

The Effect of the Temporal Structure of Spoken Words on Paired-Associate Learning

Sarah C. Creel
University of California, San Diego

Delphine Dahan
University of Pennsylvania

In a series of experiments, participants learned to associate black-and-white shapes with nonsense spoken labels (e.g., “joop”). When tested on their recognition memory, participants falsely recognized as correct a shape paired with a label that began with the same sounds as the shape’s original label (onset-overlapping lure; e.g., *joob*) more often than a shape paired with a label that overlapped with the original label at offset (offset-overlapping lure; e.g., *choop*). Furthermore, the false-alarm rate was modulated by the phonetic distance between the sounds that distinguished the original label and the lures. Greater false-alarm rates to onset-overlapping labels were not predicted by explicit similarity ratings or by consonant identification and were not dependent upon label familiarity. The asymmetry at erroneously recognizing onset- versus offset-overlapping lures remained unchanged as the presentation of the shape at test was delayed in time, suggesting that response anticipation based on the first sounds of the spoken label did not contribute much to the false recognition of onset-overlapping lures. Thus, learning 2 words whose names differ in their last sounds appears to pose greater difficulty than learning 2 words whose names differ in their first sounds because, we argue, people are biased to give more importance to the early sounds of a name than to its last sounds when learning a novel label–referent association.

Keywords: memory recognition, spoken-word recognition, temporal structure

Spoken words are perceptual events that extend over time. Seminal studies by Marslen-Wilson and colleagues (e.g., Marslen-Wilson, 1984; Marslen-Wilson & Welsh, 1978), and the substantial body of work that ensued, have established that the recognition of spoken words from memory is initiated as soon as some phonetic information from the early portion of the word becomes available to the listener: Listeners use this information immediately to generate hypotheses about which word they may be hearing. Marslen-Wilson’s original “cohort” theory gave the early portion of a spoken word a privileged status, compared with the rest of the word form, by assuming that the first sounds of a spoken stimulus contribute to establishing a set of word candidates, a *cohort*. The cohort is then progressively pruned of those candidates that become inconsistent with new phonetic information until only one remains, at which point recognition is achieved.

A deficiency of the theory that was soon recognized is that it predicted that listeners could not recognize words whose initial sound was distorted or missing. Numerous studies have demonstrated that this prediction was incorrect (Connine, Blasko, & Titone, 1993; Norris, 1982; Salasoo & Pisoni, 1985). This led the

field to embrace “continuous-mapping” models, in which any portion of a spoken stimulus can contribute to recognition, with no built-in privilege assigned to the early part of the stimulus (Allopenna, Magnuson, & Tanenhaus, 1998; Connine et al., 1993). Nonetheless, in many of the current models, the advantage that word candidates matching the first sounds of a spoken word have over those candidates matching later sounds has been maintained because of the models’ internal dynamics. As the spoken stimulus unfolds over time, words that match it immediately accrue evidence supporting them and compete with alternatives in proportion to the strength of this evidence. The earlier in the spoken stimulus that words begin to match, the more strongly they can compete with words that match the input later.

Thus, the greater contribution of early (as opposed to late) phonetic information to the recognition of familiar words is predicted by all current theories and has been empirically confirmed under a variety of conditions. For instance, Cole and Jakimik (1980) exposed listeners to a prerecorded story in which some words had been mispronounced and asked people to detect the mispronunciations. Cole and Jakimik reported faster mispronunciation detections when the mispronunciation affected the second syllable of a word than when it affected the first syllable, suggesting that listeners had a greater ability to retrieve the original word from the distorted stimulus when the early part of the word was intact. Using a priming methodology, Monsell and Hirsh (1998) showed that the first sounds of a spoken word elicit the retrieval from memory of all known words that begin with these sounds. The last sounds of a spoken word, however, elicit no similar process. In line with the Monsell and Hirsh study, Vitevitch (2002) and Benkí (2003) reported evidence of a greater influence of onset-overlapping competitors on the recognition of a spoken word. For example, Benkí presented monosyllabic consonant–

Sarah C. Creel, Department of Cognitive Science, University of California, San Diego; Delphine Dahan, Department of Psychology, University of Pennsylvania.

This work was supported by National Science Foundation (NSF) Grant 0433567 and National Institutes of Health Grant 1-R01-HD-049742-1 to Delphine Dahan and by NSF Integrative Graduate Education and Research Traineeship Grant 0504487.

Correspondence concerning this article should be addressed to Sarah C. Creel, Department of Cognitive Science 0515, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093. E-mail: creel@cogsci.ucsd.edu

vowel–consonant words in noise and found that the density of the onset neighborhood (i.e., the number and frequency of the words that share the initial consonant and vowel with the spoken item) contributed to predicting the probability of correctly identifying the spoken words, whereas the density of the rhyming neighborhood (i.e., the number and frequency of the words that rhyme with the spoken item) did not.

A study by Allopenna et al. (1998) illustrates how the differential impact of onset- and offset overlapping competitors arises during the recognition of a spoken word. Allopenna et al. asked participants to click on one of four pictures as their eye movements to the pictures were monitored. Analysis of participants' gaze revealed that in the course of hearing the name of the referent picture (e.g., "candle"), people were more likely to fixate on a picture with a name that started with the same sounds (e.g., *candy*) than on a picture with a name that rhymed with the referent's name (e.g., *handle*). The timing of participants' fixations to the onset-overlapping competitors suggests that these fixations were largely programmed while the early, ambiguous portion of the spoken word was being heard. During this brief interval, an onset-overlapping competitor is as strong a contender for recognition as the actual word. An offset-overlapping competitor, on the other hand, never achieves this status because the portion that differentiates it from the actual word is immediately available.

Influences of Temporal Structure During Word Learning

As this review of past work illustrates, most of the work pertaining to the influence of spoken words' temporal organization on their recognition has focused on highly practiced words. The present study explored how the temporal organization of spoken words affects how words and their referents are learned and remembered. We evaluated the hypothesis that when learning a new word, that is, a name–referent association, people give more weight to the name's early sounds than to its later sounds.

Variation in the weight given to different portions of a spoken word may result from a variety of factors. First, people may have a general propensity to especially attend to and remember the beginning of a sequence. Primacy effects in the free or serial recall of a list of items—a greater probability of recalling the first item of a list than items a few positions later—would be an illustration of such a propensity. Second, and in the realm of word learning, people may attend to different parts of a word as a function of their information value in distinguishing words within a language's vocabulary. Indeed, statistical analyses of the English lexicon have revealed that some positions convey more information than others. For example, Kessler and Treiman (1997) and De Cara and Goswami (2002) have shown that English monosyllabic words tend to share their vowel and final (coda) consonants with more words than they share their initial (onset) consonants and vowel with. There are also stronger statistical dependencies between the vowel and its subsequent coda consonant than between the onset consonant and the subsequent vowel, even when only monomorphemic words are considered (see also Miller, 1951, for an early report on this asymmetry). Thus, greater attention to onset consonants than to coda consonants when recognizing familiar English monosyllabic words would reflect the greater entropy associated with onset consonants. Regardless of the origin of the asymmetry favoring

words' initial sounds, the question of interest here is whether the onset–coda asymmetry extends beyond the online processing seen during recognition of familiar words, to the learning of novel name–referent associations.

Existing research on word learning and language acquisition does not provide unequivocal support for a stronger role for early information in word learning. Past work has examined the influence of word form characteristics on label–referent association learning. For example, Storkel and colleagues (Storkel, 2001; Storkel, Armbrüster, & Hogan, 2006) have documented how the similarity of novel labels to other known words and their phonological structure influence adults' and children's ability to accurately remember the novel labels. This finding is important, as it provides support for the influence of the learner's language vocabulary and its phonological structure on the learning of new words. Metsala and Walley (1998) have proposed that the similarity between the labels that children have learned for different concepts causes the representation of the labels to change and gain phonemic specificity (but see Swingley, 2003, 2009). However, whether some portions of a new label are better retained in memory than others is not directly addressed.

