Working Papers of the Cornell Phonetics Laboratory
2000, vol. 13, pp. 34–72

Phonetic priming of features in a shadowing task*

Tobey L. Doeleman

The role of features in speech perception was investigated in a series of priming experiments in which subjects shadowed CV syllables consisting of the English consonants [p, b, t, d, f, v, s, z] followed by the vowel [a]. Natural tokens were used to create sixty-four prime-target pairs which varied critically the number of features the consonants shared (0-3) and the type of featural similarity. Subjects were asked to repeat the target syllables from aurally presented prime-target pairs as fast as possible. Patterns of speech onset latencies and accuracy data were compared for the different featural similarity conditions. The results indicate that there is inhibitory phonetic priming of specific place and manner features but that the influences of the prime and target features are more complex than simple shared activation of a particular feature. Results are also discussed with reference to previous findings from a series of gating experiments (using the same stimulus set manipulation) which found that subjects given place information had significantly higher scores in a consonant identification task than subjects given manner information (Doeleman, 1998).

1 Introduction

In this paper, I present a series of priming experiments investigating subphonemic context effects in a shadowing task. Specifically, these experiments examined whether the number of features and type of featural similarity between consonants in natural speech prime-target CV syllable pairs facilitated or inhibited shadowing of the target syllable. The paper begins with a description of the methodology of priming and how it has been used to elucidate the processes involved in word recognition. Next, I discuss the motivation for the present study and describe the experimental design and procedures. Results are first analyzed for significant differences in speech onset latencies depending on the number and type of featural similarity between primes and targets. Results are then organized into prime-target matrices which are subjected to multi-dimensional scaling analysis. This method derives a multi-dimensional stimulus configuration that reveals underlying patterns in the data. Finally, multiple regression analysis is performed on the results in order to evaluate how well the data can be predicted as a linear combination of the various variables suggested by the preceding analyses. The discussion section at the

* Many thanks to Allard Jongman, Michael Spivey, Joan Sereno, Carol Krumhansl, and members of the Cornell Phonetics Laboratory for thoughtful advice and criticism.

end compares the results of each of these analyses and shows how each contributes to insight on the units of perception and the processing involved in accessing the internal representation of these consonants.

2 Background of the methodology of priming

The methodology of priming involves the examination of behavior in response to a given target stimulus for which the context has been systematically manipulated by a particular prime stimulus. Priming is evidenced by a significant difference in reaction time or accuracy in response to related versus unrelated prime-target stimulus pairs. By manipulating the relationship between the primes and targets, experimenters try to determine which aspects of this relationship give rise to an observed priming effect. In this section, the methodology of priming is examined as it has been used to investigate both the processes and the units of perception involved in word recognition. The section concludes with a detailed discussion of the relevant variables and the effect each has on the results of phonological priming studies.

2.1 Insights into the nature of processing

Priming studies provide insights into the nature of processing through an examination of the contexts in which priming effects are observed. In comparisons with unprimed conditions, primed conditions have been shown to result in faster reaction times in a naming task (Bowey, 1996) and increased accuracy on a perceptual identification task (Slowiaczek, Nusbaum, and Pisoni, 1987). These differences in responses are taken to reflect differences in the processing of the target due to the contextual paradigm of the experiment. The majority of priming studies compares reaction times for targets preceded by related and unrelated primes. Although priming effects are generally small in absolute magnitude, (e.g. reaction time differences commonly much less than 100 ms) their statistical significance supports the conclusion that the processing of the target is in some way altered in the presence of the related prime.

This difference in processing may be manifested in one of two ways, corresponding to the two main types of priming effects: facilitation (or positive priming) which refers to faster reaction times and/or greater accuracy compared to controls, and inhibition (or negative priming, sometimes called interference) which refers to slower reaction times and/or more errors compared to controls. The difference between facilitation and inhibition has been shown to depend at least in part on the magnitude of the inter-stimulus interval (ISI) and the presence of intervening stimuli between the prime and the target. In a phonological priming study using a shadowing task, Banks, Roberts, and Ciranni (1995) found inhibition effects with short ISIs, and facilitation for the same prime-target pairs when there were intervening stimuli and thus more time between the prime and the target. These results show that processing can be differentially affected in the same task by the same prime-target relationship (in this case identical words spoken by different speakers) depending on the time course of serially presented stimuli.

The fact that priming effects are observed with respect to a particular relation between the prime and the target suggests that the processing of the prime in some way influences the processing of the target. The precise nature of this influence is unclear. Generally, the related prime and target are said to share some particular feature such as a degree of phonological similarity or some semantic association. The activation of the internal phonological or semantic representation corresponding to the prime may overlap with the activation of the internal representation of the target, but since the subjects are usually asked to respond only to the targets, it is not claimed that the processing of the prime and the target are entirely parallel. In fact, priming effects have been shown to vary with different tasks (e.g. Lukatela and Turvey, 1990), thus the "processing" of the target cannot be discussed without reference to the task through which the processing is revealed. Studies using a masked priming paradigm have further demonstrated that there are measurable differences between the processing of primes and targets. Grainger and Ferrand (1996) and Ferrand, Segui, and Grainger (1995) found priming effects with visually presented masked primes which were shown not to require conscious processing (durations of 43 and 29 ms identified with only 60% and 10% accuracy respectively, in pretests). These two studies confirm a difference at least in the extent of processing between the prime and the target.

2.2 Insights into the content and form of the internal representations

Taken together, phonological priming studies suggest that there is no single discrete unit of perception in word recognition to which manipulations affect priming results. Rather, it has been shown that priming can result from an overlap of features (Connine, Blasko, and Wang, 1994, Milberg, Blumstein, and Dworetzky, 1988), phonemes (Slowiaczek and Hamburger, 1992, Goldinger, Luce, Pisoni, and Marcario, 1992), syllabic onsets (Corina, 1992), syllabic rimes (Burton, Jongman, and Sereno, 1997), syllables (Jakimik, Cole, and Rudnicky, 1985), words (Slowiaczek and Pisoni, 1986, Lukatela and Turvey, 1991), and nonword homophones (Lukatela and Turvey, 1993). Many studies have contrasted the overlap of more than one of these units, but the findings vary depending upon the experimental paradigm used.

There is evidence that the internal representation is not limited to static, abstract segment information, but also contains information about statistical properties related to segment use in the language. Bowey (1996), for example, found that irregular orthographic rime primes such as "ind" facilitated the naming of typically irregularly pronounced target words such as "grind" more so than regular rime primes (e.g. "int") facilitated the response to inconsistent typically regular words (e.g. "flint"). This result is in contrast to previous findings which showed that orthographic primes with a consistently regular pronunciation produced faster responses than primes with an atypically irregular pronunciation. Bowey points out, however, that these data can be explained with reference to statistical properties of the language and the experiment design. The "ind"-type rimes are almost always pronounced [aind] in monosyllabic English words, and all targets in the experiment were monosyllabic English words. Statistical properties such as these must therefore be included in or accessible to the internal representation.

Further evidence that the internal representation contains or has access to information about statistical factors within the experiment comes from numerous studies which have found repetition or identity priming (e.g. Rossi, Burton, Jongman, and Sereno, 1993, Gaskell and Marslen-Wilson, 1996, Marslen-Wilson, Nix, and Gaskell, 1995, Radeau, Morais, and Dewier, 1989). These studies indicate that the processing of a given target is facilitated when the target is identical to the prime, or when the target has been previously presented (e.g. either intentionally as in studies explicitly investigating repetition priming, or inadvertently as in a previous block of trials within an experiment as found by Bowey, 1996). These findings demonstrate that information about what has come before is contained in the internal representation in some manner, perhaps through lower thresholds of activation for the internal representation of repeated targets or faster access to those representations.

Additional insights into the content and form of the internal representation come from studies showing that variations within segmental units can also affect priming results. Andruski, Blumstein, and Burton (1994), for example, found that the magnitude of semantic priming in a lexical decision task is influenced by manipulations to the VOT of the initial voiceless consonant in the prime word. Marslen-Wilson, Nix, and Gaskell (1995) found that identity primes which had been gated to exclude the final consonant release still facilitated a lexical decision task for targets, but that this facilitation was reduced more for primes followed by coronal versus non-coronal targets. This suggests a difference between coronals and non-coronals either in the internal representation or in the processing of acoustic cues sufficient to access this representation. That is, it may be that the more extreme formant transitions associated with non-coronals facilitate categorization and thus activate the primes to a greater extent (or perhaps just more quickly). Rossi, Burton, Jongman, and Sereno (1993) found that isolated tokens of [f] with matching coarticulatory information primed the two-alternative forced choice decision of [fi] and [fu] tokens more so than similar prime target pairs containing the fricative [s]. An acoustic analysis of the primes revealed large coarticulatory differences between the [f] tokens excised from [i] and [u] environments, and little difference between the different [s] tokens. Thus, the same coronal versus non-coronal pattern which emerged for the degraded stimuli in Marslen-Wilson, Nix, and Gaskell (1995) is evident here with full consonants.

Less clearly interpretable are the findings of Connine, Blasko, and Wang (1994) who examined semantic priming effects with primes that had an ambiguously voiced initial phoneme. They found that the voiced interpretation facilitated the lexical decision of a target in comparison to the voiceless interpretation unless ambiguously voiced auditory controls were included as primes, in which case it was the voiceless interpretation which was the more effective prime. Together, these studies indicate that fine grain acoustic differences can differentially affect priming results, but, as is the case for larger phonological differences, these effects are dependent upon many experimental factors and are therefore not always present in the results. Gaskell and Marslen-Wilson (1996), for example, showed that easily perceptible featural changes resulting in mismatching coarticulatory information did not negate the priming results found for unchanged primes.

Certain phonological priming effects which bear upon this question of the nature of the internal representation may best be interpreted with reference to current knowledge of auditory neural mechanisms such as adaptation, facilitation, and long-lasting inhibition. Using single-cell recordings of the neural discharge patterns of a cat, Delgutte and Kiang (1984) investigated the context-dependent encoding of speech sounds using synthetic stimuli including 6 identified as /da/, /ada/, /na/, /ša/, /sa/, and /tša/ respectively. All stimuli shared the /da/ portion and differed only in the preceding context information. They found that the response patterns of an auditory-nerve fiber show a clear peak in discharge rate for the beginning of the rapid formant transitions in /da/ stimuli resulting from adaptation. This peak is much smaller in response to the stimuli for which the preceding context produces either an increase (as for /ša/ and /sa/), or a decrease (as for /ada/) in discharge rate compared to the base firing. Thus, the neural representation corresponding to the /da/ portion of the waveform, at this level of processing, contains more information to contrast these stimuli than is present in the waveform itself. Delgutte (1997) goes on to report that response patterns of a binaurally-excited inferior colliculus unit to the same 6 synthetic stimuli result in an even greater differentiation of response. The pattern for /da/ has a clear peak at the onset of the transitions, followed by an immediate dramatic decrease in firing and then relatively sustained intermediate rate corresponding to the vowel portion of the stimulus. The amplitude and shape of the peak is different for each of the other six stimuli. Thus context dependencies are robustly encoded at this level of processing.

