect on the lower quartile, but was significant with the 15 subjects ($p < .0001$) and the upper quartile ($p < .01$). Three separate analyses of variance in which the conditions factor had been reduced to two levels (so that the effect was now due to that of the difference between two conditions) revealed the following pattern of significant effects for the conditions factor (15 subjects):

Conditions 1 (/ —) and 3 (\ /):
$$F(1,14) = 15.0, p = .0019.$$

Conditions 1 (/ —) and 2 (\ —):
$$F(1,14) = 18.7, p = .0009.$$

Conditions 2 (\ —) and 3 (\ /):
$$F(1,14) = 42.0, p = .0001.$$

Averaged over all durations and sweep rates, the correct discrimination scores in percentages per condition were: Condition 1 (rising vs. level), 76.3%; Condition 2 (falling vs. level), 71.8%; Condition 3 (falling vs. rising), 83.9%. The upper and lower quartiles had the same order in their percentages, which rose from Condition 2 via Condition 1 to Condition 3.

Let us now turn to the prediction we made in relation to the discrimination experiment. This prediction was that Condition 2 (falling vs. level) would be the most difficult condition, and Condition 3 (falling vs. rising) the easiest. This is clearly borne out by the overall percentages, although the differences between the conditions are not very great. The fact that all the differences among the three conditions are significant does not provide us with such a clear pointer as we got in the previous paper with the sweep tones: there the difference between rising versus level and falling versus rising was not significant, indicating that these two conditions were to some extent equivalent. However, the difference between Conditions 1 and 2 is the smallest and least significant here, whereas the difference between Conditions 2 (falling vs. level) and 3 (falling vs. rising) is by far the greatest and most significant, indicating that level and rising stimuli are very different from each other, relatively speaking. The prediction is confirmed, therefore: Condition 2 (falling vs. level) is more difficult than Condition 1 (rising vs. level). As we saw in the identification experiment, however, the tendency to perceive level and slowly rising stimuli as falling is much less pronounced with formant sweeps than with sweep tones.

EXPERIMENT 2

The next step we took on the long road from pure tones to rapidly varying speech signals was a single-formant approximation of consonant-vowel or vowel-consonant syllables. In this approximation, the (plosive) consonant was represented by the band sweep itself, whereas a constant formant preceding or following the sweep stood for the

vowel. The attractive possibility of having steady states at both ends of a sweep had to be rejected, since it would have been meaningless in the present series of experiments, in which it would have amounted to asking subjects whether the first tone was lower or higher than the second one. In the future, we may decide also to try the route taken by Horst (1982), whose sweeps start and end at the same frequency, turning around halfway; this does correspond to certain vowel–consonant–vowel sequences, although probably not to sequences in which the consonant is a plosive.

The necessity of avoiding a simple frequency discrimination experiment, in which subjects compare steady states rather than sweeps, led to an important change in the stimulus parameters. In all our experiments up to now, the sweeps had been centered on 1300 Hz, with onset and offset frequencies determined by sweep rate and duration. With the addition of a steady tone, this was no longer feasible: since in the discrimination experiments subjects had to compare sweeps of the same duration but of different sweep rates, they would have been able to ignore the sweeps and just compare the steady tones. In order to avoid this, we decided to abandon the time center of 1300 Hz and instead to start or terminate all our sweeps, regardless of rate or duration, at 1300 Hz, so that the steady tone was always at the same frequency, namely 1300 Hz. The sweeps in Experiment 2, therefore, moved up or down from 1300 Hz, or up or down to 1300 Hz.

In Experiment 2A, the subjects had to identify and discriminate separate band sweeps, moving up or down to or from 1300 Hz, *without steady states*. This step was included in order to make possible a direct comparison with Experiment 1. Experiment 2B had the same design as Experiment 2A, but the stimuli now included the 1300-Hz steady states.

Experiment 2A
Band Sweeps to and from 1300 Hz

Method

Stimuli. The stimuli were very similar to those used in Experiment 1. The main difference was that, whereas in Experiment 1 a frequency of 1300 Hz was passed at the halfway point during the time course of the stimuli, this time the stimuli started or ended at 1300 Hz.