Phonologists have long noticed that children produce the first consonant of monosyllabic words more often and more accurately than their last consonants (e.g., Stoel-Gammon & Dunn, 1985). However, this asymmetry may arise from many different factors, including articulatory limitations. For example, 1-year-olds, who are much more likely to omit or misarticulate a coda consonant than an onset consonant, nevertheless show a similar degree of disruption in familiar-word recognition when hearing a monosyllabic word whose onset or coda has been substituted by a similar consonant (Swingley, 2009), suggesting that the difference in likelihood of correct articulation is not necessarily due to a difference in the children's knowledge of the onset versus coda of familiar words. Research by Nazzi and colleagues (Nazzi, 2005; Nazzi & Bertoncini, in press; Nazzi, Floccia, Moquet, & Butler, 2009), which specifically examines toddlers' learning of novel label–object associations, has revealed intriguing differences in children's attention to consonantal versus vocalic contrasts in the process of learning category names. However, there is no evidence for an asymmetry in 20-month-old children's encoding or retrieval of consonant information as a function of the consonant's syllabic position, at onset or coda.

Another approach to understanding how the temporal aspect of spoken words affects their memorization and recognition has been to assess the similarity structure among these words. This research, conducted on both adults and young children, has consistently found that monosyllabic words or nonsense stimuli that rhyme, thereby sharing their last sounds, are judged to sound more similar to one another than stimuli that begin with the same sounds (e.g., Hahn & Bailey, 2005; Nelson & Nelson, 1970). For example, Storkel (2002) evaluated the frequency with which preschool children considered a monosyllabic word (the test word) to sound "like" another one (the standard word). Storkel's results showed that children thought a test word sounded like the standard more often when these two words were identical or phonetically similar in their coda consonants than when they were identical or phonetically similar in their onset consonants. Thus, similarity judgments suggest that greater weight is given to the late portion of the spoken word than to its earlier portion, in opposition to what the

literature on word recognition indicates. Hahn and Bailey (2005) attributed this discrepancy to the tasks used in different studies. If the task allows, or even encourages, listeners to anticipate the identity of the word they hear, the information received earlier in time plays a disproportionate role. This asymmetry lessens, and eventually reverses, as the task relies less on listeners' incremental processing of spoken words, as observed in similarity judgments given to two stimuli that are presented sequentially.

Thus, similarity judgments given by adults on nonsense syllables, or by preschoolers on words recently acquired and whose lexicon is still developing, predict that newly learned words should be more confusable with one another when they overlap at offset (i.e., sharing the same rime) than when they overlap at onset (i.e., sharing the same onset consonant and vowel). We note that this prediction assumes that the confusability and similarity are the two sides of the same coin, a point we return to in the General Discussion.

In the current study, we examined potential word position asymmetries for adults learning novel referent–name associations. This study builds on a growing body of work in which a set of novel labels for novel objects is taught to participants (e.g., Creel, Aslin, & Tanenhaus, 2006; Magnuson, Tanenhaus, Aslin, & Dahan, 2003; Shatzman & McQueen, 2006). Magnuson et al. (2003) trained participants to associate novel labels to nonsense shapes and later tested their ability to correctly select a labeled shape among four alternatives. Participants' accuracy in selecting the correct referent for each label was very high. However, the concurrent monitoring of participants' eye gaze provided a measure of temporary confusion between a given label and the names of the incorrect shapes. Eye gaze analyses indicated that participants were more likely to briefly consider a shape with a name that began with the same sounds as the current label they heard than a shape with a name that ended with the same sounds. This result, the authors argued, arose because of the incremental evaluation of the speech signal, as found with highly practiced words. However, it is also possible that the effect was in part driven by the tendency to confuse the referents of onset-overlapping labels more than those of offset-overlapping labels. In other words, it is difficult to assess whether the confusion was local, based on the first portion of the spoken label only, or global, based on the nature of the word–referent association itself.

Following up on the Magnuson et al. (2003) study, Creel et al. (2006) evaluated how the temporal order and phonological structure of labels may affect word learning. Participants were taught a large set of label–shape associations and tested in a four-alternative forced-choice task. By contrast with Magnuson et al., the amount of training was limited, and participants frequently chose an incorrect referent for the spoken label they heard. Errors provided a measure of the confusability between the label and the name of the erroneously chosen shape, conferred by their phonological similarity. Analyses of participants' errors revealed a greater confusion between onset-overlapping labels (e.g., *bamo* and *bami*) than between offset-overlapping labels (e.g., *goti* vs. *poti*). This first demonstration was followed by a series of experiments in which the status of the overlapping segments (consonants vs. vowels) and their position were varied. Overall, consonant overlap led to greater confusion than vowel overlap, unless vowels were the first unique sound of each syllable of the labels.

Thus, the syllable position of the speech sounds shared by two labels appeared to modulate their confusability.

However, as in the Magnuson et al. (2003) study, it is difficult to distinguish between two possible interpretations: People erroneously may choose a shape with a name that overlaps with the spoken label at onset more often than one that overlaps at offset because people sometimes commit to a response based primarily on the early sounds of the label. Alternatively, or additionally, learning to distinguish onset-overlapping labels and their respective referents may be harder than learning to distinguish offset-overlapping labels because people have a bias to attend to the early portion of spoken words. Thus, people's tendency to erroneously choose a shape with an onset-overlapping name as the referent of a spoken label may reflect the transient influence of incremental processing of the speech signal, but may also reflect a long-lasting influence of the temporal organization of speech on word learning. Without a way to constrain how much of the spoken label people have heard before they can start to prepare their response, it is difficult to decide between these two accounts. The present study did just that.

We extended the Creel et al. (2006) study in several important ways. First, we examined confusability among novel labels that differed by a single consonant and compared confusability between words that shared their first sounds but differed on the last consonant (e.g., *joop* [dʒʊp] and *joob* [dʒʊb]) and words that shared their last sounds but differed on the first consonant (e.g., *joop* [dʒʊp] and *choop* [tʃʊp]). This let us test the effect of location of overlap without confounding location with sound type. Second, we varied the phonetic distance between the consonants that differed between the two overlapping labels in order to directly assess the impact of similarity on confusions and its possible interaction with the location of overlap. Finally, we diverged from Creel et al. by assessing people's memory for the taught name–referent associations using a recognition-memory task: Participants saw one of the trained shapes and heard a label; their task was to decide whether the name–shape pair was the same as one of those presented during training (i.e., whether the shape was named correctly). We examined whether interference, as measured by the false-alarm rate on rearranged label–picture pairs, is stronger when the original label and the lure overlapped in their initial sounds than when they overlapped in their final sounds. That is, if a listener who has learned that a given shape is a *joop* hears the same shape labeled “joob”—the label associated with a different shape during training—will the listener be more likely to falsely recognize this as a correct pairing than if the listener heard that shape labeled “choop”? It is important to note that by manipulating when the shape appeared with respect to the unfolding of the spoken label, we were able to manipulate how much of the spoken label people had heard before they could evaluate its association with the shape, thereby examining the degree to which response anticipation may account for an asymmetry in the confusions between onset- and offset-overlapping labels.

Relationship of Similarity to Recognition Memory

There is a substantial literature on the effect of similarity on the false-alarm rate in recognition memory (Dyne, Humphreys, Bain, & Pike, 1990). Similarity affects recognition memory in at least two ways. First, an unstudied item is more likely to be incorrectly

recognized as part of a memorized list as its similarity or relatedness to items in the study list increases. Second, the influence of similarity on false recognition has been examined in studies requiring participants to learn paired associates. Participants are more likely to erroneously recognize a rearranged pair as “old” if the substituted item and the original one resemble each other. For example, studies of face–name associations have shown that the probability of incorrectly recognizing a rearranged face–name lure increased as the similarity between the lure’s face and the original face associated with the lure’s name increased (Pantelis, van Vugt, Sekuler, Wilson, & Kahana, 2008). Although the memory processes that give rise to this effect are under debate (e.g., Malmberg & Xu, 2007; Verde, 2004), for our current purposes, it demonstrates a relationship between perceived similarity between pairs and false-alarm rate: As similarity increases, the false-alarm rate increases. In the present study, we examined whether the position of the overlap between the label originally associated with a shape at study and the label presented at test affects the false-alarm rate.

Work by Nelson and colleagues (Nelson & Borden, 1973; Nelson & Brooks, 1973; Wheeler & Nelson, 1982) provides some evidence that the sequential order of the features shared by words in studied pairs may affect associative interference. In these studies, participants studied a list of six word–digit pairs in which the six-letter words all had letters in common, either the first two, the middle two, or the last two. The probability of erroneously recognizing a rearranged pair as old was highest when the six-letter words all shared their first two letters. This suggests that the similarity among the words, which caused associative interference, is greatest when the overlap among the confused items concerns their initial portion. Although the words were presented visually, they may have been encoded as phonological stimuli with a temporal organization. Thus, although only suggestive, the results are consistent with greater interference from shared elements among pairs in which these elements occur early in the words than when they occur late.

