INTRODUCTION

The subject of this thesis is a newly discovered perceptual phenomenon which has been termed the "lag effect." The lag effect was first observed in a dichotic listening\(^1\) experiment where stop consonant-vowel syllables\(^2\) contrasting in the stop consonant were delivered to opposite ears with slight differences in the time of arrival of the syllables at the right and left ears. When listeners were asked to identify the competing stop consonants, they were more often correct in identifying the delayed syllable than the one which was first to arrive (Studdert-Kennedy, Shankweiler, and Schulman, 1970; Lowe, Cullen, Thompson, Berlin, Kirkpatrick, and Ryan, 1970; Porter, Shankweiler, and Liberman, 1969). The term "lag effect" refers to this advantage in recall accruing to the syllable at the delayed ear. This advantage is observed regardless of whether the delayed ear is the right ear or the left ear.

We still do not know what experimental conditions are required in order to elicit the lag effect. It has been established, however, that two sorts of competition are needed. There must be competition between stimuli and also competition between ears. When the temporally overlapped CV syllables were made to compete at the same ear, the result was the opposite of that observed with dichotic presentation. With monotic presentation\(^3\) leading stop consonants were more accurately identified than lagging stop consonants (Studdert-Kennedy et al., 1970; Lowe et al., 1970).

Of the two effects—the dichotic lag effect and the monotic lead effect—the dichotic effect excited greater interest because it clearly had its origin in competition arising within the central nervous system. The monotic lead effect, on the other hand, could plausibly be attributed to masking at the peripheral receptor.

In order for us to make a more precise statement about the locus of the lag effect, more information is required. The general goal of the research described in this thesis was to provide data which would aid in pinpointing more exactly the stage in the processing of the stimuli at which the lag effect originates.

It is possible that the lag effect is a phenomenon of auditory perception. If this is so, then it should be possible to obtain the lag effect with a fairly wide assortment of auditory stimuli—nonspeech stimuli as well as speech stimuli. However, Studdert-Kennedy et al. (1970) and Porter et al.

---

\(^1\)Dichotic presentation is an experimental method in which different auditory stimuli are delivered simultaneously to opposite ears.

\(^2\)The stimuli were the syllables [pa], [ta], [ka], [ba], [da], [ga].

\(^3\)Monotic presentation is the presentation of different stimuli simultaneously at the same ear.
(1969) considered the lag effect to be a more circumscribed phenomenon. They proposed that the lag effect occurs only for speech sounds and that the effect reflects some property of speech processing mechanisms. If this idea is correct, then an understanding of why the lag effect occurs could greatly increase our knowledge about the nature of the processes involved in speech perception.

The notion that the lag effect is a speech perception phenomenon seems reasonable from a theoretical point of view. There is considerable evidence that there exist distinct perceptual systems for processing speech and non-speech stimuli (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967). Research on speech perception has demonstrated that the relationship between the acoustic speech signal and the perceived phonological structure of the message is exceedingly complex, much more complex than the relation between stimulus and percept in the case of nonspeech auditory stimuli. One complexity peculiar to speech perception has to do with the problem of segmentation of the acoustic signal. The phonetic message is perceived as a string of segments; for example, the word "mat" is perceived as having three segments or phones, [m], [æ], and [t]. However, these segments are not physically marked in the acoustic signal. Rather, the signal is essentially continuous, and the minimal units into which the signal might be segmented on some purely physical basis are more on the order of syllables than individual phones. In addition to the lack of segmentation, another complexity in the relation between the acoustic speech signal and percept is the variability in the acoustic representation of a particular speech sound depending on the surrounding phones. For example, in the word "mat," the features of the acoustic signal which distinguish [m] from [n] or [t] from [k] vary greatly depending on whether the vowel is [æ] or some other vowel, and in rapid speech the acoustic cues for the vowel are themselves altered by the identity of the initial and final consonants. Thus, the entire syllable is recoded into a single acoustic unit, and the task of the listener is one of decoding the signal in order to extract the string of phonetic segments (Liberman, 1971). This parallel transmission of sequential information appears to account for the unique efficiency of speech as a vehicle for transmitting language (Liberman, Cooper, Studdert-Kennedy, Shankweiler, and Harris, 1966), but special apparatus would seem necessary for the recognition of speech sounds.