The robust encoding of onsets and the context-sensitive interactions in neural responses may be partially responsible for various phonological priming effects. Specifically, the relative position of phonological overlap (e.g. initial versus final) has been shown to be an important factor in priming studies (Praamstra, Meyer, and Levelt, 1994). This may be related to the precision with which onsets are encoded. Also, the importance of syllabic constituency in priming (versus number of phoneme overlap) (e.g. Radeau, Morais, and Segui, 1995) may be due to the way in which larger phonological units are represented in the more central areas of the auditory system.

In regards to the form and content of the units of perception in word recognition then, evidence of phonological priming in word recognition suggests that the internal representation contains (or has access to) a rich array of information about each of the phonological components discussed above, including subphonemic fine-grained acoustic properties as well as statistical properties such as frequency of occurrence in the language or the experimental paradigm. The success of this methodology in revealing subphonemic effects in word recognition indicates that this may also be an appropriate technique for investigating phoneme recognition directly. With this objective, I turn to an examination of how the relevant variables have affected the findings of phonological priming studies.

2.3 Relevant parameters in priming studies

One of the difficulties involved in presenting a coherent picture of phonological priming stems from the enormous amount of variation among the studies and their findings. Task effects, presentation factors, and stimulus factors all contribute to this variation, as well as the failure of some results to be replicated by others (e.g. Martin and Jensen's 1988 failure to replicate Hillinger's 1980 findings). This section provides a review of how variations of each of the relevant factors have contributed to the experimental findings.

2.3.1 Task effects

Different tasks have been shown to result in different priming effects for the same prime-target relationship, even when the same stimuli and experimental conditions are used. Many priming studies have employed the following tasks as measures of word recognition: lexical decision, shadowing, naming, and perceptual identification. In the case of lexical decision, subjects have been instructed to respond to both the prime and the target (e.g. Hillinger, 1980), or just the target (this is the more common procedure, e.g. Burton, Jongman, and Sereno, 1997). They are usually instructed to make a response for every target, but have also been limited to responses just to word targets (Jakimik, Cole, and Rudnicky, 1985). The most consistent findings from lexical decision results include faster reaction times in response to phonologically identical prime-target pairs (e.g. Radeau, Morais, and Dewier, 1989) and faster responses to rhyming prime-target pairs (e.g. Praamstra, Meyer, and Levelt, 1994). The same degree of initial overlap often has no significant effect on reaction times (e.g. Praamstra, Meyer, and Levelt, 1994) or results in slower reaction times (Corina, 1992).

In a shadowing task, subjects are instructed to repeat the target word as quickly and accurately as possible. Results from shadowing are generally consistent with those from lexical decision. That is, identical pairs result in faster reaction times (Radeau, Morais, and Dewier, 1989) as do rhyming pairs (Radeau, Morais, and Segui, 1995), and as these same two studies find, initial overlap results in slower reaction times or no effect.

Subjects participating in a naming task are instructed to name visually presented targets. Studies which employ this task often compare responses to pairs sharing phonological and/or orthographic factors. Pairs sharing both types of information generally result in faster reaction times (Burton, Jongman, and Sereno, 1997) .

In a perceptual identification task, subjects are asked to type the target word. Priming effects are generally reported in terms of accuracy or error data. Perceptual identification results are often in contrast to lexical decision and shadowing results for prime target pairs which share initial phonemes. Initial overlap generally results in more accurate perceptual identification responses (Goldinger, Luce, Pisoni, and Marcario, 1992). Final overlap also results in more accurate responses (Slowiaczek, Nusbaum, and Pisoni, 1987).

2.3.2 Prime and Target Presentation Factors

Phonological priming has been investigated using different experimental conditions concerning the presentation of the primes and targets. The presentation can be: (1) auditory or visual, (within a single modality or cross-modally) (2) serial (with variations in ISIs) or parallel (simultaneous presentation of prime and target); (3) clear or embedded in noise; and (4) easily perceptible or masked. Although modality conditions are usually held constant within a given study, modality effects have been found for both lexical decision and naming tasks by Burton, Jongman, and Sereno (1997). In regards to the temporal conditions, a serial presentation consisting of a prime followed by a target with longer intervals between pairs is by far the most common method, but a simultaneous presentation has been used both within the auditory modality (Banks, Roberts, and Ciranni, 1995) and cross-modally with auditory primes and visual targets (e.g. Gaskell and Marslen-Wilson, 1996). A few studies have presented targets in noise and have found that initial phonological overlap facilitates perceptual identification (Slowiaczek and Pisoni, 1987, Goldinger, Luce, Pisoni, and Marcario, 1992). Studies using masked primes with different tasks have found phonological priming effects consistent with studies which use unmasked (normal) primes: initial overlap facilitates naming, but not lexical decision, and a combination of phonological and orthographic overlap is more effective at priming than orthographic overlap alone (Grainger and Ferrand, 1996).

Variations in the ISI have repeatedly been shown to affect the presence or magnitude of priming, generally in the direction of longer ISIs resulting in little or no priming effect. The semantic facilitation found by Andruski, Blumstein, and Burton (1990), for example, for targets following primes with VOT-manipulated initial voiceless consonants was only present with short ISIs of 50 ms, whereas no effect was found with an ISI of 250 ms. Similarly, the inhibitory effect found in a perceptual identification task by Goldinger, Luce, and Pisoni (1989) for phonetically close prime-target pairs was present with an ISI of 50 ms, but there was no effect with an ISI of 500 ms. Goldinger, Luce, Pisoni, and Marcario (1992) used similar stimuli in a lexical decision task and found the same effect of ISI: at an ISI of 50 ms, subjects were less accurate for the phonetically close stimuli, but there was no effect at an ISI of 500ms.

2.3.3 Stimulus factors

Stimulus set characteristics have also been shown to influence the presence and magnitude of priming effects. For phonological priming especially, repetition of target stimuli within the experiment can confound results through a combination of repetition priming and the intended prime-target relationship effect (Bowey, 1996). The percentage of related pairs within the stimulus set has also been shown to affect results. The facilitation found by Goldinger, Luce, Pisoni, and Marcario (1992) for phonetically close prime-target pairs sharing an initial phoneme was greatly reduced when the percentage of related pairs in the stimulus set was reduced from 50% to 10%. The type of fillers (words or nonwords) used in a lexical decision task can also be important, as is the ratio of words to nonwords used as primes and/or targets. Radeau, Morais, and Dewier (1989) found differences in priming effects for word and nonword targets such that initial phonological overlap increased the reaction times for words but slowed reaction times to nonwords. Other important stimulus factors are word frequency and neighborhood density. Reaction times are faster for high frequency and low density targets (Radeau, Morais, and Segui, 1995).

The stimulus characteristics within prime-target pairs is another major factor. Much of the variation in phonological priming is due to the degree of phonological similarity in prime-target pairs and the relative position of that similarity, and these variables have been manipulated in various ways by many studies. In auditory priming, careful consideration is needed concerning the amount of phonetic similarity present as well. Rossi, Burton, Jongman, and Sereno's (1993) findings suggest stronger identity priming effects resulting from the use of the same acoustic token in both the prime and the target.

3 Motivation for the present study

Results from a series of gating experiments (Doeleman, 1998) suggested a special status for place feature information allowing it to reduce ambiguities of voicing or manner and improve identification scores for the English consonants [p, b, t, d, f, v, s, and z]. Subjects in the Place group, who heard the stimuli blocked by consonant place ([p, b, f, v] in one block and [t, d, s, z] in the other), had significantly higher identification scores than subjects who heard the same eight stimuli unblocked or blocked by manner ([p, b, t, d] and [f, v, s, z]). Since the stimuli were presented aurally, it was hypothesized that the higher identification scores for the subjects who heard the stimuli blocked by the place feature could be due to phonetic priming of this feature.

The purpose of the present study is to explore the priming of phonetic features in a shadowing task. Using the methodology described in the previous section, it is possible to systematically manipulate featural similarity between the prime and target stimuli. One hypothesis that may account for the facilitation in identification found in the previous study is that the consonant representations were phonetically primed by the place feature shared by successively presented stimuli. A testable prediction of this hypothesis, then, is that prime-target syllable pairs in which the consonants share place should give rise to a phonetic priming effect when compared to unrelated pairs or pairs which do not share place.

Another aim of these priming experiments is to further explore the effect of the stimulus blocking manipulation used in the gating experiments. At this point, it is unclear whether the results from the gating study were due to the subjects' prior knowledge of the particular feature shared among the consonants in the response sets, or whether identification was aided or hindered by the featural differences remaining between members of the response sets. In other words, it may be the case that the facilitative effect for the Place group was due to the fact that, for example, the consonants in the response set [p, b, f, v] shared the place feature labial. Alternatively, it could be that the facilitative effect was the result of less confusability among these consonants due to the remaining differences in voicing and manner. By performing this same feature-blocking manipulation in the priming experiments, it is possible to test these two hypotheses. If the first hypothesis is true, that is, if prior knowledge of the place feature of the response facilitates access of the target consonant, then the reaction times for the group presented with stimuli blocked by target place should be shorter than those of the other groups.

A broader objective of the present study is to systematically explore possible any phonetic priming of features. The design of the experiment allows an examination of several planned comparisons. The first is whether there is any correlation between the number of shared features (0-3) and the existence or extent of phonetic priming. The second is whether responses to the target stimuli depend upon the type of featural similarity between the prime and the target. The various featural similarity conditions included in the experiment are: identity, shared manner and voicing, shared manner and place, shared voicing and place, voicing only, place only, manner only, and unrelated. Data from these conditions can be also combined in order to examine the effect of sharing a particular feature (such as place) or a particular feature specification (such as labial).