In the present study, we examined whether the false-alarm rate on rearranged label–picture pairs is stronger when the original label and the lure overlap in their initial sounds than when they overlap in their final sounds. In addition to manipulating the location of overlap between the original label and the lure, we assessed the impact of their similarity at each overlap location by varying the phonological distance between the sounds by which the two items differed. The two sounds were perceptually close, that is, sharing most of their phonological features (e.g., *j* and *ch*, as in *joop* or *choop*), or perceptually distant, that is, sharing fewer features (e.g., *z* and *p*, as in *zutch* and *putch*). This manipulation allowed us to concurrently assess the effects of featural distance, in phonetic space, and location of overlap on false-alarm rate, as well as any interaction between these factors.

In Experiment 1, we first examined the effect of overlap position on false-alarm rate to rearranged pairs. In the three following experiments, we evaluated various explanations of the results of Experiment 1. In Experiment 2, we assessed the clarity of the labels used in Experiment 1, and the perceived similarity between them, when no paired-associate learning is required, using a transcription task and explicit similarity judgments. In Experiment 3, we compared false-alarm rates to rearranged pairs and pairs with new labels, which had not been studied during training. Finally, In Experiment 4, we examined the role of response anticipation in this task.

Experiment 1

Method

Participants. Eighteen students from the University of Pennsylvania participated in the experiment for course credit. None reported a history of hearing problems, and all identified themselves as native speakers of English. The data from two participants were subsequently excluded because they revealed poor learning (see below).

Stimuli. The verbal stimuli consisted of 32 consonant–vowel–consonant items. None of the items were words of English, with the possible exceptions of *sav* (i.e., *salve*, a low-frequency word, according to Kučera and Francis’s, 1967, word count) and *puđ* (a slang term that does not occur in Kučera and Francis). These items were constructed from a set of eight onset consonants (/b, p, ɖʒ, ʃ, k, g, s, z/), eight vowels (/æ, aɪ, aʊ, u, i, ʊ, ɔɪ, ʌ/), and eight coda consonants (/b, p, ɖʒ, ʃ, t, d, f, v/), each of which occurred equally often across the set (see Table 1). The 32 items consisted of eight groups of four items in which two onset consonants, one vowel, and two coda consonants were used (/bæf/, /bæv/, /sæf/, /sæv/). Within each set of four items, the two onset consonants were phonetically close, differing by a single phonetic feature (which was always voicing; e.g., /git/ vs. /kit/), or phonetically distant, differing by three features (voicing, place of articulation, and manner of articulation; e.g., /bæv/ vs. /sæv/). The same was true of the two coda consonants. Each item in the set (e.g., /bæv/) overlapped with another on its first two segments (/bæf/) and with another on its last two segments (/sæv/). The phonetic similarity of the onset and coda consonants within a set was varied across the eight sets. The use of four-item sets enabled us to have a given label play the role of both onset- and offset-overlapping labels, thus controlling for the influence that a label’s similarity with existing English words may have had on the learning of this label. For example, the label /bæv/ was an onset-overlapping competitor when the original label was /bæf/ but an offset-overlapping competitor when the original label was /sæv/.

The 32 items were read by a female native speaker of English from western Pennsylvania in random orders and recorded on a computer at a sampling rate of 22,050 Hz. Average item duration was 596 ms. Tokens were selected that were uniform in terms of prosody and did not contain anomalous noise artifacts. After extraction, the tokens were normalized to 70 dB SPL with Praat signal processing software (Boersma & Weenink, 2006).

The visual stimuli to serve as the referents of the items were a set of 32 black-and-white shapes (see Figure 1 for examples). There were four assignments of pictures to items, distributed equally across participants, and each assignment ensured that the referents of similar-sounding names did not look particularly similar.

Procedure. Participants were tested individually. They sat at a comfortable distance from a computer monitor and wore headphones adjusted to a comfortable volume. Stimulus presentation was controlled by DMDX software (Forster & Forster, 2003). The training phase consisted of a series of trials in which a shape appeared on the computer monitor concurrently to the presentation of what was described to participants as its name. Participants were asked to learn the names of the shapes. To ensure participants’ sustained attention during training, we also asked participants to judge whether the shape appeared on the left or right side of the

Table 1
Vocabulary Used in Experiments 1–4

Item (IPA gloss) ^a	Rearranged lures				New lures ^b				Pronunciation notes
	Onset-overlapping competitor		Offset-overlapping competitor		Onset-overlapping competitor		Offset-overlapping competitor		
	Close	Distant	Close	Distant	Close	Distant	Close	Distant	
chowb (/tʃaʊb/)	tʃaʊp		ɔʃaʊb			tʃaʊge		saʊb	Vowel as in <i>out</i>
chowp (/tʃaʊp/)	tʃaʊb		ɔʃaʊp	tʃaʊt				kaʊp	
jowb (/dʒaʊb/)	dʒaʊp		tʃaʊb		dʒaʊd			kaʊb	
jowp (/dʒaʊp/)	dʒaʊb		tʃaʊp			dʒaʊtʃ		gaʊp	
choob (/tʃʊp/)	tʃʊp		ɔʃʊb			tʃʊf		kʊb	
choop (/tʃʊp/)	tʃʊb		ɔʃʊp		tʃʊt			zʊp	
joob (/dʒʊb/)	dʒʊp		tʃʊb			dʒʊf		sub	
joop (/dʒʊp/)	dʒʊb		tʃʊp			dʒʊv		gʊp	
geege (/gɪdʒ/)		gɪt	kɪdʒ			gɪf	bɪdʒ		Word-initial <i>g</i> is a hard <i>g</i>
geet (/gɪt/)		gɪdʒ	kit		gɪp			ɔʃɪt	
keege (/kɪdʒ/)		kit	gɪdʒ		kɪtʃ		pɪdʒ		
keet (/kɪt/)		kɪdʒ	gɪt			kɪf		ɔʃɪt	
goodge (/gʊdʒ/)		gʊt	kʊdʒ		gʊtʃ			tʃʊdʒ	Vowel as in <i>hood</i>
goot (/gʊt/)		gʊdʒ	kʊt		gʊp		bʊt		
koodge (/kʊdʒ/)		kʊt	gʊdʒ			kʊv	pʊdʒ		
koot (/kʊt/)		kʊdʒ	gʊt		kʊp			ɔʃʊt	
baf (/bæf/)	bæv			sæf		bæp	pæf		
bav (/bæv/)	bæf			sæv		bæb	gæv		
saf (/sæf/)	sæv			bæf		sæb	zæf		
sav (/sæv/)	sæf			bæv		sætʃ	zæv		
bife (/baɪf/)	baɪv			saɪf		baɪb	paɪf		
bive (/baɪv/)	baɪf			saɪv		baɪp	paɪv		
sife (/saɪf/)	saɪv			baɪf		saɪp	zaɪf		
sive (/saɪv/)	saɪf			baɪv		saɪb	zaɪv		
poɪtʃ (/pɔɪtʃ/)		pɔɪd	zɔɪtʃ		pɔɪge		kɔɪtʃ		
poɪd (/pɔɪd/)		pɔɪtʃ	zɔɪd		pɔɪb			tʃɔɪd	
zoɪtʃ (/zɔɪtʃ/)		zɔɪd	pɔɪtʃ		zɔɪtʃ			ɔʃɔɪtʃ	
zoɪd (/zɔɪd/)		zɔɪtʃ	pɔɪd		zɔɪt		sɔɪd		
putʃ (/pʌtʃ/)		pʌd	zʌtʃ			pʌv	bʌtʃ		Vowel as in <i>hut</i>
pud (/pʌd/)		pʌtʃ	zʌd			pʌv		tʃʌd	
zutʃ (/zʌtʃ/)		zʌd	pʌtʃ		zʌdʒ			gʌtʃ	
zud (/zʌd/)		zʌtʃ	pʌd		zʌb		sʌd		

Note. Data indicate various incorrect labels that the word’s shape was paired with at test.
^a Novel word plus its gloss in the International Phonetic Alphabet (IPA). ^b New lures apply to Experiment 3.

screen and to indicate their response by pressing the left or right arrow on the computer’s number keypad. The training phase consisted of 16 blocks of 32 trials each, with each shape–label combination occurring once per block. Presentation of the 32 trials in a block was randomized separately for each participant.