One of the most convincing sources of evidence of the existence of distinct speech and nonspeech modes of perception is research indicating that speech and nonspeech sounds are analyzed in opposite brain hemispheres. It has long been known that in man the left cerebral hemisphere is more important than the right cerebral hemisphere for the perception and production of language. This conclusion was based originally on studies of the effects of unilateral brain injury and on diagnostic procedures like the paralysis of one hemisphere by the injection of intracarotid sodium amytal (Wada and Rasmussen, 1960) and the electrical stimulation of the cortex (Penfield and Roberts, 1959). More recently an experimental technique has become available for investigating lateralization of function in normal people. Use of this technique has led to the conclusion that the analysis of speech sounds takes place predominantly in the left hemisphere along with other linguistic functions, while analysis of nonspeech auditory stimuli takes place predominantly in the right hemisphere.

These conclusions are based on the results of certain dichotic listening experiments. Kimura (1961a,b) discovered that when digits are presented in
competition at opposite ears, the digits presented at the right ear are more accurately identified than those simultaneously presented at the left ear. A right-ear superiority has also been demonstrated for dichotically presented meaningful words other than digits (Dirks, 1964; Curry and Rutherford, 1967; Borkowski, Spreen, and Stutz, 1965) and for nonsense words (Kimura, 1967; Curry and Rutherford, 1967). It is now generally accepted that the right-ear advantage in report of dichotically presented verbal material is a consequence of the left-hemispheric specialization for language. Kimura (1961b, 1967) attributed the ear asymmetry to the fact that the contralateral pathway from ear to cortical processing areas is stronger than the ipsilateral path. The superiority of the crossed path would give stimuli at the right ear a perceptual advantage over stimuli at the left ear when the stimuli are competing for the linguistic processing areas in the left hemisphere.

This interpretation of the ear effect, which links the right-ear advantage to left-hemispheric language "dominance," received strong support from the finding that people with reversed (i.e., right-hemisphere) language representation as assessed independently by the sodium amytal test also had a reversed (i.e., left-ear) advantage on the dichotic digits task (Kimura, 1961a). Further support came from the finding that most people show a left-ear advantage when the stimuli are dichotically presented nonspeech sounds like melodies, sonar signals, or environmental sounds (Kimura, 1964; Curry, 1967; Chaney and Webster, 1966; Spreen, Spellacy, and Reid, 1970; Darwin, 1969). The left-ear effect for dichotically presented nonspeech sounds implies that the right hemisphere is of greater importance than the left for nonspeech auditory perception. The results of research on effects of right and left temporal-lobe damage also support the view that the right hemisphere is specialized for processing nonspeech material (Milner, 1962; Shankweiler, 1966).

The fact that the direction of ear asymmetry is different for speech and nonspeech auditory stimuli does not necessarily mean that the identification of speech sounds takes place in the left hemisphere. Left-hemispheric specialization may pertain only to the higher (semantic and syntactic) levels of language. The actual analysis of the acoustic signal and the identification of the sounds might take place predominantly in the right hemisphere along with other auditory analysis. It was soon discovered that neither grammatical structure nor meaningfulness was a necessary condition for obtaining the right-ear effect. The right-ear advantage could be reliably obtained when the dichotically presented stimuli were CV or CVC syllables contrasting between ears in the initial or final consonant. Report from the right ear is more accurate than report from the left ear for competing stop consonants (Shankweiler and Studdert-Kennedy, 1966, 1967; Studdert-Kennedy and Shankweiler, 1970), fricatives (Darwin, 1971b), or liquids (Haggard, 1969). The finding of a right-ear superiority when single phonetic segments differ between ears and a left-ear superiority when nonspeech stimuli differ between ears is strong evidence in support of the distinction between speech and nonspeech perceptual modes. This result also ties the speech perception mode closely to the remainder of the linguistic system.

