

## Lexical ambiguity and acoustic distance in discrimination

Chelsea Sanker\*

**Abstract.** This work presents a perceptual study on how acoustic details and knowledge of the lexicon influence discrimination decisions. English-speaking listeners were less likely to identify phonologically matching items as the same when they differed in vowel duration, but differences in mean F0 did not have an effect. Although both are components of English contrasts, the results only provide evidence for attention to vowel duration as a potentially contrastive cue. Lexical ambiguity was a predictor of response time. Pairs with matching duration were identified more quickly than pairs with distinct duration, but only among lexically ambiguous items, indicating that lexical ambiguity mediates attention to acoustic detail. Lexical ambiguity also interacted with neighborhood density: Among lexically unambiguous words, the proportion of ‘same’ responses decreased with neighborhood density, but there was no effect among lexically ambiguous words. This interaction suggests that evaluating phonological similarity depends more on lexical information when the items are lexically unambiguous.

**Keywords.** lexical ambiguity; acoustic distance; auditory perception

**1. Introduction.** In perceptual tasks, how does acoustic distance in different characteristics influence discrimination and how might that interact with lexical ambiguity? Listeners can be sensitive to acoustic distance within phonological categories (Lieberman et al. 1957; Pisoni & Tash 1974), though much of the work on acoustic distance in discrimination in just a few characteristics, particularly VOT. It is unclear whether acoustic distance would have similar effects across different characteristics.

Lexical ambiguity can influence processing of words (e.g. Kellas et al. 1988; Borowsky & Masson 1996), but most studies on ambiguity effects use orthographic stimuli, leaving open questions about how lexical ambiguity influences perception of acoustic input. Sanker (2019) demonstrates differences in responses to acoustically presented pairs of lexically ambiguous words and lexically unambiguous words; some differences can be attributed to acoustic differences between homophone mates in production (cf. Guion 1995; Gahl 2008), though other differences suggest an effect of ambiguity itself in how an acoustic stimulus is evaluated.

This work presents a perceptual study on how sub-phonemic details and knowledge of the lexicon influence decisions in a discrimination task. English-speaking listeners were less likely to identify phonologically matching paired items as the same when they differed in vowel duration, but differences in mean F0 did not have an effect. Distance in duration also had an effect on response time, but only among lexically ambiguous items, suggesting that lexical ambiguity mediates attention to acoustic detail.

**1.1 PRODUCTION OF ACOUSTIC DETAIL.** Despite their matching phonological identity, homophone mates can exhibit relatively consistent phonetic differences in production due to factors such as lexical frequency (e.g. Guion 1995; Gahl 2008) and part of speech (e.g. Sorensen et al. 1978; Conwell 2017). However, differences are found most reliably in natural speech (e.g. Gahl 2008; Lohman 2017), and can disappear when words are produced outside of their

---

\* Author: Chelsea Sanker, Yale University ([chelsea.sanker@yale.edu](mailto:chelsea.sanker@yale.edu)).

natural contexts. Differences correlated with lexical frequency can largely be eliminated by producing words in isolation or in frame sentences rather than in meaningful sentences (Guion 1995). Similarly, effects of part of speech can largely be explained as the result of prosodic structure, because they can be eliminated by producing all words in the same position in the sentence or the phrase (Sorensen et al. 1978; Conwell 2017). Productions are also influenced by predictability in context; controlling for predictability can eliminate other factors as predictors of phonetic characteristics in natural speech (Jurafsky et al. 2002).

Even if phonetic differences across phonologically identical words are driven by predictability and similar factors, it is plausible that listeners might learn these differences. The existence of these differences in production does not necessarily indicate that homophone mates have different phonetic details in their representations, but some theories predict that listeners do associate phonetic details with particular items (e.g. Johnson 1997; Pierrehumbert 2002). In particular, Exemplar Theory posits that each token of a word enters the cloud of that word's representation; differences in phonetic convergence based on lexical frequency have been taken as evidence for these lexically-specific representations (e.g. Goldinger 1998; Babel 2010). However, other evidence from convergence indicates that phonological shifts are consistent across the same segment in different words (Nielsen 2011; Pardo et al. 2012), which suggests that phonetic details in the representation are at the phonological level rather than the lexical level.

1.2 PERCEPTION OF ACOUSTIC DETAIL. Listeners are sensitive to acoustic detail in perception. Response times are faster for discriminating between paired items that are more acoustically distinct (Pisoni & Tash 1974). Identification decisions about sounds that are closer to a category boundary are slower (Pisoni & Tash 1974) and less accurate (Liberman et al. 1957) than decisions about sounds that are more central within a category; greater activation with greater prototypicality is also reflected in stronger priming effects produced by stimuli that are phonetically prototypical within their categories (Andruski et al. 1994; Ju & Luce 2006). Eye tracking similarly reflects greater uncertainty about categorization with items closer to a category boundary (McMurray et al. 2002). However, much of the work on acoustic distance in discrimination uses VOT contrasts; it is unclear whether effects would be consistent across different characteristics.

Listeners can be sensitive to acoustic detail and homophone mates can exhibit acoustic differences, so can listeners distinguish between homophone mates? Sanker (2019) found that pairs of homophone mates (e.g. *sun-son*) are more likely to be identified as phonologically distinct than pairs of the same word (e.g. *sun-sun*) when the stimuli were extracted from meaningful sentences. This result likely reflects the greater acoustic distance between homophone mates than between pairs of different speakers' productions of the same word; this difference was largely apparent in vowel formants, though also found to a smaller degree in several other characteristics. When stimuli were made from productions in isolation, the two pair types did not differ from each other in acoustic distance between paired items or in response patterns, which further indicates that the discrimination results cannot be interpreted as evidence for distinct acoustic details in the lexical representations of homophone mates.