Of the various tasks which have been used in priming experiments, the most consistent results have been found for lexical decision and shadowing. For the present study, for purposes of relating the results to those of the gating experiments detailed in Chapter Two, the prime-target stimulus pairs consist of sequences of syllables from the same consonant inventory used in the gating experiments. Since these stimuli are not all lexical items, a lexical decision task is inappropriate. Rather, the task chosen was that of shadowing.

4 Experiments 1a, 1b, 1c, and 1d: An investigation of phonetic priming

The role of features in speech perception was investigated in a series of priming experiments in which subjects shadowed CV syllables consisting of the English consonants [p, b, t, d, f, v, s, z] followed by the vowel [a]. Natural tokens were used to create sixty-four prime-target pairs which varied critically the number of features the consonants shared (0-3) and the type of featural similarity. Subjects were asked to repeat the target syllables from aurally presented prime-target pairs as fast as possible. Patterns of speech onset latencies and accuracy data are compared for the different featural similarity conditions.

4.1 Methods

4.1.1 Stimuli

The recorded stimuli were CV utterances consisting of the consonants [p, t, f, s, b, d, v, z,] followed by the vowel [a]. The stimuli were produced by a male native speaker of American English (different from the speaker recorded in the gating study). The speaker was recorded in a soundproof booth, using a cardioid microphone (Electrovoice, model RE 20) and a high-quality cassette recorder (Carver, model TD-1700). The speaker produced three repetitions of each target syllable in isolation at a normal speaking rate. The recordings were then digitized on a SUN SPARCstation2 at 22050 Hz and stored as files to be processed by the commercial software package WAVES+/ESPS (Entropic, Inc.). Two tokens of each syllable were chosen for inclusion in the present study, one to be used as consistently as a prime and the other as a target. The amplitude of the vowel portion of each syllable was standardized to 70 dB and the syllables were presented to listeners in an identification task to ensure that the syllables contained clear exemplars of the intended consonants.

4.1.2 Subjects

Thirty-eight students at Cornell University served as paid subjects in this study. All subjects were native speakers of American English who reported no history of speech or hearing disorders. Each subject participated in one of four experimental groups. Subjects with greater than 10% error rates and subjects who consistently confused one or more targets in the experiment (generally [ba] heard as [va]) were excluded from the analysis. This was the case for 10 subjects, leaving 28 subjects - 6 or 8 in each experiment group.

4.1.3 Design and Procedure

The stimuli described above were transferred directly from the SUN SPARCstation2 to a Swan 386/25 PC. Using BLISS software (Mertus, 1989), prime-target pairs were created by combining the eight different syllables in all possible orders. Table 1 shows the sixty-four resulting prime-target pairs categorized by the number of features shared and the type of featural similarity between the consonant in the prime syllable and the consonant in the target syllable. Five repetitions of each stimulus pair were presented auditorily over Sony headphones, with a 50 ms inter-stimulus interval between the prime and the target and 3 second inter-trial interval. The 50 ms ISI was chosen on the basis of the studies described in section 3.2.3.2, which found a priming effect at this ISI, but not at longer ISIs of 250 or 500 ms.

Prime – Target Stimulus Pairs

Identity	**Two Features Shared**			**One Feature Shared**			**Unrelated**
	M & V	**V & P**	**M & P**	**M**	**V**	**P**	
pa-pa	ta-pa	fa-pa	ba-pa	da-pa	sa-pa	va-pa	za-pa
ba-ba	da-ba	va-ba	pa-ba	ta-ba	za-ba	fa-ba	sa-ba
ta-ta	pa-ta	sa-ta	da-ta	ba-ta	fa-ta	za-ta	va-ta
da-da	ba-da	za-da	ta-da	pa-da	va-da	sa-da	fa-da
fa-fa	sa-fa	pa-fa	va-fa	za-fa	ta-fa	ba-fa	da-fa
va-va	za-va	ta-sa	fa-va	sa-va	da-va	pa-va	ta-va
sa-sa	fa-sa	ba-va	za-sa	va-sa	pa-sa	da-sa	ba-sa
za-za	va-za	da-za	sa-za	fa-za	ba-za	ta-za	pa-za

Table 1. Prime-target pairs categorized by number of features shared and type of featural similarity (M = manner, V = voicing, P = place).

Subjects began the experiment with a short identification task which included two repetitions of each of the sixteen acoustically different prime and target syllables, presented in isolation. This task served to familiarize the subjects with the stimuli to be used in the priming experiment and ensure that errors in shadowing were task- specific and not due to a general inability of the subject to perceive the intended target consonant.

For the shadowing task, subjects were told that they would hear sequences of the syllables from the identification task and were instructed to repeat the second (target) syllable as fast as possible. A microphone attached to the headphones transmitted the subjects' responses simultaneously to a voicekey and a high-quality analog tape recorder. Speech onset latencies (SOLs) were measured from the onset of the target syllable to the onset of the subjects' response. To encourage rapid responses, subjects received verbal feedback from the experimenter regarding their average SOLs after practice trials and throughout the experiment, during the pauses between blocks of stimuli.

The twenty-eight subjects whose performance on identification and shadowing met inclusion criteria each participated in one of four experiments. The eight subjects in Experiment 1a, designated the Unblocked Group, heard five repetitions of the sixty-four stimulus pairs blocked by repetition. In a manipulation similar to that done in previous gating experiments (Doeleman, 1998), the subjects in Experiments 1b, 1c, and 1d also heard five repetitions of the same 64 prime-target pairs, but the stimuli were presented in blocks corresponding to a particular feature of the target consonant. Thus the six subjects in Experiment 1b, designated the Voicing group, shadowed all stimuli blocked by target voicing, with voiced targets in one block and those with voiceless targets in a separate block. The six subjects in Experiment 1c, the Place group, heard the stimulus pairs blocked by labial targets and alveolar targets. The eight subjects in Experiment 1d, the Manner group, shadowed all stimuli blocked by target manner, that is, those with stop consonant targets in one case and fricative targets in the other. Subjects in Experiments 1b, 1c, and 1d were explicitly told that although the first syllable of each pair they would hear could be any of the eight syllables heard in the identification task, the second syllable would only be one of four syllables. The target syllables were specified aurally before each block, and the subjects heard practice trials containing only those targets.

4.2 Analysis of results

The final data set consisted of 8960 responses from which extreme outliers (SOLs less than 150 ms or greater than 1200 ms) were excluded. Following Rogers and Storkel (1998), the *transformed speech onset latency* (TSOL) served as the primary dependent variable in all analyses (except that regarding the number of errors). This measure of response latency was obtained by first calculating the median speech onset latency (SOL) for each subject for each target and then subtracting this amount from all SOLs corresponding to that target for that subject. The aim of this transformation was twofold. First, it served to eliminate variability in the data set which is due to absolute differences in the timing of articulation of the eight consonants. Second, this procedure also minimized intersubject variability in the data set since the transformations were no longer measures of absolute reaction times.

4.2.1 Error analysis

Subjects were monitored for accuracy both during the experiment and from the audio recordings of their responses, the latter by both the experimenter and an assistant. SOLs corresponding to errors were subsequently removed from the data set and an analysis of variance (ANOVA) was carried out on the error data in order to investigate possible main effects of subject group and featural similarity condition. Figure 1 shows the number of errors for each subject group across the eight featural similarity conditions. The chart legend also indicates the overall percentage of errors for each group. Results of the ANOVA show no significant difference in error rates for the different subject groups, and no effect of featural similarity condition on error patterns.

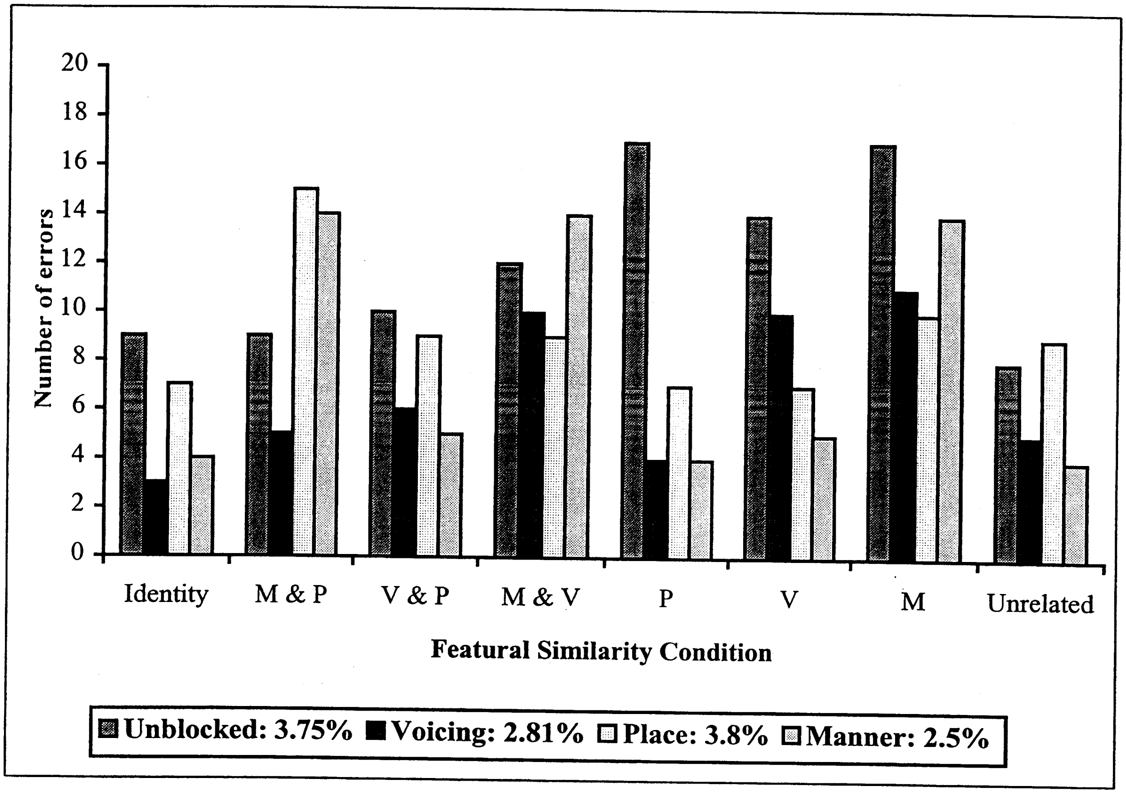

Figure 1. Error rates for each subject group across featural similarity conditions.

4.2.2 Featural similarity conditions

Figure 2 shows the mean TSOL for each subject group for each featural similarity condition. An ANOVA revealed no significant differences between the mean TSOLs of the different subject groups. Based upon this result, the data were collapsed across all four subject groups for all subsequent analyses.