There was another difference, however, also born of necessity. In Experiment 1, the usual high-frequency roll-off of −6 dB/oct had been employed, resulting in lower amplitudes for higher frequencies, and amplitude equalization had simply consisted of giving *the most intense period* of each stimulus the same maximum amplitude. Although this would still have been perfectly acceptable in Experiment 2A, we noticed that it provided an extra cue as soon as 1300-Hz steady states were added (as in Experiment 2B below): steady-state amplitude varied systematically as a function of sweep direction, rate, and duration, since it depended on the relative amplitude of the 1300-Hz constant formant, in comparison with the rest of the stimulus. For example, in a sweep *falling* toward 1300 Hz, the latter frequency had the highest amplitude of the signal, but it had the lowest amplitude in a sweep *rising* towards it, so that falling stimuli were always louder than rising ones. It was

therefore decided to provide *every single stimulus* period with the same amplitude, regardless of frequency. We realized that this took us one small step away from speech again, but we felt we had no alternative.

We did conduct an extra identification experiment, however, in order to find out whether this latest change in amplitude made any difference for perception. This extra experiment was identical to the identification experiment of 2A: band sweeps to or from 1300 Hz without steady states. The only difference was that in the extra experiment the high-frequency roll-off was maintained, whereas in Experiment 2A it was discarded by giving every single period the same amplitude. The results turned out to be identical to those of Experiment 2A, which are to be described below, indicating that removal of the amplitude cue does not affect identification of sweep direction.

Procedure. The identification and discrimination procedures were the same as in Experiment 1, with one minor difference: the subjects now had to press appropriately labeled keys on a terminal keyboard.

Subjects. The stimuli moving towards 1300 Hz were identified by 14 subjects and discriminated by 14 other subjects; the stimuli moving away from 1300 Hz were identified by 16 subjects and discriminated by 16 other subjects. The four subexperiments were run separately; the subjects were university students who entered their names for as many experiments as they liked (including the four subexperiments making up Experiment 2B, which is to be described below). The payment system was the same as in Experiment 1.

Results and Discussion

The results are displayed in Figures 3 and 4, in four panels per figure. Identification is shown in Figure 3, discrimination in Figure 4. The two panels on the left in both figures represent the stimuli with the same terminal frequency of 1300 Hz, whereas the panels on the right contain the results for the stimuli with the same initial frequency of 1300 Hz (remember that each panel is the outcome of a separately run subexperiment, involving different subjects in principle, although not necessarily in practice).

The top panels in both Figure 3 and Figure 4 represent all subjects, and the bottom panels represent the longest and shortest durations of the upper (continuous line) and lower (dashed line) quartiles. It was felt that this was sufficient, and that displaying the full quartiles, as was done for Experiment 1, could be dispensed with here.

As in Experiment 1, the identification scores (Figure 3) are expressed in terms of the percentages of "down" responses, whereas the discrimination scores (Figure 4) give correct response percentages.

The identification results (transformed to correct scores and excluding 0 oct/sec) were subjected to a four-way analysis of variance: subjects × direction (up, down) × duration × sweep rate. In both cases (*to* 1300 and *from* 1300 Hz), all main effects were significant at the 1% level, except for that of the factor *direction*.