The testing phase immediately followed. On each trial, a shape appeared concurrently with the onset of a spoken stimulus. The participants’ task was to indicate whether the spoken word was the correct name for the shape by pressing the right or left arrow on the number keypad. Participants were asked to respond on every trial, even if unsure, and without lingering too much on their choice. However, it was not presented as a speeded task. Testing consisted of four blocks of 32 trials. Within each block, each picture and each label occurred only once, either as a correct pairing or as one of three types of lure pairings. Lures corresponded to rearranged pairs; one third corresponded to an onset-overlapping lure, another third to an offset-overlapping lure, and the last third to an unrelated lure (i.e., a pair in which the shape’s given label shared no phonemes with the original label). Within participants, the same labels occurred in the role of both onset-

overlapping and offset-overlapping lures. As mentioned above, this design ensures that differences in responding to onset- and offset-overlapping lures cannot be accounted for by their similarity to other words in the set or to existing English words. The block in which a given shape was presented along with one of the four possible labels was counterbalanced across participants—that is, one participant might get the shape originally named /bæf/ associated with its onset-overlapping label /bæv/ in Block 1 of the test, another participant would get that combination in Block 2, and so on. Presentation order within a block was randomized. Thus, the test pair was a rearranged pair on 75% of the trials. The disparity between the number of correct and incorrect pairs was intended to maximize the false-alarm rate.

Results and Discussion

Performance of 16 of our participants revealed a reasonably good level of learning, with studied pairs being identified as correct more often than rearranged unrelated pairs (69%, *SD* = 0.19, for the studied pairs and 17%, *SD* = 0.18, for the rearranged

unrelated pairs). The remaining two participants failed to demonstrate this learning, categorizing rearranged unrelated pairs as correct at least as often as correct pairs were judged correct. Their data were excluded from further analyses.¹

Figure 2 presents the false-alarm rate for rearranged pairs with an onset-overlapping and offset-overlapping label as a function of the phonetic distance of the differing sound. Apparent from the graph, lures with an onset-overlapping label were erroneously recognized as a studied pair more often than lures with an offset-overlapping label. Furthermore, within each kind of lure, there was a strong impact of the featural distance between the lure label and its correct counterpart, with phonetically more similar lures associated with a higher rate of false-alarm responses than phonetically less similar lures.

To confirm these observations, we conducted a repeated measures analysis of variance (ANOVA) with Location of Overlap (onset, offset) and Featural Distance (close, distant) as within-participant factors. (Note that in this and all following experiments, data were arcsine-square-root transformed to account for possible statistical distortions of percentage data, which are not normally distributed.) We also report item analyses, in which an item corresponds to a given shape-label pair, with Location of Overlap and Featural Distance as between-items factors. There was a main effect of Location of Overlap, $F(1, 15) = 29.73, p < .001$, and $F(1, 60) = 40.28, p < .001, \eta_G^2 = .143$. There was also a main effect of Featural Distance, $F(1, 15) = 22.39, p = .001$, and $F(1, 60) = 73.34, p < .001, \eta_G^2 = .235$. The interaction of these two factors was not significant, $F(1, 15) < 1$ and $F(1, 60) = 2.09, p = .15, \eta_G^2 = .001$.

Participants erroneously recognized a lure with a label that overlapped with the correct label at onset more often than a lure with an offset-overlapping label. It is important to note that the rate of false alarms was equally affected by the featural distance between the first or last sound of the spoken word and the original label. Although only suggestive, this lack of interaction speaks

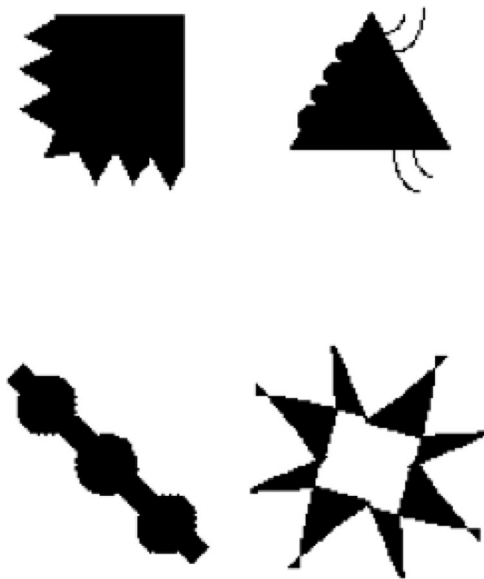


Figure 1. Sample pictures used in Experiments 1, 3, and 4.

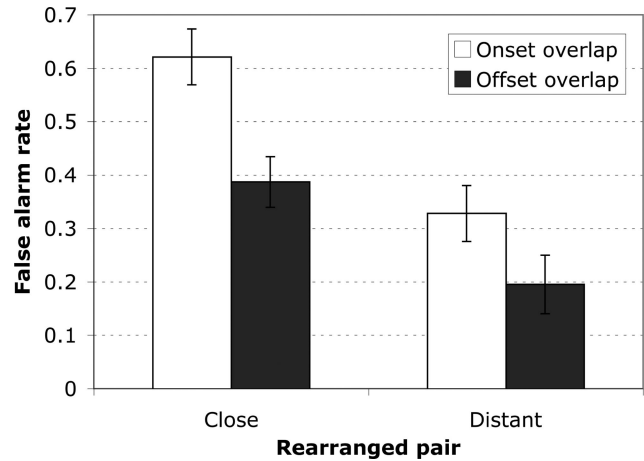


Figure 2. Experiment 1, false-alarm rates to onset-overlapping (white) and offset-overlapping (dark) lure labels. Error bars correspond to standard errors.

against the possibility that the large false-alarm rate on onset-overlapping competitors mainly results from participants preparing their response based on the initial portion of the spoken word. Instead, we argue, learning to differentiate two words and their attributes is more difficult and error prone when the two words start with the same sounds than when they end with the same sounds because of participants' greater attention devoted to the early portion of word, even though the composition of our paired-associate set made onset and coda consonants equally informative.

A possible objection to this conclusion may be raised: The greater false-alarm rate on onset-overlapping lures over offset-overlapping lures may have arisen from a greater difficulty in perceiving and discriminating coda consonants, compared with the discrimination of onset consonants. Such a difference in the perception of onsets and codas may have resulted from two aspects of our stimuli. First, we used slightly different sets of sounds as onsets and codas. The specific sets of sounds, although matched in terms of their phonetic distance, may differ in terms of their acoustic similarity and confusability; that is, the onset sounds may be more similar to one another than the coda sounds are. Second, even if the same sounds had been used at onset and coda, it is possible that the observed asymmetry in false-alarm rate reflects a difference in the ease with which onset and coda consonants were identified during study or at test. Indeed, there is some evidence that, especially in continuous and/or casual speech, the strength of articulation for coda consonants may be reduced compared with that of onset consonants, rendering coda consonants harder to discriminate than onset consonants (Redford & Diehl, 1999). Although our speaker produced what seemed like clear and well-articulated tokens in isolation, coda consonants may have been harder to perceive than onset consonants. We addressed this concern in Experiment 2 by having participants transcribe the stimuli

¹ In this experiment, as well as in subsequent experiments that involved paired-associate learning, the data from participants who were excluded for poor performance showed no significant difference in false-alarm rate between onset- and offset-overlapping lures.

used in Experiment 1 and rate the similarity between the onset-overlapping labels and the offset-overlapping labels. We reasoned that if the coda consonants were harder to perceive than the onset consonants, participants' transcriptions should reflect these misperceptions, and participants should rate onset-overlapping items as more similar to one another than offset-overlapping items.

Experiment 2

Method

Participants. Twenty-five students took part. Eight of them were students from the same pool as in Experiment 1 and completed the similarity ratings only. The remaining 17 were students from the University of California, San Diego, psychology participant pool and participated in both the transcription task and the similarity-rating task. One participant reported reversing the similarity rating scale and was eliminated from this task, leaving 24 participants in the rating task. Another three participants misunderstood instructions for the transcription task and were eliminated from that task, yielding 14 participants in the transcription task.