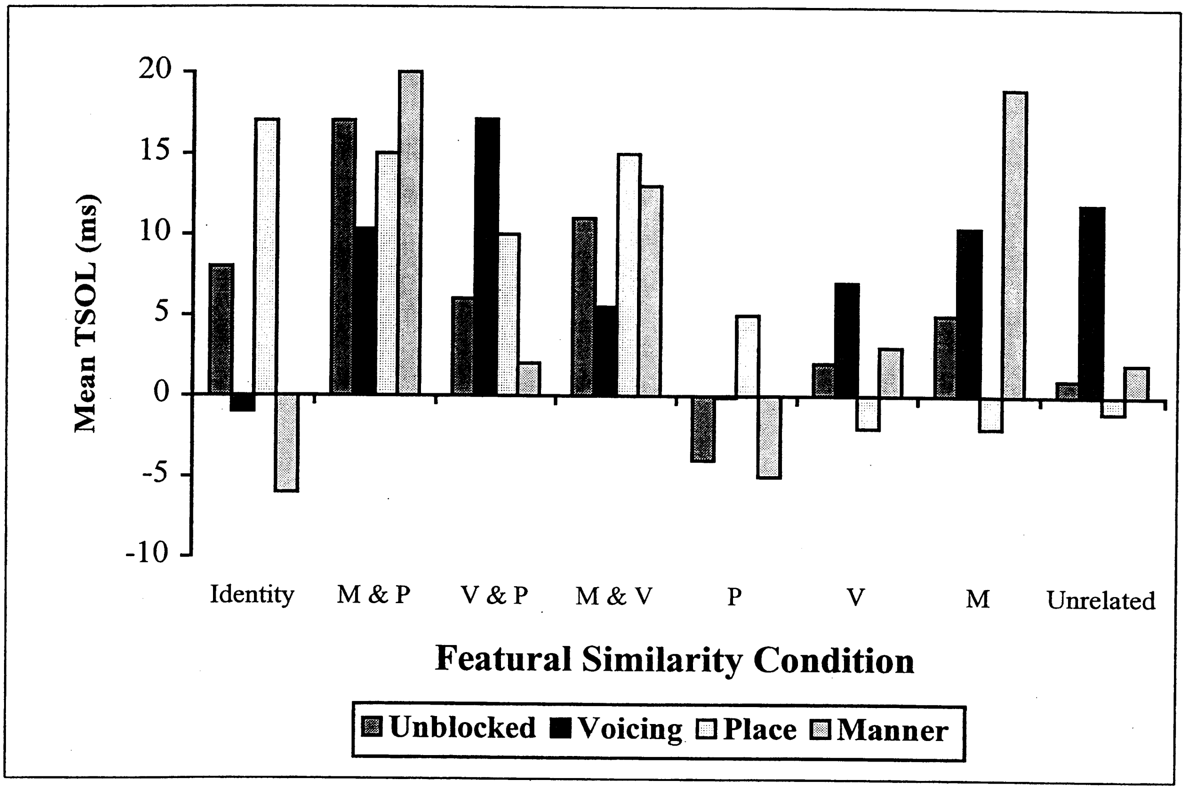

Figure 2. Mean TSOLs for each subject group for each featural similarity condition: Identity = shared manner, place, and voicing, M&P = shared manner and place, V&P = shared voicing and place, M&V = shared manner and voicing, P = shared place only, V = shared voicing only, M = shared manner, Unrelated = no features shared.

Figure 3 shows the mean TSOLs for each featural similarity condition collapsed across the four subject groups. A one-way ANOVA revealed a main effect for condition [F(7, 189) = 4.719, p = .000]. A Bonferroni's post-hoc analysis showed that the latencies for the shared manner and place condition (M&P) were significantly longer than those of the identity, shared place only (P), shared voicing only (V), and the unrelated conditions. Latencies for the shared manner and voicing condition (M&V) were also significantly longer than those for shared place only (P).

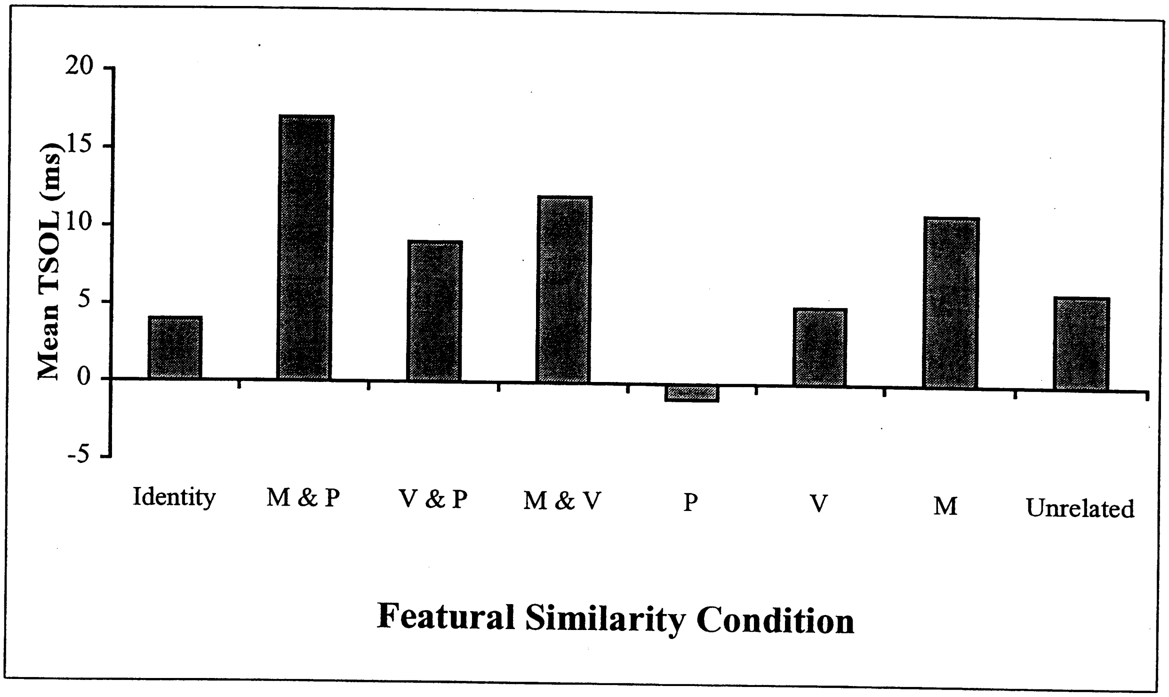

Figure 3. Mean transformed speech onset latencies (TSOLs) for each featural similarity condition.

4.2.3 Number of shared features

Figure 4 shows the mean transformed speech onset latencies for the data when categorized strictly by the number of features shared between the prime and the target. A one-way ANOVA revealed a main effect for the number of features [$F(3, 81) = 5.838$, $p = .001$] and a Bonferroni's post-hoc analysis showed that the latencies for the two-features-shared condition were significantly longer than the one-feature-shared and no-features-shared conditions.

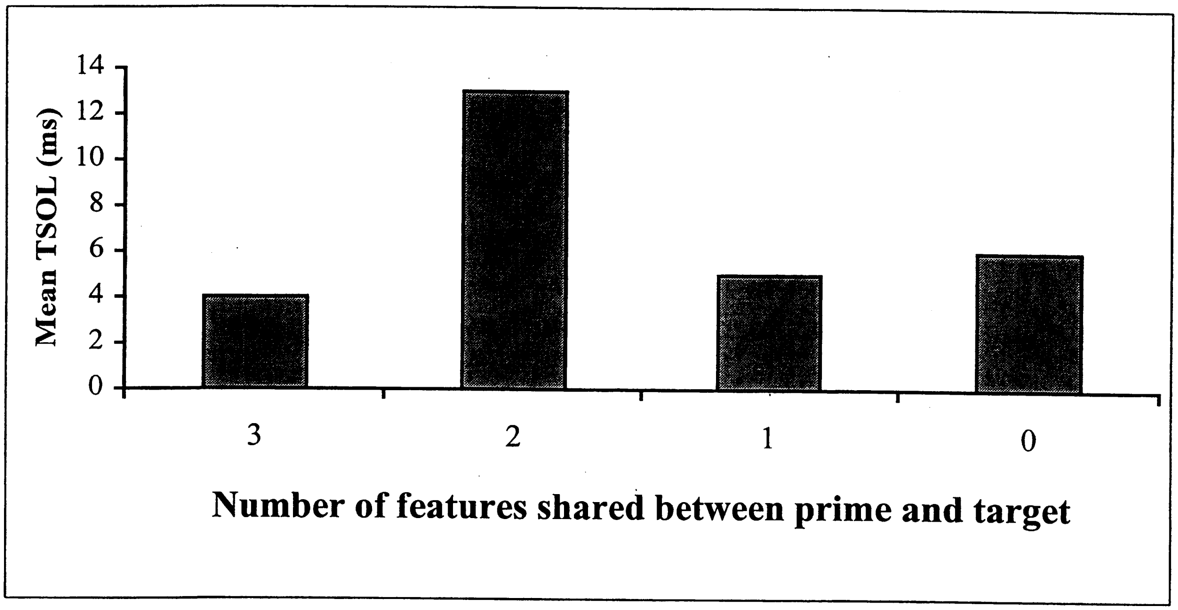

Figure 4. Mean transformed speech onset latencies (TSOLs) for all data categorized according to the number of features shared between the prime and the target.

4.2.4 Effect of sharing voicing, place, or manner

Figure 5 shows the mean TSOLs for the data grouped according to whether or not the consonants in the prime-target pairs share each of the features: voicing, place, or manner. A one-way ANOVA revealed a main effect of shared feature only for the manner feature [$F(1, 27) = 13.462$, $p = .001$]. TSOLs for prime-target pairs which shared manner were significantly longer than those which did not. No such effect was found for shared place or shared voicing.