The two left-hand panels of Figure 3 make it clear that identifying the direction of sweeps moving towards the same terminal frequency is an impossible task, due to cue reversal. The falling sweeps are heard as rising, and the rising sweeps as falling. This happens to such an extent
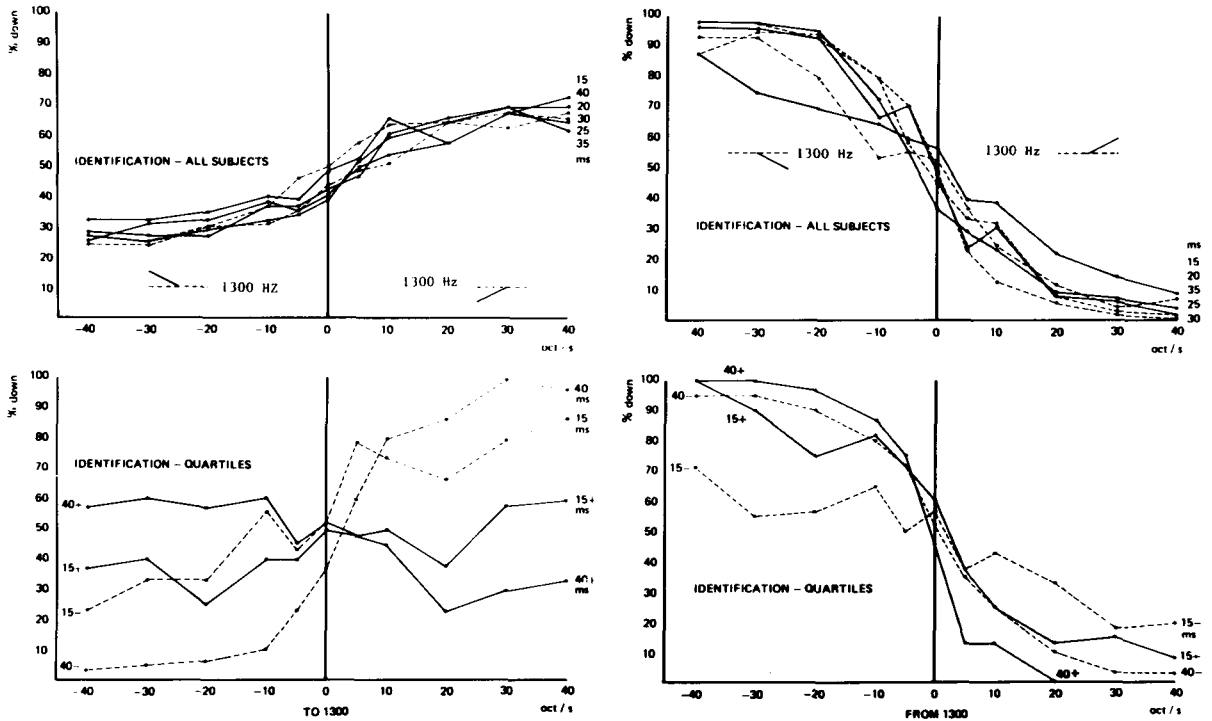
Figure 3. Average identification percentages of "down" responses in Experiment 2A; left: *to* 1300 Hz; right: *from* 1300 Hz. The upper panels show the overall scores, the lower panels show the results for the upper (continuous lines) and lower (dashed lines) quartiles. The horizontal axes represent sweep rate in octaves/second. The parameter within the panels is duration in milliseconds.
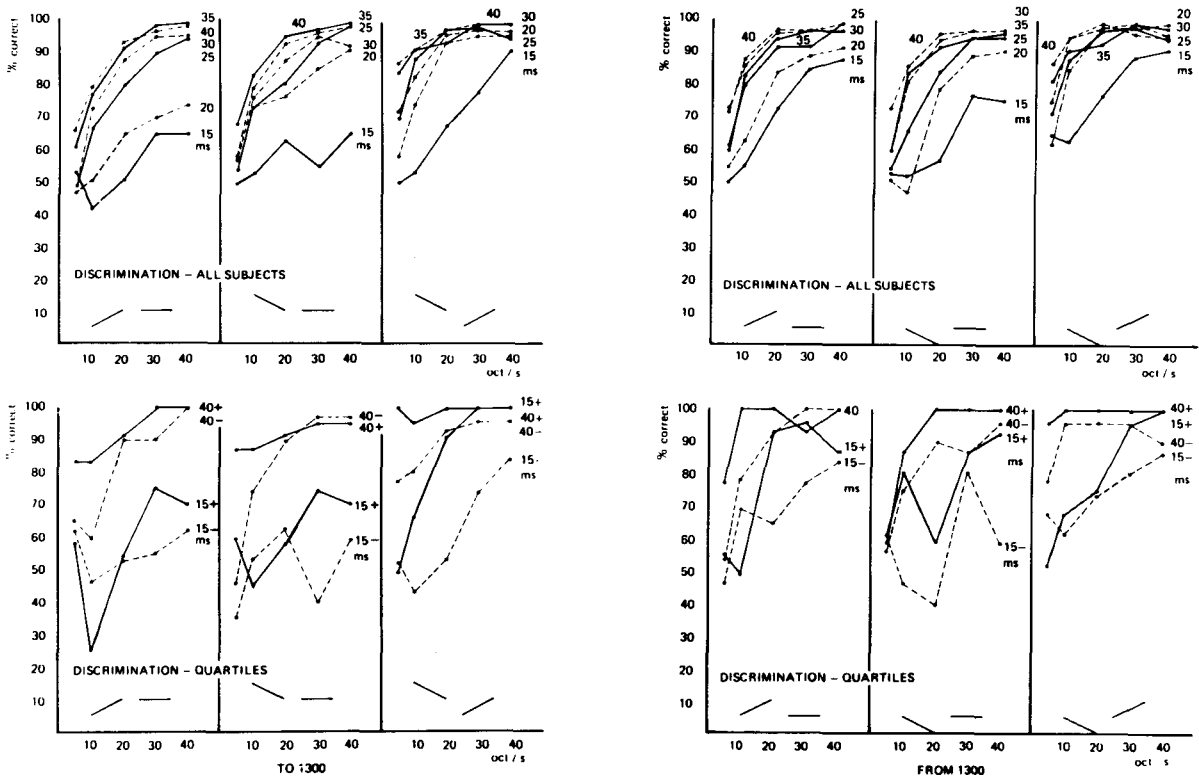