Stimuli. Using the same words as those in Experiment 1, we created 128 pairs of spoken labels following the same structure as that of Experiment 1's test trials: Each of the 32 labels was paired with itself, an onset-overlapping label, an offset-overlapping label, or an unrelated label. The two spoken items were separated by a 500-ms interval of silence.² The ordering of the trials differed between the two sets of students. For University of Pennsylvania students, trials occurred in blocks of 32, during which each label occurred once as the first label in a pair and once as the second label within a pair. Order of trials within a block was randomized for each participant, the order of presentation of the blocks was counterbalanced across participants, and the entire set of blocks was run twice, for a total of 256 trials. For University of California students, the order of trials was completely randomized without blocking, and each of the 128 trials was heard just once.

Procedure. On each trial of a similarity-rating task, participants heard two syllables spoken successively and were asked to rate their similarity on an 8-point scale. Participants were instructed to respond with 1 if they judged the two syllables to be *identical*. In any other case, they were to rate their similarity, with 2 associated with a *very similar* judgment and 8 with a *very dissimilar* one.

On each trial of the transcription task, participants heard a single spoken word and provided a transcription by entering it on the computer keyboard. Participants were able to hear the stimulus as often as necessary before providing a transcription. The 32 items were presented two times across two blocks of trials with random orders.

Results and Discussion

Similarity-rating task. We distinguished trials that received a similarity rating of 1 (i.e., trials in which participants judged the items to be identical) from other trials and analyzed these data separately.

First, we examined how the *identical* ratings were distributed across pair types. Although participants gave this rating to the pairs in which one of the 32 labels was paired with itself 95% of the time, they did so on only 1.0% of the onset-overlapping pairs, 1.1% of the

offset-overlapping pairs, and 0.2% of the unrelated pairs. Although such ratings were more frequent when the pair contained overlapping labels than when it contained unrelated labels, $\chi^2(1) = 6.145$, $p = .013$, they were equally frequent for onset- and offset-overlapping pairs, $\chi^2(1) = 0.048$, $p = .83$. This suggests that coda consonants were not harder to discriminate than onset consonants.

Figure 3 displays the mean similarity ratings for pairs judged to be nonidentical as a function of the location of overlap and the feature distance between the sounds that differed between the two items. As apparent on the graph, two labels that differed by a phonetically close consonant were rated as more similar than two labels that differed by a phonetically distant consonant. Effect of phonetic distance was expected to be reflected in similarity judgments and is consistent with the confusability data from Experiment 1. The location of overlap, however, had an effect on ratings such that two onset-overlapping labels were rated less similar to each other than two offset-overlapping labels. Thus, the location of the overlap between two labels affected similarity judgments in the opposite manner that it affected confusability in Experiment 1's recognition-memory task.

An ANOVA on the similarity ratings confirmed these assessments. Ratings of 1 (*identical*) were removed prior to analysis. The effect of Featural Mismatch was significant, $F(1, 23) = 66.19$, $p < .001$, and $F(1, 60) = 94.36$, $p < .001$, $\eta_G^2 = .11$, indicating higher similarity ratings for featurally close words ($M = 3.06$, $SD = 0.92$) than featurally distant words ($M = 3.72$, $SD = 0.90$). There was also a significant effect of Location of Overlap, $F(1, 23) = 8.83$, $p = .007$, and $F(1, 60) = 18.19$, $p < .001$, $\eta_G^2 = .024$, with offset-overlap pairs rated as more similar than onset-overlap pairs. Finally, there was an interaction between the two factors, $F(1, 23) = 10.11$, $p = .004$, and $F(1, 60) = 8.87$, $p = .006$, $\eta_G^2 = .011$, with a greater decrement due to larger Featural Mismatch for onset-overlapping pairs.

Transcription task. Transcriptions were coded to reflect whether the onset and coda consonants had been correctly identified and, when they were identified, whether the transcription happened to be the consonant used in the label's counterpart (e.g., transcribing the label "bav" as "baf").

Labels' onset consonants were accurately transcribed 95.6% of the time and coda consonants 98.7%. The difference in accuracy between onset and coda consonants was largely due to one label, "zoich," whose onset consonant was misheard as *s*. Exclusion of this label brought consonant accuracy to 98.7% and 98.6% for onset and coda consonants, respectively. It is important to note that misidentified consonants were almost never heard as the consonant used in the label's counterpart (0% of the onsets, 0.3% of the codas).

Experiment 2's results speak to the perception and discrimination of onset versus coda consonants. Both the transcription data and the rates with which people failed to discriminate overlapping pairs in the similarity-rating task provide no evidence that coda consonants are harder to perceive or identify than onset consonants. Thus, the difference in false-alarm rate between onset- and offset-overlapping lures from Experiment 1's recognition-memory

² University of Pennsylvania students were tested on the tokens used in Experiment 1, and University of California students were tested on the tokens used in Experiment 3. No difference emerged between the two groups, and their data were merged.

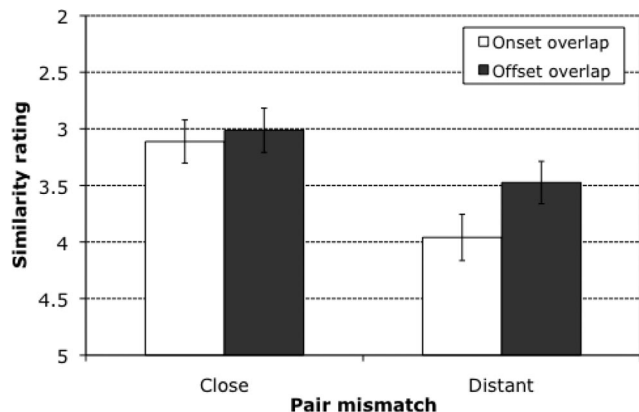


Figure 3. Experiment 2, similarity ratings for onset-overlapping (white) and offset-overlapping (dark) label pairs. Error bars correspond to standard errors.

task cannot be attributed to misperceptions affecting coda consonants (i.e., the consonants that differentiate onset-overlapping labels) more than onset consonants (i.e., the consonants that differentiate offset-overlapping labels).

Similarity judgments revealed an effect of the featural distance between the sounds that differed between the two items when perceived as nonidentical: The more phonetically distant the sounds were, the less similar the two items were judged. It is important to note that two offset-overlapping items (*joop*, *choop*) were judged more similar to each other than two onset-overlapping items (*joop*, *joob*). This tendency is consistent with previous research on word similarity using explicit similarity judgments on adults (e.g., words, Nelson & Nelson, 1970; nonsense syllables, Hahn & Bailey, 2005) and children (e.g., Storkel, 2002). This tendency is the opposite of what would be expected if similarity ratings were predictive of false-alarm rates in Experiment 1's recognition-memory task. Why people rate rhyming words as more similar to one another than words that start with identical sounds is a question we return to in the General Discussion. For now, we can conclude that there is no evidence that the coda consonants in our stimuli were more similar or harder to differentiate than the onset consonants. This lends some support to our proposal that learning two words whose names differ in their last sounds poses greater difficulty than learning two words whose names differ in their first sounds because people are biased to more strongly associate the early sounds of the name with a referent than the last sounds when learning a label–referent association.

Experiment 3 provided another test of our hypothesis that people give greater importance to the early sounds of a label than its last sounds when learning a label–referent association by assessing listeners' recognition of new lures, in addition to rearranged lures. New lures consisted of studied shapes paired with novel, previously unheard labels, which differed from the original shape label's onset or coda consonant by one or multiple phonetic features. Thus, new lures departed from the original labels to the same degree as rearranged lures did. However, new labels are novel strings, and this lack of familiarity should lower false-alarm rates. The question of interest here is whether the lack of familiarity with the new labels will affect false alarms on onset- and offset-overlapping new lures in the same

way. If the false-alarm asymmetry observed in Experiment 1 resulted from people encoding or retrieving labels' original coda consonants less accurately than onset consonants, swapping the codas with a different consonant should lower participants' familiarity with the resulting new labels less than swapping the onsets. Thus, the decline in false-alarm rates between rearranged and new onset-overlapping lures should be smaller than the difference between rearranged and new offset-overlapping lures. If, on the other hand, Experiment 1's false-alarm asymmetry resulted from a stronger association between a referent and the label's first sounds than between the referent and the label's last sounds, the effect of location of overlap between the original label and the lure should be comparable for rearranged and new lures.

Experiment 3

Method

Participants. Thirty-six students from the University of Pennsylvania took part in the experiment for course credit. None of them had participated in the first two experiments. The data from four participants were subsequently excluded because of evidence of poor learning.