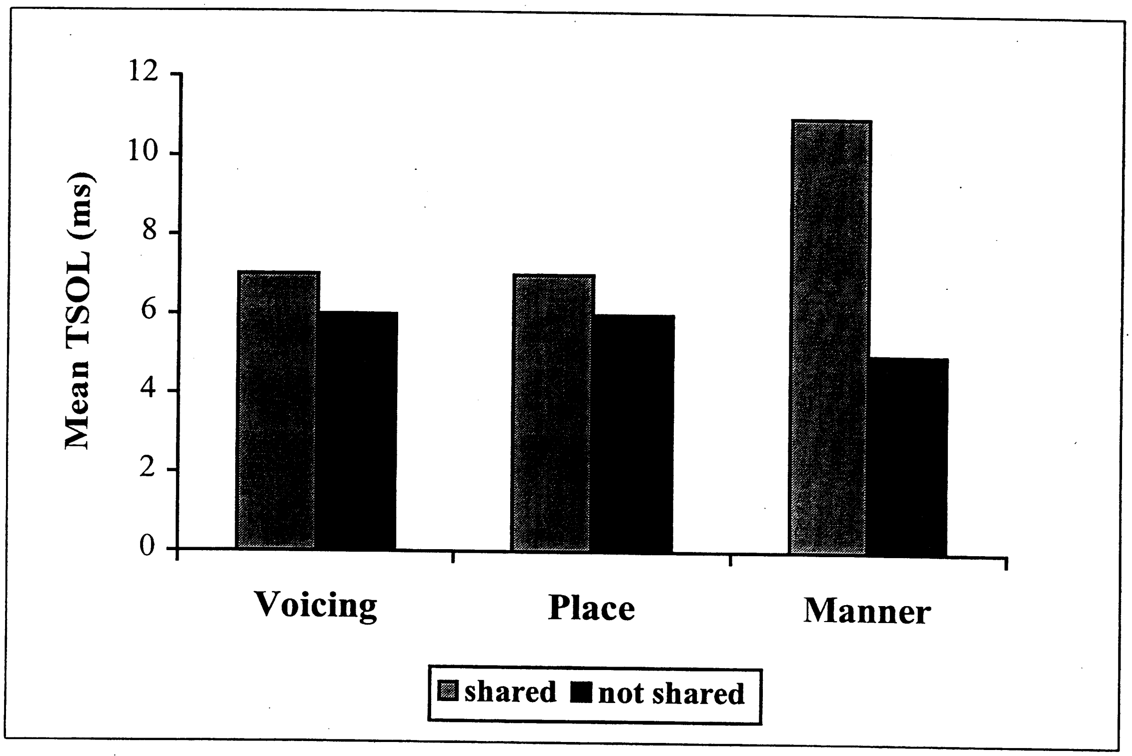

Figure 5. Mean TSOLs for prime-target pairs sharing (and not sharing) each of the features: voicing, place, and manner.

4.2.5 Analysis of the prime-target relationship in terms of particular feature specifications

In order to further investigate the effect of shared manner illustrated in the previous section, the prime-target relationships within the stimulus set were further categorized by the particular voicing, place, and manner specifications of the prime and the target. One advantage of this analysis is that it may reveal differences in response latencies which are asymmetrical with respect to the particular feature specifications of the consonants included in the experiment.

Figure 6 shows the mean TSOLs for the particular manner specifications (i.e. stop or fricative) of the primes and targets. A one-way ANOVA revealed a main effect for manner specification [$F(3,81) = 4.061$, $p = .01$]. A Bonferroni's post-hoc analysis revealed that the latencies for the stop-prime/stop-target condition were significantly longer than those for the two unshared manner conditions. Results from this analysis suggest that the conclusion reached by the shared feature analysis was rather misleading.

That is, it is not truly the case that latencies are longer strictly due to shared manner since the latencies for the fricative/fricative condition were not significantly longer than the two unshared manner conditions.

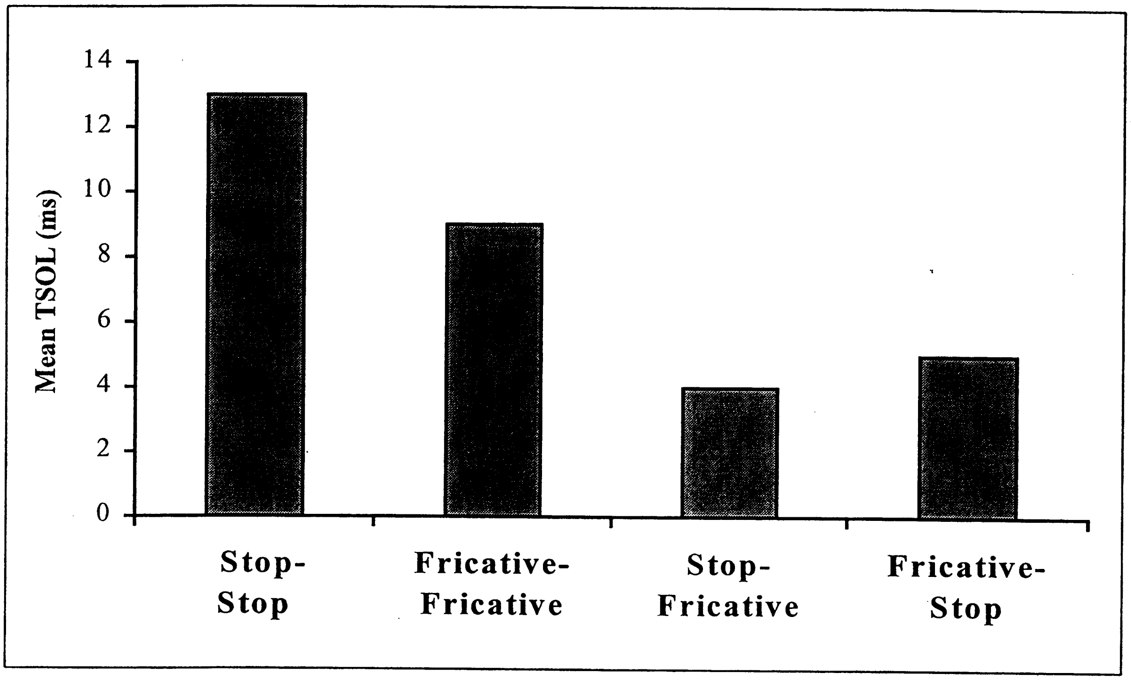

Figure 6. Mean TSOLs for all prime-target pairs categorized by manner specifications. The first manner specification of each pair refers to that of the prime; the second refers to that of the target.

Figure 7 shows the mean TSOLs for the particular place feature specifications (i.e. labial or alveolar) of the primes and targets. A one-way ANOVA revealed a main effect for place specification [$F(3,81) = 6.696$, $p = .001$]. A Bonferroni's post-hoc analysis revealed that the latencies for the alveolar-prime/labial-target condition were significantly longer than those for the labial/labial and labial/alveolar conditions.

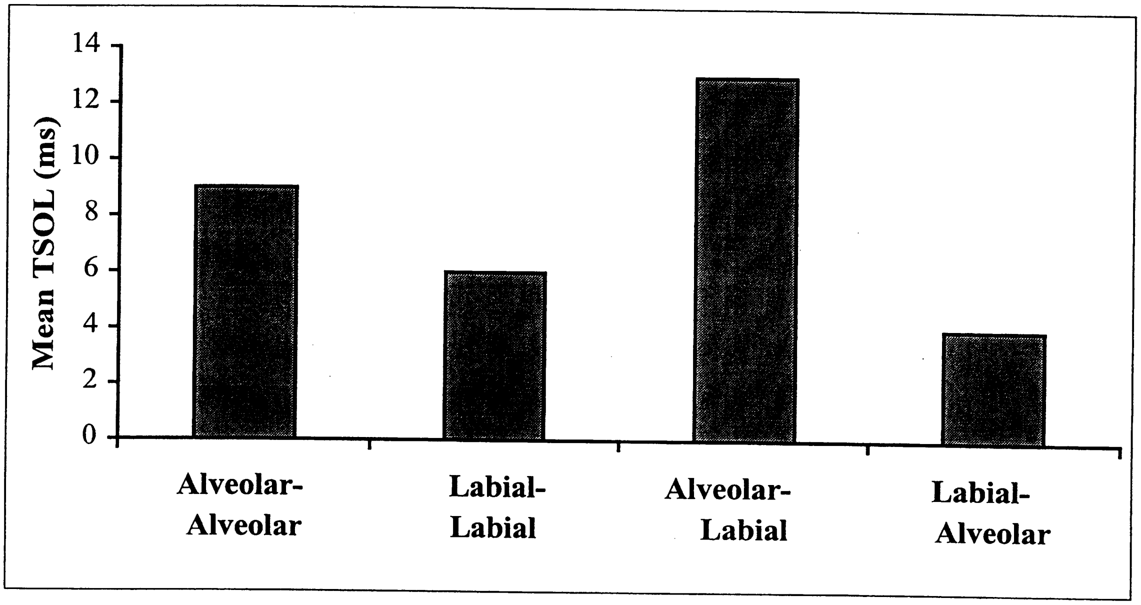

Figure 7. Mean TSOLs for all prime-target pairs categorized by place specifications.

Figure 8 shows the mean TSOLs for the particular voicing feature specifications (i.e. voiced or voiceless) of the primes and targets. A one-way ANOVA revealed no main effect for voicing specification.

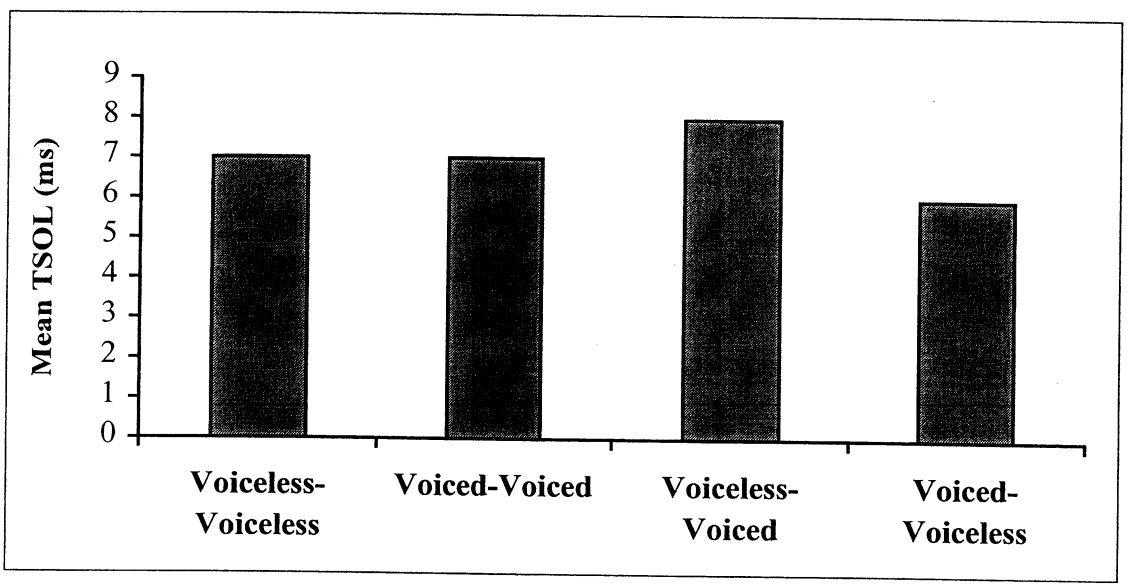

Figure 8. Mean TSOLs for all prime-target pairs categorized by voicing specifications of the primes and targets.