Figure 4. Correct discrimination scores in Experiment 2A. Otherwise same as Figure 3.

that the highest correct score that can be obtained (by the "upper" quartile in the bottom left panel) is not much different from 50%, whereas the "lower" quartile show a complete reversal. The quotation marks are there on purpose: it is clear that the "upper" quartile are just guessing, whereas the "lower" quartile very consistently give different but wrong labels to different stimuli. The cause of this reversal may have to be sought in the findings reported by Brady, House, and Stevens (1961) and by Nabalek, Nabalek, and Hirsh (1970, Experiment 1): under certain conditions (product of duration and frequency range above a certain minimum level), the pitch of a sweep corresponds to a frequency that is close to its terminal frequency. This means that all stimuli finishing at 1300 Hz had a very similar pitch, which in the case of *rising* sweeps was a little lower than 1300 Hz, whereas with *falling* sweeps it was a little higher. Apparently, a high-pitch stimulus is labeled as "rising," and a low-pitch stimulus as "falling." If this is true, it could mean that all our results originate from such a mechanism. So far, all rising stimuli had always ended up higher than all falling stimuli; now that this relationship had been changed, however, the responses had also changed.

The right-hand panels of Figure 3 show the identification results for stimuli moving away from 1300 Hz. The pattern is highly regular (no cue reversal) and shows no evidence at all of the asymmetry between falling and rising sweeps that was such a prominent feature of the Schouten (1985) results.

A five-way analysis of variance was applied to the discrimination data: subjects × conditions (up vs. level, down vs. level, down vs. up) × position (odd one out in first or in second pair of a trial) × duration × sweep rate. All main effects were significant well beyond the 0.1% level; most were even significant (at the 1% level) among the two quartiles. The effect of position meant that more correct responses were given when the deviating stimulus occurred in the *first* pair than when it occurred in the second pair; the difference was always well over 10%. Subjects apparently respond "1" when they are guessing.

Looking now at the discrimination data for both sets of stimuli, we see that discrimination is hardly, if at all, affected by the cue reversal we observed in identification: there are very few differences between the two sets as far as discrimination is concerned, whereas their identification was very different. The only thing that catches the eye is the relatively low score obtained for the short stimuli (15 and 20 msec) in the rising (to 1300 Hz) versus level condition. This lack of difference between the two sets of discrimination results means that a pitch extraction mechanism cannot be the whole story: if it were, the scores in the left-hand panels of Figure 4 would have been much lower than those in the right-hand panels, where the terminal frequencies are different.

The identification data (Figure 3) have already shown us that in the present stimuli there is no asymmetry between the perception of rising and that of falling sweeps; there is, therefore, no reason to regard the level stimuli (0 oct/sec) as more like the falling than like the rising stimuli. In our previous discrimination experiments, this asymmetry had resulted in Condition 3 (falling vs. rising) being closer to Condition 1 (level vs. rising) than to Condition 2 (level vs. falling). There is no evidence of this at all in the present discrimination experiments; in one case, even the reverse seems true.

Experiment 2B
Band Sweeps with Steady States at 1300 Hz

This experiment, in which steady states were added to the band sweeps to or from 1300 Hz, was run concurrently with Experiment 2A.

Method

Stimuli. The same stimuli were used as in Experiment 2A, except that in this case a 100-msec band centered at 1300 Hz was added to the sweep, either preceding or following it.