Stimuli. For each of the 32 items from Experiment 1, we created two additional items to be used at testing as new lure labels. One such item overlapped with its base item at onset and the other at offset. None of these new items made up existing English words. To maximize the false-alarm rates to new lures, we ensured that the consonants that differed between the original and new labels were consonants that had appeared in that position in other studied items (at onset or coda). The consonant that differed between the new labels and their original counterparts was either phonetically close (i.e., differing by one phonological feature) or distant (i.e., different by two features or more), with slightly more offset-overlapping lures being phonetically distant to their base stimulus than phonetically close (18 vs. 14). Table 1 displays the complete set of stimuli.

Because a large number of stimuli were added for use in the test phase, all words were rerecorded by the same speaker as in Experiment 1 to ensure acoustic uniformity across items. Average word duration was 591 ms.

Procedure and design. The training phase proceeded as in Experiment 1, with the exception that the new recordings for the shapes' labels were used. For testing, we elected to present each participant with only half of the four possible overlapping lures (two rearranged, two new) for each of the 32 items. Participants were randomly assigned to one of two lists, which varied which lures were presented. Thus, for each participant, a given shape was presented four times, once with its correct label, once with an unrelated label, once with a (rearranged or new) onset-overlapping label, and once with a (rearranged or new) offset-overlapping label, yielding 128 testing trials. Among the 32 onset-overlapping lure trials, half of them consisted of a familiar shape associated with a familiar label (i.e., a rearranged pair), and the other half consisted of a familiar shape associated with a novel label (i.e., a new pair). Rearranged and new lures were equally distributed across blocks of testing. Instructions to participants were identical to those in Experiment 1.

Results and Discussion

Overall, participants recognized the pairs they had studied as correct more often than rearranged unrelated pairs ($M = 0.72$,

$SD = 0.16$ vs. $M = 0.22$, $SD = 0.19$), except for four participants, whose data were excluded from further analyses.

Figure 4 presents mean false-alarm rates for both rearranged and new lures. Data are plotted as a function of whether the label presented at test overlapped with the original one at onset or at offset and whether the consonants that differed between the two were phonetically close (i.e., differing by one phonological feature) or distant. As apparent on the graph, the false-alarm rates on the rearranged lures closely replicate the pattern found in Experiment 1. Moreover, false-alarm rates on new lures were noticeably lower than those on rearranged lures, an effect that illustrates the contribution of overall familiarity of the test pair on false-alarm rate. Critically, the decline in false-alarm rate between rearranged and new lures was similar for onset- and offset-overlapping lures.

A three-way ANOVA conducted on the false-alarm rates confirmed these observations. There were main effects of Location of Overlap, $F(1, 31) = 46.31$, $p < .001$, and $F(1, 120) = 19.06$, $p < .001$, $\eta_G^2 = .077$; Featural Distance, $F(1, 31) = 42.51$, $p < .001$, and $F(1, 120) = 32.84$, $p < .001$, $\eta_G^2 = .083$; and Lure Type, $F(1, 31) = 29.73$, $p < .001$, and $F(1, 120) = 51.26$, $p < .001$, $\eta_G^2 = .123$, with no significant interaction between these factors.

Thus, the lack of familiarity with the new labels affected the false-alarm rate to onset- and offset-overlapping lures to the same degree—an unpredicted outcome if participants were less accurate in their encoding or retrieval of labels' coda consonants than their encoding or retrieval of onset consonants. The result, however, is compatible with our claim that people create a stronger association between the shape and the first sounds of its label than between the shape and its label's last sounds.

One account of the results of Experiments 1 and 3 that we have yet to address concerns response anticipation. As pointed out earlier, when the presentation of the shape precedes or is concurrent with the onset of the spoken item, an onset-overlapping rearranged or new lure is temporarily consistent with the correct pair. This is never the case for offset-overlapping lures. The tendency for participants to false-alarm more frequently on onset-overlapping lures may be the result of this temporary match,

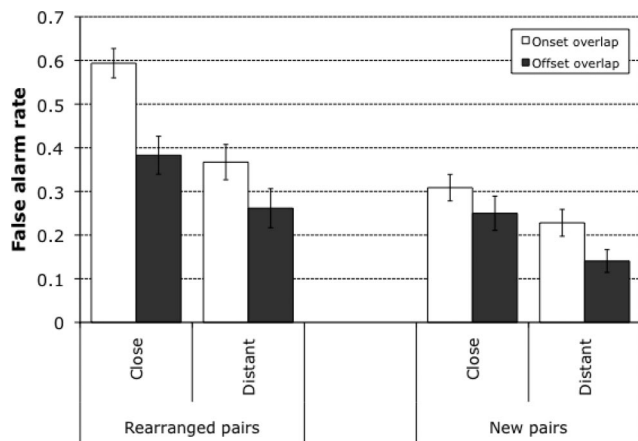


Figure 4. Experiment 3, false-alarm rates to onset-overlapping (white) and offset-overlapping (dark) lure labels as a function of whether the lure consisted of a rearranged pair or a new pair and the featural distance between the consonants that differed between the lure's label and the original label (close vs. distant). Error bars correspond to standard errors.

essentially an effect of online processing. Even though participants are not under time pressure to respond, their decision may nonetheless be influenced primarily by their evaluation of the information they receive early on. One way to circumvent this is to hold off the presentation of the shape until the entire spoken word has been heard. Without any information about the referent, information about a familiar spoken label cannot be evaluated.

In Experiment 4, we examined the contribution of response anticipation to the onset- versus offset-overlapping lure asymmetry observed in Experiment 1 by varying the timing of the visual stimulus with respect to the auditory stimulus at test: The visual shape appeared before the onset of the spoken item, simultaneously with the onset of the item, at the offset of the item's vowel, or after the offset of the item. We asked whether this manipulation affected the degree of onset–offset asymmetry in false alarms. If the asymmetry reflects response anticipation, we reasoned, the earlier the shape becomes available before the spoken label begins, the stronger the impact of the first sounds of the spoken label there should be. This is because participants may be able to generate the label associated with the visible shape, and the match between the shape's label and the first sounds of the spoken label should cause people to false-alarm the onset-overlapping rearranged pairs. Conversely, the greater the delay between the presentation of the spoken label and that of the shape, the smaller the false-alarm rate difference should be between onset- and offset-overlapping lures.

Experiment 4

Method

Participants. Seventy-five students from the University of Pennsylvania participated in the experiment for course credit or for a small monetary compensation. None of these participants had taken part in the previous experiments. The data for two participants were excluded because they were not native speakers of English, and the data of nine further participants were excluded from the analyses because of poor learning (see below).

Stimuli. Visual and auditory stimuli were the same as in Experiment 1.

Procedure. The training phase was identical to that of Experiment 1. The testing phase, however, differed from that of Experiment 1 in the timing with which the visual stimulus appeared with respect to the spoken item. The presentation of the visual shape was varied across four stimulus onset asynchrony (SOA) values. The shape was displayed (a) 667 ms before the onset of the spoken item;³ (b) concurrently with the onset of the spoken item; (c) synchronized with the end of the vowel and onset of the coda consonant, an estimate of the point of disambiguation between the original shape label and the onset-overlapping lure, which was 377 ms on average after the item's onset; or (d) 200 ms after the offset of the coda consonant, which corresponded to 796 ms on average after the item's onset. The end of the vowel was located by visually inspecting each waveform and spectrogram for a drop in amplitude and disappearance of formant bands and then

³ This time interval had originally been set to 200 ms, but an error in programming changed it to the default DMDX frame duration value of 667 ms (40 frames).

auditorily checking that the preceding portion did not contain audible coda consonant material and that the following portion did not contain audible vowel material.

An equal number of the 32 testing trials on each block were assigned to each of the four SOA conditions. Across blocks, an equal number of SOA trials were assigned to each Location of Overlap and Featural Distance condition.

Except for the timing of the shape presentation with respect to that of the spoken stimulus, the training and testing procedures were identical to those of Experiment 1.

Results and Discussion

Data from nine participants for whom the rate of recognition of the correct pairs was not greater than that of erroneous recognition of rearranged unrelated pairs were excluded. The remaining participants displayed satisfactory learning during training (correct: $M = 0.68$, $SD = 0.17$; unrelated: $M = 0.18$, $SD = 0.17$). Performance on correct and unrelated pair trials was unaffected by varying the SOA.