If there are distinct speech and nonspeech modes of perception, then it should be possible to discover perceptual phenomena which occur for speech stimuli but not for nonspeech stimuli. The dichotic right-ear effect is one such phenomenon. Another phenomenon which apparently is found only with speech
sounds is the phenomenon of categorical perception. Categorical perception is demonstrated in an experiment where listeners are required to discriminate small variations along some acoustic dimension which serves as an acoustic cue for phonetic classification of the stimuli. If the stimuli are encoded speech sounds like stop consonants, listeners are unable to make auditory judgments about stimuli which are acoustically different but fall within the same phonetic class. That is, listeners perceive the phonetic category and fail to discriminate the within category acoustic variations. If the same acoustic dimension is removed from the speech context, these auditory distinctions are readily made (Eimas, 1963; Liberman, Harris, Kinney, and Lane, 1961; Mattingly, Liberman, Syrdal, and Halwes, 1971).

Both categorical perception and a dichotic right-ear advantage are reliably obtained when the stimuli are the stop consonants. These same stimuli were the ones used in the experiment in which the lag effect was discovered. In fact, when syllables contrasting in the stop consonant are presented dichotically with slight differences in onset time, both a right-ear advantage and a lag effect are observed. They appear to be separate phenomena superimposed on each other. Since it is known that the right-ear effect and categorical perception are specifically speech perception phenomena, it would seem reasonable to suppose that the lag effect might be a speech perception phenomenon as well.

In addition to the theoretical considerations which might lead one to suppose that the lag effect is a speech perception phenomenon, there is also some empirical evidence relevant to this hypothesis. In order to test directly the notion that the lag effect is a speech perception phenomenon, one would like to compare perception of dichotically presented time-staggered speech and nonspeech stimuli. This experiment has not yet been performed. However, Massaro (1970) performed a similar experiment with tones and reports a phenomenon similar to the lag effect. Massaro (1970) delivered a 20-msec test tone to one ear and a longer masking tone to the other ear. When the onset of the masking tone followed the onset of the test tone by 20 to 350 msec, the masking tone interfered with recognition of the test tone. However, no interference was observed when the masking tone preceded the test tone.

The phenomenon described by Massaro is similar to the lag effect, but it differs from the lag effect in one respect which seems significant. This difference is in the time course of the two phenomena. For the dichotic tones, the interference was greatest with the shortest delays between stimulus onsets and decreased monotonically with longer delays. For the dichotic stop consonants, the lag effect in the data of Studdert-Kennedy et al. (1970) was visible with a 10-msec difference in stimulus onsets, but the interference in report of the leading stop increased with longer delays. The lag effect was greatest with a delay of about 50 msec between syllables and then decreased with still longer delays. The effect was still visible with a delay of 120 msec, but with delays of 180 msec or more, the lag effect is apparently no longer in evidence (Berlin, Loovis, Lowe, Cullen, and Thompson, 1970). In the Massaro (1970) experiment the interference persisted to a certain extent even with delays as long as 350 msec between stimulus onset. Because of the difference in the time course of the interference effects observed with speech and nonspeech, it seems unlikely that the identical processes underlie the two phenomena.
Also, it is evident from other research that the phenomenon described by Massaro (1970) is not in itself a completely general phenomenon of auditory perception of dichotic stimuli. Darwin (1971a) opposed stop consonant-vowel syllables at one ear to bursts of noise of the same duration at the opposite ear. He found that when syllables were opposed to noise, the effect of delay between ears was the opposite of that observed when syllables were competing against other syllables. Syllables were better identified when they preceded the noise onset, although they were better identified when they followed the onset of a competing syllable.