Procedure. The procedure did not differ from that in Experiment 2A, except that subjects were now instructed to listen for (and identify or discriminate) changes taking place at the beginning or end of a stimulus.

Subjects. Subjects were recruited in the same way as in Experiment 2A. There were 14 subjects identifying, and 15 subjects discriminating stimuli moving *towards* 1300 Hz; in the case of stimuli moving *from* 1300 Hz, there were 14 subjects for identification and 13 for discrimination.

Results and Discussion

The results are presented in Figures 5 and 6 in the same way as in Figures 3 and 4: identification in Figure 5, discrimination in Figure 6; *to* 1300 Hz on the left, *from* 1300 Hz on the right. In the left-hand panels of Figure 5, duration had so little effect on the overall identification results that, for the sake of clarity, only the shortest and the longest duration are shown.

The identification results (Figure 5) were submitted to a four-way analysis of variance: subjects × direction (up or down) × duration × sweep rate. Not unexpectedly, the results for the set of stimuli moving towards 1300 Hz contained only one significant effect: that of subjects. The data in the right-hand panels exhibited three highly significant main effects; again direction turned out to have no effect.

The discrimination data (Figure 6) were submitted to a five-way analysis of variance: subjects × condition (up vs. level, down vs. level, up vs. down) × position (first or second pair) × duration × sweep rate. All main effects were highly significant. As in Experiment 2A, more correct responses were given when the deviating stimulus occurred in the first pair of a trial.

Let us now compare the results for sweeps with and without steady-state extensions, by looking at the corresponding panels in Figures 3 and 5 and Figures 4 and 6.