Figure 5 presents the false-alarm rate to rearranged onset- and offset-overlapping lures as a function of the featural distance between the sounds that differed between the original label and the lure, across the four SOA conditions. As in Experiments 1 and 3, false alarms to onset-overlapping lures were more frequent than to offset-overlapping lures, and false alarms were less frequent as the featural distance between the lure and the original label increased. Varying the SOA had a very modest impact on the false-alarm rates. These observations were confirmed by a three-way ANOVA conducted on the false-alarm rates, with Location of Overlap, Featural Distance, and SOA as within-participant factors. We found a main effect of Location of Overlap, $F(1, 63) = 41.43$, $p < .001$, and $F(2, 60) = 82.8$, $p < .001$, $\eta_G^2 = .057$, and a main effect of Featural Distance, $F(1, 63) = 50.67$, $p < .001$, and $F(2, 60) = 84.03$, $p < .001$, $\eta_G^2 = .059$. There was also a small effect of SOA, with a mild decrease of all false alarms as the SOA increased, $F(3, 189) = 4.57$, $p = .004$, and $F(2, 180) = 2.30$, $p = .08$, $\eta_G^2 = .008$. There was no interaction between the Location of Overlap and SOA conditions, suggesting that the

tendency to false-alarm more on onset-overlapping lures than on offset-overlapping lures remained fairly stable despite large changes in the timing of the shape presentation.

To further assess the stability of the difference in false-alarm rates between onset- and offset-overlapping lures across SOA conditions, we computed, for each participant, the size of the difference between the two overlap conditions for each SOA condition, with SOA expressed in absolute time (in milliseconds) with respect to the onset of the spoken item. For the SOA conditions in which the exact timing varied across items (i.e., item's vowel offset and 200 ms after the item's offset), we used the value averaged across items. For each participant, we computed the slope of the regression line expressing the correlation, across the four data points, between the SOA value (in milliseconds) and the false-alarm rate difference. If there is a linear trend toward a decrease in the discrepancy between early- and late-overlap false alarms, we reasoned, the slope should be negative and significantly less than zero. The mean slope across participants was $-.00002$ ($SD = .00021$). A t test confirmed that the slope was indistinguishable from zero, $t(63) = 0.77$, $p = .45$.

Across a substantial range of SOA values, the tendency to erroneously recognize onset-overlapping lures as correct pairs more often than offset-overlapping lures remained relatively constant. We found no evidence that the asymmetry declined when most or even all of the spoken label had been heard before the shape appeared. The appearance of the shape determines the point at which the information about the spoken label can be evaluated and contribute to a decision. The results of Experiment 4 are important because they suggest that, for the most part, participants' propensity to confuse an onset-overlapping lure for a correct pair is not the result of response anticipation based on the temporary match between the spoken input and the visual shape at test. Thus, we argue, the larger confusion between onset-overlapping labels arises from the greater contribution of a spoken name's early sounds to word learning.

General Discussion

The present study addresses how the temporal organization of spoken words affects their encoding in and retrieval from memory. In a series of experiments, we demonstrated that spoken labels and their attributes are more confusable when their names overlap early than when they overlap late. Experiment 1 established the basic phenomenon, and Experiments 2, 3, and 4 explored factors that might have contributed to the confusion asymmetry between onset- and offset-overlapping items. Experiment 2 verified that this confusion asymmetry did not result from an asymmetry in perceiving or discriminating the consonants that distinguished the onset- and offset-overlapping labels. Participants transcribed onset and coda consonants equally well. When presented with two stimuli sequentially and asked to judge their similarity, listeners judged two onset-overlapping labels to be identical very rarely and just as rarely as they judged two offset-overlapping labels to be identical. Furthermore, similarity judgments on stimuli pairs judged not to be identical showed that onset-overlapping stimuli were judged less similar to one another than offset-overlapping stimuli. This finding is the opposite of what would be expected if degree of perceived similarity predicted recognition memory. We return to the discrepancy between these two measures below. For

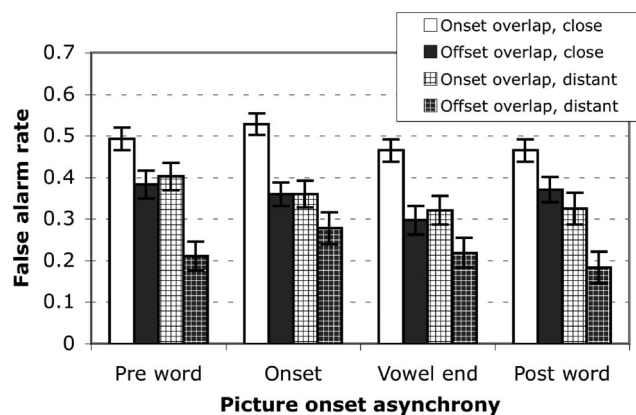


Figure 5. Experiment 4, false-alarm rates at each picture onset asynchrony. Onset-overlapping bars are white, and offset-overlapping lure label bars are dark. Solid bars are close lure labels, crisscrossed bars are distant labels. Error bars correspond to standard errors.

now, we can conclude that the tendency to confuse two onset-overlapping labels cannot be attributed to the difficulty in perceiving the coda consonants that distinguish them.

Experiment 3 replicated Experiment 1's results on rearranged lures and revealed an analogous asymmetry between onset- and offset-overlapping new lures. Although new lures led to significantly lower false-alarm rates than rearranged lures did, the degree to which participants erroneously recognized the lure as old was modulated by the location of overlap between the new label and the shape's original label, lending support to the hypothesis that the early portion of a spoken word is given greater weight than its late portion when learning to associate a referent with a name.

Finally, in Experiment 4, we tested the role that response preparation and anticipation play in accounting for the observed asymmetry. We varied the timing with which the shape became available with respect to the auditory presentation of the label at test. If the asymmetry was mainly caused by response anticipation, we reasoned, the disproportionate contribution of the early portion of the spoken label to decision making and response preparation should decrease as the presentation of the shape is delayed in time, leading to a decrease in the difference in false-alarm rates between onset- and offset-overlapping lures. Experiment 4 revealed that the confusion asymmetry between onset- and offset-overlapping items is extremely robust across a range of stimulus-presentation conditions. Even when the correctness of the label could not be gauged against a picture referent until after the end of the word, onset-overlapping lures were more confusable with a correct pair than were offset-overlapping lures. Thus, there was no evidence that people accept onset-overlapping lures as correct pairs more often than offset-overlapping lures because the false alarms are, for the most part, the result of response anticipation based on the early sounds of the lure's label.

Taken together, the results suggest that, in the early stage of word learning at least, the early portion of a name is more strongly associated with its referent than the late portion. As discussed in the introduction, this asymmetry may reflect a general tendency for people to attend to and/or remember the beginning of a sequence, akin to the primacy effect observed in free or serial recall of a list of items. Alternatively, or in addition to this general constraint, participants' greater attention to and retention of early phonetic information may reflect its greater contribution to lexical contrast in English. Indeed, the English vocabulary is characterized by many monosyllabic rhyming words. Determining which one is being uttered requires listeners to obtain detailed perceptual information about the onset consonant of a spoken word. Monosyllabic words that differ only in their coda consonants, however, are less frequent; consequently, discriminating among them depends on a less detailed analysis of their coda consonants. It is worth mentioning that although these characteristics are true of English, they do not apply to the lexicon we exposed participants to. The set of labels we exposed people to at training rendered onset and coda consonants equally informative because each label had the same number of onset- and offset-overlapping competitors. Thus, our account assumes that the statistics of English vocabulary continued to constrain participants' attentional focus when listening to spoken names even when the set of names that people were exposed to did not share the same properties.

How these results may extend to other languages would shed light on their underlying causes. Ziegler and Goswami (2004; see

also Peereman, Dubois-Dunilac, Perruchet, & Content, 2004) found that, as in English, monosyllabic words in French, Dutch, and German have more offset-overlapping neighbors (i.e., words that share all their sounds except for their onset consonants) than they have onset-overlapping neighbors (i.e., words that share all their sounds except for their coda consonants). If the bias we observed here (i.e., for listeners to learn and remember the first sounds of labels for new referents better than their last sounds) originates from the statistical properties of the lexicon of one's native language, a similar bias should be observed with adult French, Dutch, or German listeners. This account also predicts that the bias to attend to the early part of a word more than its later portion becomes more pronounced as listeners accumulate the statistics of their language's lexicon. Thus, this account is consistent with the absence of an onset-consonant bias in toddlers' learning new name-referent associations, as reported by Nazzi and colleagues (Nazzi & Bertoncini, in press; Nazzi et al., 2009).