There was also a significant interaction of the manner and voicing specifications [$F(9, 243) = 3.325$, $p = .001$]. Figure 9 shows a composite of the latency information from

Figures 7 and 8, illustrating this interaction. A Bonferroni's post-hoc analysis shows that the mean latencies for stimulus pairs consisting of alveolar stop primes with labial stop targets are significantly longer than those of many other stimulus pairs, including all other alveolar primes with labial targets.

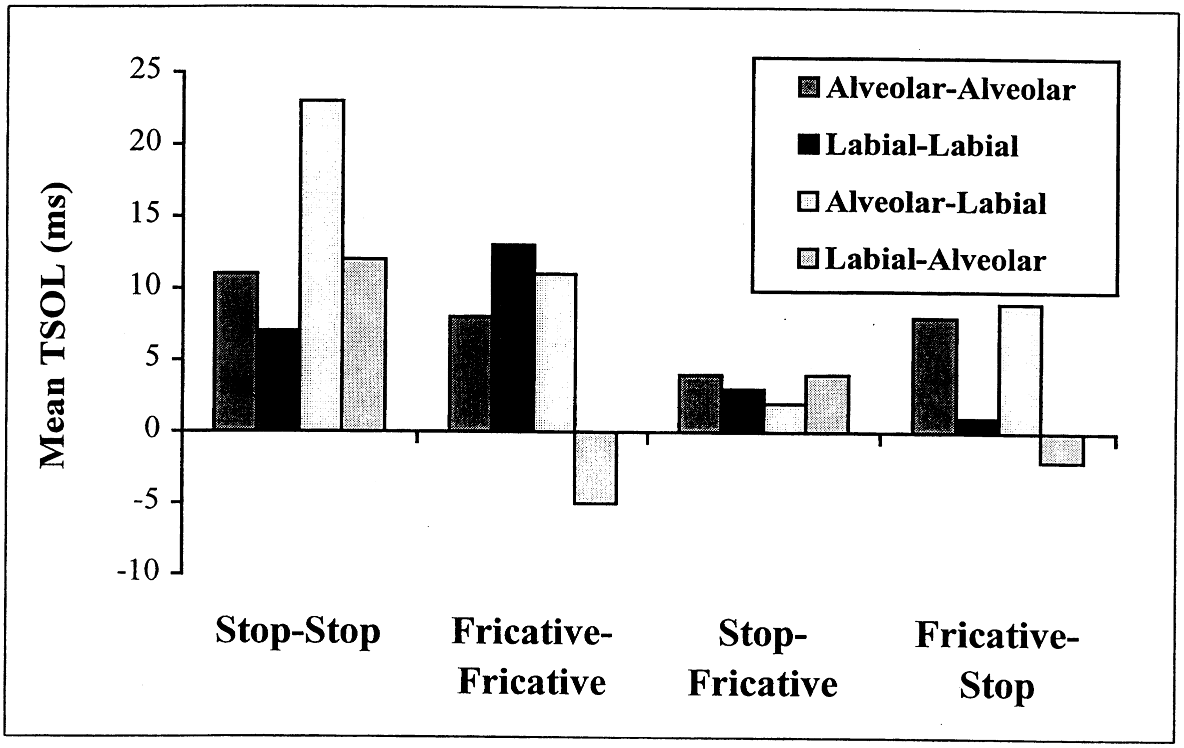

Figure 9. Mean TSOLs for specific place feature conditions for each specific manner feature condition.

4.2.6 Multi-dimensional scaling

In light of the asymmetries in the data that were revealed by the previous analysis, it is clear that a characterization of feature interactions which ignores the particular feature specifications of the consonants involved is inadequate to account for the data. In an effort to explore how the feature interactions might correspond to organizational properties of the internal representations of the consonants, a non-metric multi-dimensional scaling analysis was performed on the results. The aim of this analysis is to see if the latency data is meaningfully related in any way to the notion of perceptual distances between specific phoneme representations by examining whether the dimensions along which the data pattern correspond to any feature specifications. The resulting stimulus configuration may also reveal organizational differences between particular feature specifications (e.g. stops versus fricatives).

A particular type of multi-dimensional scaling, referred to as individual differences scaling, or INDSCAL, was chosen for the present study for several reasons. First of all, this more powerful analysis has already been used successfully with speech sounds by Wish and Carroll (1974) and Soli and Arabie (1979) on Miller and Nicely's (1955) confusion matrices (although a crucial difference between those analyses and the present analysis lies in the size of the matrices – larger matrices, such as those from Miller and Nicely (1955), have a greater number of degrees of freedom which may have contributed to the success of the analyses). Second, this analysis accounts not only for differences in response bias, but also for individual differences in the perceptual or cognitive processes that generate the dissimilarity data. This method of analysis also reports the importance of each dimension for each subject, as well as the overall importance of each dimension. Although multi-dimensional scaling has previously been used on confusion data, applying this analysis to the speech onset latency data obtained in the current priming experiments may result in a convergence of processing data.

In order to perform this analysis, the transformed speech onset latencies were organized into matrices, with rows corresponding to the prime syllable and columns corresponding to the target syllable. The latencies were scaled up to eliminate negative numbers in the data and were then transposed so that longer latencies indicated smaller "distances", or more similarity between the phoneme representations.

The assumption that longer latencies indicate smaller distances is consistent with neighborhood effects in word recognition. Several prominent models of word recognition, including the Neighborhood Activation Model (Luce, Pisoni, and Goldinger, 1990) and the COHORT model (Marslen-Wilson, 1987), are designed to account for evidence showing that words with many phonologically similar neighbors take longer on average to access than words with few neighbors. Thus, dense neighborhoods have an inhibitory effect on the access of a given item, and reflect similarity among neighbors. The pattern of interference for shared stop feature found in the previous analysis supports the assumption that longer latencies in the priming data indicate greater similarity.

A three-dimensional solution was chosen based on the following four criteria from Shepard (1974): goodness of fit, statistical stability, interpretability, and visualizability. Goodness of fit is measured by the stress value and is indicated by a sharp drop in stress, such as the drop from .272 to .137 observed between two- and three-dimensional solutions. The three-dimensional solution also maintains a relatively high RSQ value of .87. Furthermore, although no easily interpretable dimensions are evident in the four- or two-dimensional solutions, the three-dimensional solution reveals a stimulus configuration which corresponds to linguistic features. Figure 10 shows the derived stimulus configuration for this three-dimensional solution.

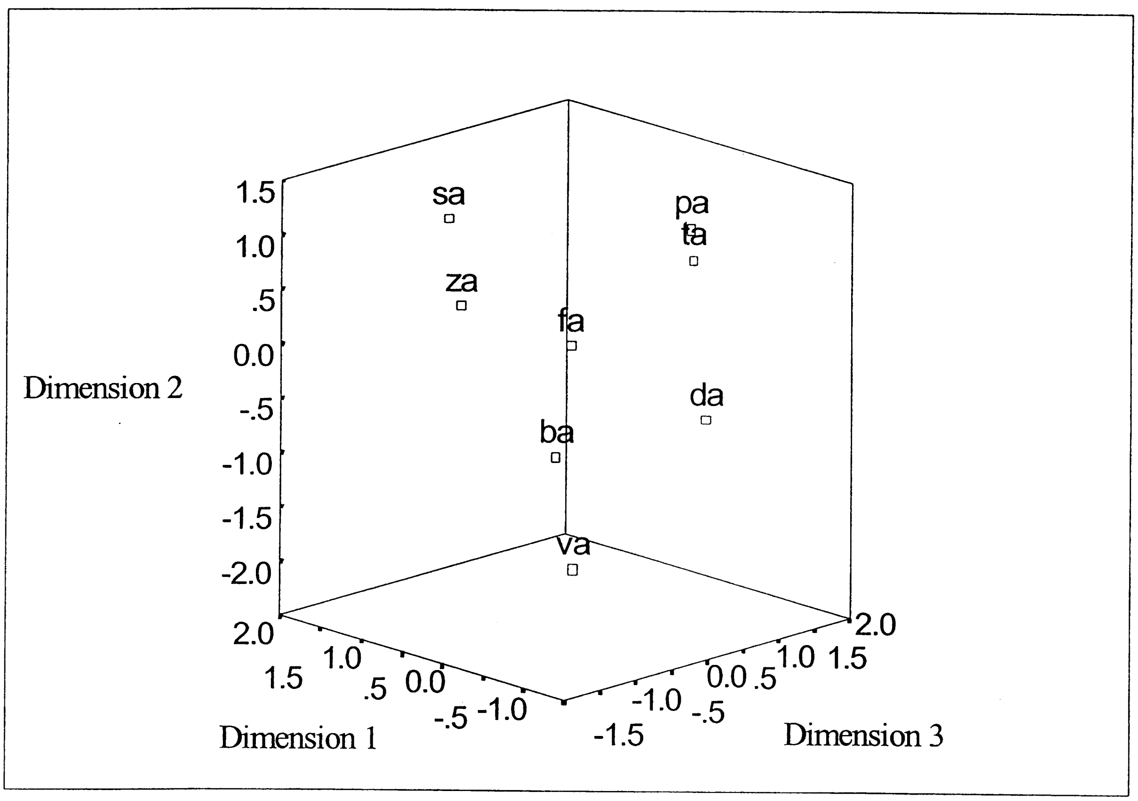

Figure 10. Derived stimulus configuration for three-dimensional non-metric multi-dimensional scaling model applied to transformed speech onset latencies from priming experiments.

Although the first dimension corresponds clearly to the manner distinction, the derived stimulus configuration in Figure 10 does not illustrate this well. For this reason, the stimulus coordinates for the three-dimensional solution depicted in Figure 10 are given in Table 2.