Identification—to 1300 Hz (Figures 3 and 5, left-hand panels). The addition of a steady state here does not greatly
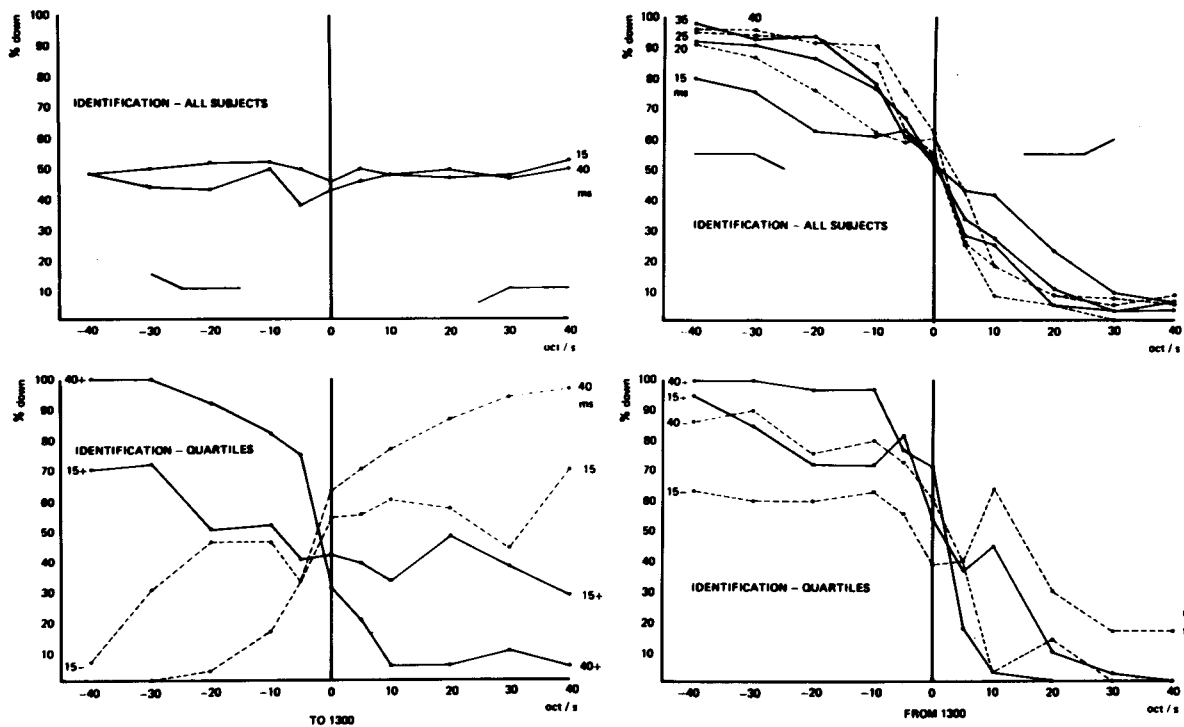
Figure 5. Identification scores from Experiment 2B (sweeps plus steady states). Otherwise same as Figure 3.
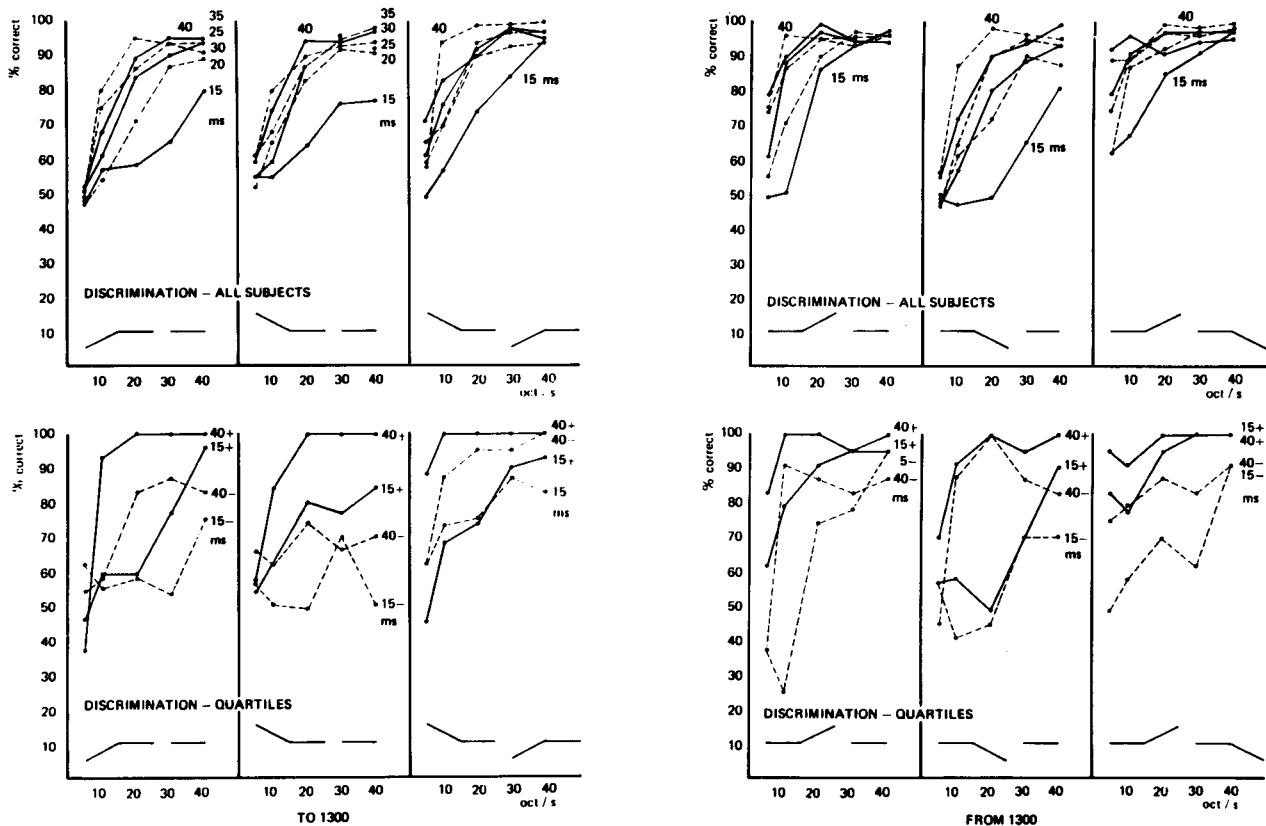

Figure 6. Discrimination scores from Experiment 2B (sweeps plus steady states). Otherwise same as Figure 4.

affect the "lower" quartile, who keep using the "wrong" labels consistently. However, it does seem to have a positive influence on the "upper" quartile, who move from guessing to using the correct labels. The net overall result is an almost straight line, which does not mean much. It should also be noted that the difference may simply be due to the small number of subjects making up a quartile: there may be two listening strategies or response biases, distributed unevenly over the two samples.