We now turn to the discrepancy between the asymmetry we observed in the recognition-memory task and the one reported in the similarity-judgment task here (Experiment 2) and in the literature (e.g., Hahn & Bailey, 2005; Nelson & Nelson, 1970). When listeners are given two stimuli and asked to evaluate their similarity, stimuli that overlap at offset are judged more similar than stimuli that overlap at onset, whether the items are words or nonwords. Experiment 2 confirmed this tendency on the present stimuli. However, in the recognition-memory task, labels that overlap at onset are confused with one another more than labels that overlap at offset. Thus, it appears that the similarity judgments do not predict confusability among labels in a straightforward manner. Rather, we propose, similarity judgments reflect the outcome of a similarity comparison, which, according to theories of similarity comparisons (Hahn & Bailey, 2005; Hahn, Chater, & Richardson, 2003; Markman & Gentner, 1993; Tversky, 1977), requires the extraction of a structure in the stimuli to be compared. Accordingly, similarity judgment reflects the quality of the structural alignment between the stimuli. Thus, words sharing a rime (i.e., the phonological unit that comprises the vowel and any coda consonant) may be viewed as more similar than words sharing their onset and nucleus because the former share elements within the same abstract structure, the rime. There is some evidence that the sensitivity to similarity between rhyming words develops later than that between onset-overlapping words, the former being often linked to the development of literacy (De Cara & Goswami, 2002; Jusczyk, Goodman, & Baumann, 1999).

The organization of monosyllabic words into onset and rimes is not universal, and in some languages, such as Korean, onset and nucleus are grouped into one unit, the body (e.g., Yoon, Bolger, Kwon, & Perfetti, 2002). What determines the internal constituency of words within a language is not well understood. Nonetheless, some have proposed that phonological units reflect the statistical and phonotactic properties of a language's speech sounds. For example, monosyllabic English words can be decomposed into onsets and rimes because of the stronger cohesion between the nucleus and the coda than between the onset and the nucleus, with the degree of cohesion among segments determined by their transitional probabilities. What our recognition-memory task indicates is that the phonological components of words have no special status when learning names for new referents. However, this result does not rule out the possibility that a language's phonotactics

influence which elements of a word listeners have learned to pay greater attention to. On the basis of English phonotactics, English speakers may have learned to devote special attention to the syllable onsets because of their higher entropy or the lower cohesion between the onset and the following vowel. Investigating the phenomenon reported in the present study in other languages, such as Korean, would provide critical data to these questions.

The present work provides evidence for a substantial influence of the temporal structure of spoken words on people's learning of name–shape paired associates, over and above effects of online processing. A pair that differs from a trained one in the last sound of its name is more likely to be erroneously recognized as one of the trained pairs than a pair that differs from a trained one in the first sound of its name. We interpret this result as evidence that the beginnings of the names were learned and associated with their shapes more accurately than the ends of the names. Further work will explore the nature of such biases for early information, which may characterize associative learning with speech as well as with other classes of temporally structured events.

References

- Alloppena, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439.
- Benkí, J. R. (2003). Quantitative evaluation of lexical status, word frequency, and neighborhood density as context effects in spoken word recognition. *Journal of the Acoustical Society of America*, 113, 1689–1705.
- Boersma, P., & Weenink, D. (2006). *Praat: Doing phonetics by computer*. Retrieved from <http://www.praat.org>
- Cole, R. A., & Jakimik, J. (1980). How are syllables used to recognize words? *Journal of the Acoustical Society of America*, 67, 965–970.
- Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, 32, 193–210.
- Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2006). Acquiring an artificial lexicon: Segment type and order information in early lexical entries. *Journal of Memory and Language*, 54, 1–19.
- De Cara, B., & Goswami, U. (2002). Similarity relations among spoken words: The special status of rimes in English. *Behavior Research Methods, Instruments, & Computers*, 34, 416–423.
- Dyne, A. M., Humphreys, M. S., Bain, J. D., & Pike, R. (1990). Associative interference effects in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 813–824.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116–124.
- Hahn, U., & Bailey, T. M. (2005). What makes words sound similar? *Cognition*, 97, 227–267.
- Hahn, U., Chater, N., & Richardson, L. B. (2003). Similarity as transformation. *Cognition*, 87, 1–32.
- Jusczyk, P. W., Goodman, M. B., & Baumann, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language*, 40, 62–82.
- Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language*, 37, 295–311.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, 132, 202–227.
- Malmberg, K. J., & Xu, J. (2007). The influence of averaging and noisy decision strategies on the recognition memory ROC. *Psychonomic Bulletin & Review*, 13, 99–105.
- Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25, 431–467.
- Marslen-Wilson, W. D. (1984). Function and process in spoken word-recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of language processes* (pp. 125–150). Hillsdale, NJ: Erlbaum.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition and beginning literacy* (pp. 89–120). Hillsdale, NJ: Erlbaum.
- Miller, G. A. (1951). *Language and communication*. New York, NY: McGraw-Hill.
- Monsell, S., & Hirsh, K. W. (1998). Competitor priming in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1495–1520.
- Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, 98, 13–30.
- Nazzi, T., & Bertoncini, J. (in press). Consonant specificity in onset and coda positions in early lexical acquisition. *Language and Speech*.
- Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: Cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology*, 102, 522–537.
- Nelson, D. L., & Borden, R. C. (1973). Interference produced by phonetic similarities: Stimulus recognition, associative retrieval, or both? *Journal of Experimental Psychology*, 97, 167–169.
- Nelson, D. L., & Brooks, D. H. (1973). Functional independence of pictures and their verbal memory codes. *Journal of Experimental Psychology*, 98, 44–48.
- Nelson, D. L., & Nelson, L. D. (1970). Rated acoustic (articulatory) similarity for word pairs varying in number and ordinal position of common letters. *Psychonomic Science*, 19, 81–82.
- Norris, D. (1982). Autonomous processes in comprehension: A reply to Marslen-Wilson and Tyler. *Cognition*, 11, 97–101.
- Pantelis, P. C., van Vugt, M. K., Sekuler, R., Wilson, H. R., & Kahana, M. J. (2008). Why are some people's names easier to learn than others? The effects of face similarity on memory for face–name associations. *Memory & Cognition*, 36, 1182–1195.
- Peereman, R., Dubois-Dunilac, N., Perruchet, P., & Content, A. (2004). Distributional properties of language and sub-syllabic processing units. In Bonin P. (Ed.), *Mental lexicon: Some words to talk about words* (pp. 215–235). New York, NY: Nova.
- Redford, M. A., & Diehl, R. L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*, 106, 1555–1565.
- Salasoo, A., & Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language*, 24, 210–231.
- Shatzman, K. B., & McQueen, J. M. (2006). Prosodic knowledge affects the recognition of newly acquired words. *Psychological Science*, 17, 372–377.
- Stoel-Gammon, C., & Dunn, C. (1985). *Normal and disordered phonology in children*. Austin, TX: Pro-Ed.
- Storkel, H. L. (2001). Learning new words: Phonotactic probability in

- language development. *Journal of Speech, Language, and Hearing Research*, 44, 1321–1337.
- Storkel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language*, 29, 251–274.
- Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, 49, 1175–1192.
- Swingle, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech*, 46, 265–294.
- Swingle, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, 60, 252–269.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.
- Verde, M. F. (2004). Associative interference in recognition memory: A dual-process account. *Memory & Cognition*, 32, 1273–1283.
- Vitevitch, M. S. (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 270–278.
- Wheeler, J. W., & Nelson, D. L. (1982). Developmental trends in the phonemic organization of individual words. *American Journal of Psychology*, 95, 223–233.
- Yoon, H.-K., Bolger, D. J., Kwon, O.-S., & Perfetti, C. A. (2002). Sub-syllabic units in reading: A difference between Korean and English. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Studies in Written Language and Literacy: Vol. 11. Precursors of functional literacy* (pp. 139–163). Amsterdam, the Netherlands: Benjamins.
- Ziegler, J. C., & Goswami, U. (2004). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, 3–29.

Received February 7, 2009

Revision received August 10, 2009

Accepted August 13, 2009 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.