Stimulus	Dimension 1	Dimension 2	Dimension 3
pa	.0310	.8455	1.4740
ba	-.8347	-.5380	-1.3855
ta	-1.1989	1.0780	.1136
da	-1.2891	-.3892	.1834
fa	1.3084	-.4803	1.2718
va	.1865	-2.0561	.0204
sa	1.5474	.9076	-.1923
za	.2494	.6326	-1.4853

Table 2. Stimulus coordinates for the three-dimensional solution.

While the second and third dimensions are seemingly difficult to interpret from both the stimulus configuration in Figure 10 and from the coordinates in Table 2, visualizability and interpretability can be improved by examining these dimensions in a two-dimensional space. This technique has previously been used to illuminate the feature distinctions between consonants (e.g. Soli and Arabie, 1979). Figure 11 shows the two-dimensional spatial representation obtained from the stimulus coordinates of the second and third dimensions found in Table 2. From this perspective, these dimensions reveal the place and voicing distinctions of the eight consonants.

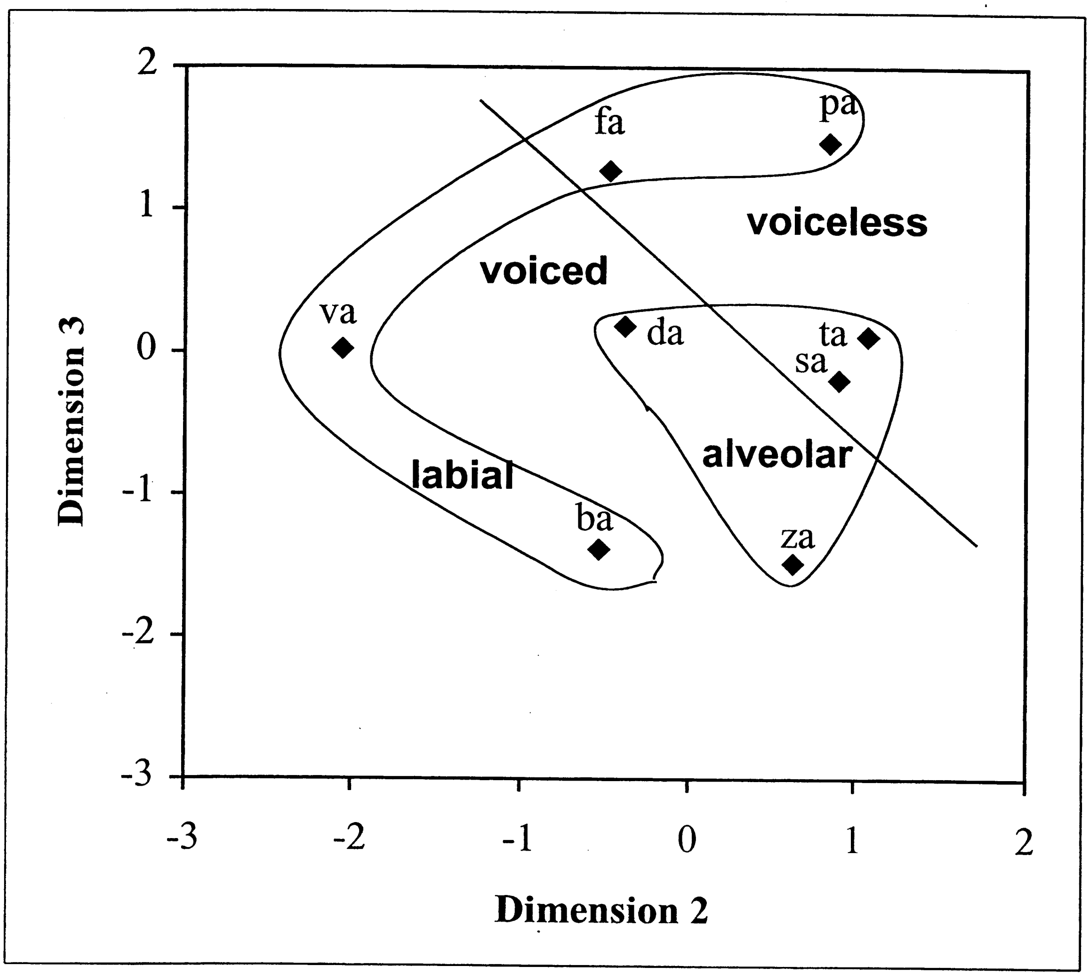

Figure 11. A two-dimensional spatial representation of the second and third dimensions of the three-dimensional solution.

In general, then, the multi-dimensional scaling analysis revealed manner as the primary dimension in the stimulus configuration, with place and voicing characterized by the combination of the second and third dimensions. The overall importance of the first dimension, corresponding to manner, was .4215, while the overall importance of the second and third dimensions was .2456 and .2046 respectively.

One advantage of multi-dimensional scaling that has been pointed out by Shepard (1974), is that this method provides an objective and uniform way of comparing the underlying patterns revealed from different experimental studies. To this end, this method was used to re-analyze the confusion matrices from the previous gating experiments (Doeleman, 1998). When the confusion matrices for the 20% to 40% duration conditions (those with ample non-random confusions) were subjected to this multi-dimensional scaling analysis, the resulting stimulus configuration was consistent with respect to the relevant parameters observed in the stimulus configuration derived from the priming data. The four-dimensional solution, which had a stress value of .14 (with a drop from .22 for the three-dimensional solution) and an RSQ of .87. The first dimension corresponded clearly to the manner distinction, with positive values indicating stop consonants and negative values indicating fricatives. Dimension two plainly coincides with the place distinction, with positive values indicating alveolars and negative values indicating labials. The fourth dimension is evidently related to the voicing distinction, with positive values indicating voiceless consonants and negative values indicating voiced consonants. While the third dimension was not easily interpreted from the stimulus coordinates, it became meaningful when plotted against the first dimension. From this perspective, the voiced and voiceless alternative for each place and manner feature cluster together. The relative proximity of the voiced and voiceless members of each pair indicated by these dimensions might also reflect differences in relative similarity between pairs.

4.2.7 Multiple-regression analysis

The multi-dimensional scaling analysis described in the previous section takes into account the asymmetries in the speech onset latency data, but the resulting stimulus configurations do not clearly illustrate these asymmetries. In a further attempt to understand this aspect of the results, the data were subjected to multiple-regression analysis.

This method evaluates how well the data can be modeled by a linear combination of the variables suggested by the analyses presented so far, and determines the relative importance of each. Thus, it was possible to investigate the contribution of the feature specifications of both the prime and the target independently on the TSOL data. TSOLs corresponding to the identity conditions were eliminated from this analysis. For this analysis, fricatives, alveolars, and voiced consonants were coded with a 'one' while stops, labials, and voiceless consonants were coded with a 'zero'. Table 3 summarizes the results of the multiple-regression analysis using the voicing, place, and manner features of the prime and target stimuli to predict the latency data.

Factor	Standardized Coefficient	Significance
R = .51, F(6, 49)		.016*
Voicing of prime	.021	.864
Place of prime	-.347	.007**
Manner of prime	-.225	.076
Voicing of target	.149	.235
Place of target	.103	.408
Manner of target	-.258	.042*

Table 3. Summary of multiple-regression analysis.

With 26% of the variance accounted for by the regression, the significant predictors of the latency data are the place feature of the prime and the manner feature of the target. The manner feature of the prime was also marginally significant, and should not be ignored. The results of this analysis, then, are consistent with the previous analyses in that the relevant factors continue to be place and manner features, while voicing plays no significant role in the underlying pattern of the data.

4.3 General discussion

One aim of these experiments was to further understand the nature of the feature interaction found in a previous series of gating experiments. One hypothesis tested was that the facilitation for the Place group in the gating experiments might be accounted for by phonetic priming of the place feature. Results from the priming experiments presented here lead to a rejection of that hypothesis. Speech onset latencies for prime-target pairs which shared place of articulation were not significantly different from those for pairs not sharing place. A second hypothesis tested was that the blocking procedure used in the gating experiments may have facilitated responses for the Place group by providing the subjects with prior knowledge about the place feature of the possible responses. This hypothesis is also rejected on the basis of the priming experiment results. The same blocking procedure was used in the priming experiments, but no effect for group was found. That is, speech onset latencies for subjects whose response set was blocked by the place feature were not statistically different from subjects whose response set was blocked by voicing or place, or from subjects whose response set was not blocked at all.

The broader research questions this study sought to answer were the following: 1) Is there any correlation between the number of shared features (0-3) and the existence or extent of phonetic priming? 2) Are responses to the target stimuli dependent upon the type of featural similarity between the prime and the target? If so, what is the effect of sharing a particular feature (such as place) or a particular feature specification (such as labial)? The data were analyzed in three distinct ways using analyses of variance of planned comparisons, multi-dimensional scaling, and multiple-regression. Results from these analyses yielded robust and consistent findings which have serious implications for theories concerning the form, content, and organization of the internal representations of these phonemes. These implications will be discussed in the following sections, each of which covers one type of analysis.

4.3.1 Featural similarity conditions

Results from the analyses of featural similarity conditions show that speech onset latencies are not systematically affected in a linear fashion by the number of shared phonetic features, but are affected by the type of featural similarity. While it is true that latencies for the two-feature shared condition were significantly longer than the three-, one-, and no-feature shared conditions, it is not the case that the more features shared, the longer the latencies. As far as which type of featural similarity corresponds to differences in speech onset latencies, at first glance it appears that shared manner is the only important factor. More detailed analyses of the prime-target relationship in terms of specific features shared and direction of featural change, however, indicate that the above conclusion is premature. Upon closer inspection, it becomes apparent that there are the following asymmetries in data: speech onset latencies are significantly longer for prime-target pairs of shared stop consonants and prime-target pairs with a change of place feature in the direction of alveolars to labials. There is also an important interaction of these two influential factors.

The finding of an inhibitory rather than facilitative effect in the present study is unexpected in the face of previous priming studies, but is consistent with previous studies examining the role of phonetic feature overlap in speech production tasks (e.g. Meyer and Gordon, 1985, Sevald and Dell, 1994, Rogers and Storkel, 1998). This effect is generally interpreted as evidence for the sequential programming of phonemes in successive utterances.