Identification—from 1300 Hz (Figures 3 and 5, right-hand panels). We can be brief here: the data in the right-hand panels of Figures 3 and 5 are virtually identical: a steady state has no effect.

Discrimination—to 1300 Hz (Figures 4 and 6, left-hand panels). The differences between sweeps with and without steady states are negligible, apart from an improved performance at the lowest durations (15 and 20 msec) in the up versus level condition.

Discrimination—from 1300 Hz (Figures 4 and 6, right-hand panels). Again, adding a steady state has no effect.

GENERAL DISCUSSION

Let us attempt to summarize our findings so far.

When investigating the perception of pure-tone sweeps (Schouten, 1985), we observed a clear asymmetry between upward- and downward-moving sweeps, which we, very tentatively, attributed to cochlear stimulation patterns. This still awaits investigation: at present, this explanation is no more than just a hypothesis that cannot as yet be tested by calculating the stimulation patterns caused by our stimuli (E. de Boer, personal communication). One of our next steps will be an attempt to construct a model of cochlear mechanics capable of dealing with our dynamic stimuli.

The asymmetry we observed is unlikely to play a role in speech perception, since it disappears as soon as the stimuli become more complex. It is still present when a third response category is introduced (Schouten, 1986), and one can just about discern it still in the band sweeps around 1300 Hz in the present Experiment 1. It has to be admitted, however, that if it had not been for the earlier asymmetrical results, we would not have detected an asymmetry here.

The asymmetry has completely disappeared when the band sweeps start or end at the same frequency. Identification of the direction of sweeps moving towards 1300 Hz proved to be a very difficult task, possibly because the perceived pitches were very similar; this makes it likely that at least some subjects use a strategy of listening for the pitch of a sweep tone or band, rather than for the movement itself; a relatively high pitch is then called "up," and a relatively low pitch "down." This would explain the results shown in the left-hand panels of Figures 3 and 5. Otherwise, the patterns are highly regular, with identification and discrimination as fairly straightforward functions of sweep rate and duration.

Figure 7 is a summary of the discrimination results in terms of 75% correct "thresholds." Each point in this figure represents the minimum sweep rate (in octaves/sec) necessary for discrimination to reach 75% at a particular duration, and was calculated by describing each discrimination curve in Figures 2, 4, and 6, and in Schouten (1985) as a second-order polynomial. In view of the shape of most of these curves, this seemed a reasonable thing to do; in those cases where a curve obviously did not have the required shape, no data point was entered into Figure 7.

The three panels on the right in Figure 7 do not show any experimental data, but rather serve as a reference: for each duration, it is shown how many octaves/second are required to obtain a sweep covering 300 Hz. The top panel contains one line for sweeps around 1300 Hz; the middle and bottom panels contain two lines, one for sweeps below, and one for sweeps above 1300 Hz.

What is surprising is the strong resemblance between the 75% data and the 300-Hz curves. If we disregard the most difficult conditions (those of which it could be said that the durations are too short for human listeners), nearly every curve seems to be characterized more by a constant frequency range than by anything else. This could indi-
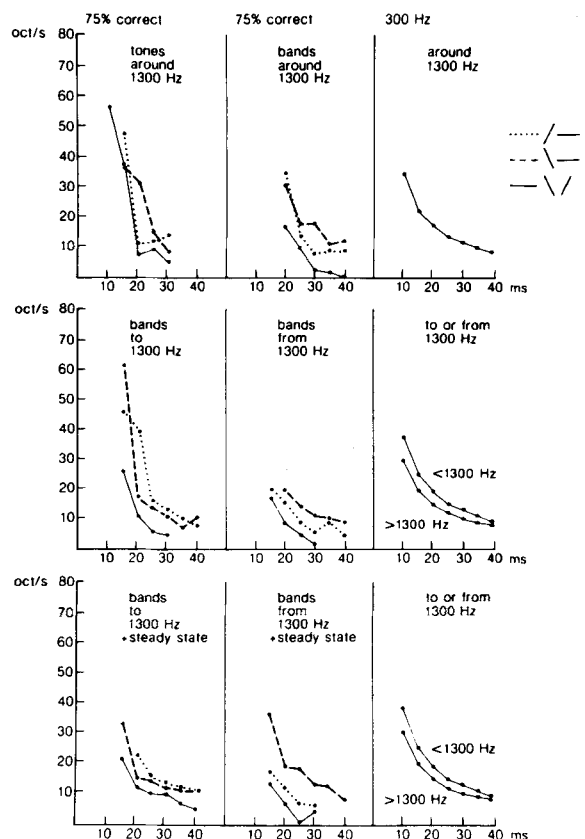