If the lag effect and the "retroactive interference" effect described by Massaro (1970) were both instances of the same phenomenon which occurs when auditory stimuli are presented with temporal delays between ears, then one would expect to observe the same effect when speech stimuli are opposed to noise, since both the syllables and the noise are auditory stimuli. Darwin's (1971a) result would seem to suggest that the lag effect arises not merely out of competition between stimuli at the two ears but that, in addition, the stimuli must be competing for the same processor. Stop consonants presented at both ears give a lag effect. Dichotically presented tones give a similar effect. But stop consonants opposed to noise do not give the effect. Thus, even if there is an effect with nonspeech stimuli which is similar to the lag effect, the fact that such interference is not obtained when one of the stimuli is a syllable and the other a nonspeech sound implies that a distinction between speech and nonspeech is involved in these phenomena.

Certain other experiments would seem to implicate speech processing in the lag effect. These experiments indicate that the lag effect is sensitive to the linguistic composition of the competing syllables. Day (1968, 1969) presented dichotically words differing between ears in the initial consonants. The word at one ear began with a stop consonant while the word at the other ear began with a liquid. The relative onset time of the two words was varied. When asked to report what they heard, about half the subjects reported hearing one or both of the stimulus words on each trial. For these subjects, there was apparently a preference for the lagging stimulus (Day, personal communication). However, the remaining subjects reported hearing a single stimulus which was in fact a fusion of the two stimulus words. For example, when "back" was delivered to one ear and "lack" was delivered to the other ear, these subjects reported hearing "black" regardless of whether the stop led the liquid or the liquid led the stop.

While nearly all subjects have a lag effect with stop-stop competition, at least half the subjects did not show a lag effect with stop-liquid combinations but gave a "fusion" response instead. Stops and liquids are rather similar classes of sounds acoustically, so if the lag effect were thought to arise in the purely auditory analysis of the stimuli, it would be difficult to explain why the effect is so much more frequent with stop-stop than with stop-liquid combinations. The explanation may lie in the linguistic difference between the stop-stop and stop-liquid conditions. Stop-stop sequences never occur in syllable-initial position in English, but initial stop-liquid clusters are common. It is only in the latter condition that fusion responses are possible. It may be that a decision is made whether to combine the dichotic stimuli or to treat them as competing stimuli. Such a decision would have to be made at a rather high level of analysis,
after a partial phonetic analysis has revealed whether the stimuli from the
two ears are capable of being fused. This reasoning would place the origin
of the lag effect at a stage in processing which follows considerable pho-
netic analysis of both syllables.

The lag effect may be one of a class of perceptual interference effects
which occur when stimuli are in competition for the same central processor.
The processors involved may be quite specific—processors of speech sounds,
or perhaps even certain classes of speech sounds, and processors of nonspeech
sounds. When different processors are required as in the case where speech
and noise were presented to opposite ears, there is no evidence of the lag
effect. It should also be pointed out that the nonspeech experiment of
Massaro (1970) did not use precisely the same experimental technique as that
used with the stop consonants, and more research is clearly needed to deter-
mine whether there is a lag effect with competing nonspeech stimuli.

The four experiments described in this thesis are all concerned with
the dichotic lag effect. The major goal was to establish the level of pro-
cessing at which the lag effect originates and, specifically, to evaluate
the notion that the lag effect is related to phonetic recognition processes.
In this connection, the relation between the lag effect and the dichotic
right-ear effect is of particular interest. Since it is known that the
right-ear effect is obtained only for speech stimuli, it would be expected
that some relation between the two effects would be seen if the lag effect
also had its origin in speech decoding processes.

Other questions dealt with in the thesis will be mentioned briefly.
One problem was to determine whether the lag effect depends on any particular
strategy of report. If the lag effect proved to be independent of the method
of recall, this would be strong evidence that the effect is a genuine percep-
tual phenomenon. If, on the other hand, a particular recall procedure were
required, one might conclude that the lag effect is related to the organization
of responses rather than to perception.

Finally, questions were raised about the subjective experience of people
listening to syllables competing at the two ears. Can listeners hear both
stimuli or are they aware of only one? Can they tell which stimulus is ar-
riving at each ear? Can they tell which of the syllables is the first to ar-
rive and which is delayed?