The effect of shared manner reported here is consistent with the findings of Rogers and Storkel (1998) who investigated the effects of phonological similarity on speech production latencies in a naming task. Subjects were asked to name visually presented monosyllabic words as fast as they appeared on a screen. The words contrasted only in the initial consonant, and the type of phonological similarity was carefully controlled to create pairs with the following features shared: voicing and manner, place and manner, voicing only, manner only, no features shared. Two aspects of the results are consistent with the present findings. First, the number of features shared did not correlate with the observed differences across similarity conditions. Second, the results indicated that shared manner most frequently yielded inhibition in the form of longer speech onset latencies. The study differed from the present one in that the two manner classes included in the study were stops and affricates, and differences in latencies within the shared manner category were not explored.

The lack of any significant difference between the identity condition and the unrelated condition in the present study is inconsistent with previously reported accounts of identity priming in a shadowing task (e.g. Radeau, Morais, and Dewier, 1989) but this may be due to statistical characteristics of the stimulus set. As discussed earlier, the facilitation found by Goldinger, Luce, Pisoni, and Marcario (1992) for phonetically close stimulus pairs was greatly reduced when the percentage of related pairs in the stimulus set was reduced from 50% to 10%. In the present study, the identity condition makes up only 12% of the stimulus set. Subjects' unconscious awareness of this factor, therefore, may have led them to a strategy of expecting the target syllable to differ from that of the prime. Also, it should be noted that most reported accounts of identity priming have used prime-target pairs that were visually or acoustically identical whereas the present study used different acoustic exemplars of the same syllable spoken by the same speaker.

4.3.2 Multi-dimensional scaling

The multi-dimensional scaling analysis of the priming data provides a geometric representation of the underlying patterns in the data. The three-dimensional stimulus configuration resulting from this analysis shows that when speech onset latencies are interpreted as a reflection of similarity of phoneme representations, then the first dimension along which the data pattern corresponds to the manner distinction. Thus the primary categorization of the eight consonants included in the study is that of stops and fricatives. The second and third dimensions which emerge from the analysis together serve to categorize the consonants according to the voicing and place features. There is an interesting asymmetry in the spatial representation of the place distinction which may reflect the asymmetrical patterns in the data.

The general findings from this analysis are consistent with results of the same analysis applied to the confusion matrices from the gating experiments described earlier. The underlying patterns in the data were shown to correspond to a four dimensional solution in which the first dimension again corresponded to the manner distinction. The second and fourth dimensions coincided clearly with the place and voicing distinctions, respectively, and the third dimension (when plotted against the first) serves to group the voiced and voiceless alternatives for each place and manner specification.

While there is remarkable consistency between the scaling analyses of the two different data sets, it should be noted that these findings are not consistent with previous analyses of speech sounds. Shepard (1972) was the first to use non-metric multi-dimensional scaling to examine the psychological representation of speech sounds by analyzing the confusion matrices from Miller and Nicely (1955). The matrices included in his analysis consisted of only those in which bandwidth was 200 to 6500 Hz and in which the only manipulation consisted of differences in the signal-to-noise ratio. The first dimension on which the data patterned was one of nasality versus voicing (voiced and voiceless) and the second dimension was that of voicing. Based on independent analyses of each of the six matrices, Shepard concluded that although the variations in the signal-to-noise ratio affected the *number* of confusions, the internal *pattern* of confusions remained stable. The apparent inconsistency between these findings and my own can be explained by differences in the stimulus set characteristics. The matrices analyzed by Shepard corresponded to confusions of consonants embedded in noise. The process of embedding the consonants in noise differentially affected the acoustic cues for the different features. Nasality and voicing cues were less masked by this process than place or durational cues.

4.3.3 Multiple-regression

Results of the multiple-regression analysis show that the speech onset latency data can be predicted fairly well by a linear combination of certain variables. The most important variables were shown to be the place feature of the prime stimulus and the manner feature of the target stimulus, with the manner feature of the prime also playing a minor role. The advantage of multiple-regression lies in this method's ability to clarify the asymmetrical nature of the data. These asymmetries have implications for speculations about the nature of the internal representations of these phonemes. The longer latencies for targets following alveolar primes (especially for stop targets following alveolar primes) indicate a difference in representation for alveolars versus labials, as well as a difference between stops and fricatives. This difference may reflect differences with respect to the internal organization of these categories, and does not imply that the category distinctions are any less robust.

5 Conclusions

The three different types of analysis performed on the latency data yielded consistent results in terms of identifying the relevant parameters and their relative importance. Each analysis is important to an understanding of the data because each provides a different insight into the processes involved in the experimental task. The analyses of variance for the planned comparisons showed that a main effect for number and type of shared feature, and shared manner emerged as the most influential factor. This result is consistent with previous findings in speech production. Subsequent ANOVAs on more detailed feature categorizations of the stimuli, however, revealed dissimilar influences of alveolar versus labial primes and brought to light the fact that the shared manner effect primarily obtained for shared stop pairs. The multi-dimensional scaling analysis provided a geometric representation of the underlying patterns in the latency data. For this method, the speech onset latencies were interpreted as measures of similarity which then correlate with relative proximity of phoneme representations. The resulting three-dimensional stimulus configuration derived from this analysis revealed manner as the most important dimension, and place and voicing as a combination of the second and third dimensions. The distinct patterns of spatial representation for alveolars and labials may also reflect the corresponding asymmetries in the data. The precise nature of these asymmetries was further illuminated by the multiple-regression analysis. This analysis showed that the data could be predicted fairly well by a linear combination of the same variables which emerged as important in the other analyses. In addition, however, this analysis revealed that the manner specification of the target is a much more important predictor than the manner specification of the prime, and the place specification of the prime is the most important factor of all. These results indicate that there is inhibitory phonetic priming of specific place and manner features but that the influences of the prime and target features are more complex than simple shared activation of a particular feature. Furthermore, although manner appears to be the primary distinction for categorization of these eight consonants and voicing seems to be the tertiary distinction, the asymmetries associated with the place distinction make it difficult to model the data with a simple geometric representation.

6 References

Andruski, J. E., Blumstein, S. E., and Burton, M. (1994) The effect of subphonetic differences on lexical access. *Cognition* 52, 163-187.

Banks, W. P., Roberts, D., and Ciranni, M. (1995) Negative priming in auditory attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21 (6), 1354-1361.

Bowey, J. (1996) Phonological recoding of nonword orthographic rime primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (1), 117-131.

Burton, M. W., Jongman, A., and Sereno, J. A. (1997) Phonological and orthographic priming effects in auditory and visual word recognition. (unpublished manuscript).

Connine, C. M., Blasko, D. G., and Wang, J. (1994) Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception and Psychophysics,* 56 (6), 624-636.

Corina, D. (1992) Syllable priming and lexical representations: Evidence from experiments and simulations. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pp. 779-784. Bloomington: Indiana University Press.

Delgutte, B. (1997) Auditory Neural Processing of Speech. In W. J. Hardcastle and J. Laver (Eds.) *The Handbook of Phonetic Sciences,* pp. 507-538. Blackwell Publishers Ltd.: Cambridge, MA.

Delgutte, B., and Kiang, N. Y. S., (1984) Speech coding in the auditory nerve: IV. Sounds with consonant like dynamic characteristics. *Journal of the Acoustical Society of America*, 75 (3), 897-907.

Doeleman, T. L. (1998). The perception of features in the identification of English consonants. In N. Adisasmito-Smith and W. Ham (Eds.), *Working Papers of the Cornell Phonetics Laboratory, 12,* 139-193.

Ferrand, L., Segui, J., and Grainger, J. (1995) Masked phonological priming in picture naming. *L'année psychologique,* 95, 645-659.

Gaskell, M. G., and Marslen-Wilson, W. D. (1996) Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 22(1), 144-158.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., and Marcario, J. K. (1992) Form-based priming in spoken word recognition: the roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (6), 1211-1238.

Grainger, J. and Ferrand, L. (1996) Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, 35, 623-647.

Hillinger, M. L. (1980) Priming effects with phonemically similar words: The encoding-bias hypothesis reconsidered. *Memory and Cognition*, 8 (2), 115-123.

Jakimik, J., Cole, R. A., and Rudnicky, A. I. (1985) Sound and spelling in spoken word recognition. *Journal of Memory and Language,* 24, 165-178.

Lukatela, G., and Turvey, M. T. (1990) Phonemic similarity effects and prelexical phonology. *Memory and Cognition*, 18 (2), 128-152.

Lukatela, G., and Turvey, M. T. (1991) Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance,* 17 (4), 951-966.

Lukatela, G., and Turvey, M. T. (1993) Similar attentional, frequency, and associative effects for pseudohomophones and words. *Journal of Experimental Psychology: Human Perception and Performance,* 19 (1), 166-178.

Marslen-Wilson, W. D., Nix, A., and Gaskell, G. (1995) Phonological variation in lexical access: Abstractness, inference and English place assimilation. *Language and Cognitive Processes*, 10, 285-308.

Martin, R. C., and Jensen, C. R. (1988) Phonological priming in the lexical decision task: A failure to replicate. *Memory and Cognition*, 16 (6), 505-521.

Meyer, D. E., and Gordon, P. C. (1985) Speech production: Motor programming of phonetic features. *Journal of Memory and Language,* 24, 3-26.

Milberg, W., Blumstein, S., and Dworetzky, B. (1988) Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, 26 (4), 305-308.

Miller, G. A. and Nicely, P. E. (1955) An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* 27, 623-638.

Praamstra, P., Meyer, A. S., and Levelt, W. J. M. (1994) Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience,* 6 (3), 204-219.

Radeau, M., Morais, J., and Dewier, A. (1989) Phonological priming in spoken word recognition: Task effects. *Memory and Cognition,* 17 (5), 525-535.

Rogers, M. A., and Storkel, H. L. (1998) Reprogramming phonologically similar utterances: The role of phonetic features in pre-motor encoding. *Journal of Speech, Language, and Hearing Research* 41, 258-274.

Rossi, R., Burton, M., Jongman, A., and Sereno, J. (1993) Coarticulatory information in word recognition: A phonetic priming study. *Journal of the Acoustical Society of America,* 93 (4), 2393.

Sevald, C. A., and Dell, G. S. (1994) The sequential cueing effect in speech production. *Cognition,* 53, 91-127.

Shepard, R. N. (1974) Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 39 (4), 373-421.

Slowiaczek, L. M., and Hamburger, M. (1992) Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (6), 1239-1250.

Slowiaczek, L. M., Nusbaum, H. C., and Pisoni, D. B. (1987) Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (1), 64-75.

Slowiaczek, L. M., and Pisoni, D. B. (1986) Effects of phonological similarity on priming in auditory lexical decision. *Memory and Cognition*, 14 (3), 230-237.