Figure 7. Seventy-five percent correct "thresholds" for all the discrimination data gathered so far (all subjects). Dotted lines: rising versus level; dashed lines: falling versus level; continuous lines: rising versus falling. By way of reference, the panels on the right show combinations of sweep rate and duration leading to a sweep *range* of 300 Hz.

cate that sweep perception is determined mainly by the difference between initial and final frequency, and that sweep rate, duration, and perhaps even sweep form matter much less than one might have thought.

Our next, and probably final, step will be to combine band sweeps into more complex combinations, and to require subjects to use speech labels to identify the resulting stimuli. Our point of departure was that rapid frequency transitions play an important part in the perception of speech, particularly of plosive consonants. Since identification of sweep direction has turned out to be a difficult, sometimes even impossible task, plosive consonant perception must be based at least partly on sweep discrimination, and therefore on the simple difference between the initial and final frequencies of a sweep.

REFERENCES

BRADY, P. T., HOUSE, A. S., & STEVENS, K. N. (1961). Perception of sounds characterized by a rapidly changing resonant frequency. *Journal of the Acoustical Society of America, 33,* 1357-1362.

DELATTRE, P. C., LIBERMAN, A. M., & COOPER, E. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America, 27,* 769-773.

GARDNER, R. B., & WILSON, J. P. (1979). Evidence for direction-specific channels in the processing of frequency modulation. *Journal of the Acoustical Society of America, 66,* 704-709.

HORST, J. W. (1982). *Discrimination of complex signals in hearing.* Unpublished doctoral dissertation, University of Groningen.

NABELEK, I. V. (1978). Temporal summation of constant and gliding tones at masked auditory threshold. *Journal of the Acoustical Society of America, 64,* 751-763.

NABELEK, I. V., NABALEK, A. K., & HIRSH, I. J. (1970). Pitch of tone bursts of changing frequency. *Journal of the Acoustical Society of America, 48,* 536-553.

PISONI, D. B. (1976). *Some effects of discrimination training on the identification and discrimination of rapid spectral changes* (Research on Speech Perception Progress Report No. 3, pp. 121-141). Bloomington, IA: Indiana University.

POLS, L. C. W., & SCHOUTEN, M. E. H. (1985). Perception of plosive consonants in ambiguous sentences. *Journal of the Acoustical Society of America, 78,* 33-39.

POLS, L. C. W., & SCHOUTEN, M. E. H. (1987). Perception of tone, band, and formant sweeps. In M. E. H. Schouten (Ed.), *The psychophysics of speech perception* (NATO ASI Series 39, pp. 231-240). Dordrecht, The Netherlands: Martinus Nijhoff.

SCHOUTEN, M. E. H. (1980). The case against a speech mode of perception. *Acta Psychologica, 44,* 71-98.

SCHOUTEN, M. E. H. (1985). Identification and discrimination of sweep tones. *Perception & Psychophysics, 37,* 369-376.

SCHOUTEN, M. E. H. (1986). Three-way identification of sweep tones. *Perception & Psychophysics, 40,* 359-361.

SHORE, S. E., & CULLEN, I. K. (1984). Cochlear microphonic responses of the peripheral auditory system to frequency-varying signals. *American Journal of Otolaryngology, 5,* 34-42.

SMOORENBURG, G. F., & CONINX, F. (1980). Masking of short probe sounds by tone bursts with a sweeping frequency. *Hearing Research, 3,* 301-316.

Announcement

19th Annual Meeting of the Society for Computers in Psychology
Atlanta, Georgia
November 16, 1989

The 19th Annual Meeting of the Society for Computers in Psychology will be held at the Hyatt Regency Hotel in Atlanta, Georgia on November 16, 1989. As usual, this meeting will feature papers on applications of computers to all areas of psychology—experimental, clinical, and educational.

For further information, contact Paula Goolkasian, Department of Psychology, University of North Carolina at Charlotte, Charlotte, NC 28223.