Even clearer evidence against homophone mates having distinct phonetic details in their representations is provided by lexical identification results; perceptual identification tasks requiring listeners to decide between homophone mates produce accuracy that is at chance or only marginally above chance (Bond 1973; Sanker 2019). Lexically ambiguous acoustic stim-

uli activate all of the associated homophone mates, which is demonstrated in priming of orthographic lexical decisions after exposure to ambiguous acoustic stimuli (Onifer & Swinney 1981; Grainger et al. 2001).

1.3 EFFECTS OF LEXICAL AMBIGUITY. Even if homophone mates cannot be distinguished from each other, lexical ambiguity seems to have an effect on lexical access and phonological processing. Lexical ambiguity in a discrimination task with acoustic stimuli produces slower responses and fewer identifications of phonologically matching items as being the same (Sanker 2019), which could indicate that listeners' awareness of ambiguity in the lexicon impedes their decisions about such items; with lexically unambiguous stimuli, they can approach the task lexically, while they cannot determine that lexically ambiguous stimuli match based on narrowing both down to being the same word. The pattern of responses to lexically ambiguous items could also indicate that this lexical uncertainty causes listeners to attend more closely to acoustic detail, because they are expecting differences.

Responses to lexically ambiguous words and lexically unambiguous words differ in a range of tasks. Some studies have found that responses are faster for homophones than non-homophones in orthographic lexical decisions (e.g. Rubenstein et al. 1970; Jastrzembski 1981; Kellas et al. 1988; Borowsky & Masson 1996) and naming tasks (e.g. Hino et al. 2002), though other studies have found the opposite effect (e.g. Pexman et al. 2001). However, the apparent effects of lexical ambiguity might be indirect, due to characteristics which correlate with lexical ambiguity. Gernsbacher (1984) suggested that faster responses for lexically ambiguous words were the result of familiarity, as words with multiple meanings are often more familiar. However, experiments controlling for factors such as familiarity and frequency still found effects of lexical ambiguity in lexical decision tasks (e.g. Kellas et al. 1988; Borowsky & Masson 1996). Faster responses for lexically ambiguous words have been interpreted as the result of multiple lexical entries with the same phonological form all contributing to the activation of that phonological form (Jastrzembski 1981; Kellas et al. 1988).

The effects of lexical ambiguity in processing depend on whether the task motivates semantic processing or not, and whether the phonological advantage offsets the effects of semantic competition (Joordens & Besner 1994). The predicted results thus differ by task, with facilitation only when the response is consistent for all homophone mates, e.g. lexical decision and naming. In semantic decision tasks, homophones have a consistent disadvantage, at least when the meanings are in disagreement (Hino et al. 2002; Siakaluk et al. 2007); the disadvantage can be intensified by priming their homophone mates (Pylkkänen et al. 2006).

1.4 LEXICAL FREQUENCY AND NEIGHBORHOOD DENSITY. Words with homophones have higher frequency and greater neighborhood density on average than other words. While studies often control for frequency, they less often control for neighborhood density, resulting in different neighborhood densities of the lexically ambiguous and lexically unambiguous (e.g. Rubenstein et al. 1970; Hino & Lupker 1996). As a result of this relationship, some of the effects of frequency and neighborhood density might appear to be effects of lexical ambiguity. It is additionally unclear whether effects of frequency and neighborhood density would be the same for lexically ambiguous words and for lexically unambiguous words; most studies do not look for an interaction.

A range of experiments demonstrate faster processing for high frequency words, usually with orthographic stimuli, e.g. for lexical decision (Stanners et al. 1975; Murray & Forster

2004) and semantic categorization (Monsell et al. 1989; Lewellen et al. 1993). The same effect is also found in picture naming (Oldfield & Wingfield 1965; Carroll & White 1973). Higher frequency words also have a processing advantage with acoustic stimuli; in lexically ambiguous items, the higher frequency meaning will be more strongly activated (Simpson & Burgess 1985; Binder & Rayner 1998). When stimuli are partially obscured by noise or contain segments manipulated to be acoustically ambiguous, listeners are often more likely to interpret them as higher frequency words (Howes 1957; Connine et al. 1993), though Samuel (1981) did not find an effect of word frequency in phoneme restoration.

In discrimination tasks, response times increase with higher neighborhood density (Vitevitch & Luce 1999; Luce & Large 2001); because there are more phonologically similar competitors, listeners take longer to determine whether paired forms from a dense neighborhood are distinct words or not. In contrast, high neighborhood density may facilitate reading tasks (Borowsky & Masson 1996; Mulatti et al. 2006), though this apparent effect might be better described as an effect of phonotactic probability, which is highly correlated with neighborhood density. In lexical decision tasks, neighborhood density does not seem to be a predictor of response latency (Borowsky & Masson 1996; Vitevitch & Luce 1999); lexical neighbors do not serve as competitors in such decisions.

**2. Methods.** 48 native speakers of American English (6 male; mean age 20.7) participated in the task and were paid for participation. Stimuli were pairs of monosyllabic English words, produced in isolation by two female native speakers of English. The juxtaposed items in each pair differed in speaker, to encourage phonological decisions rather than decisions about phonetic identity. The full list of stimulus words is given in Tables 3-4 at the end of the paper.

Among the phonologically matching pairs, there were two conditions of lexical ambiguity. All participants heard all of the words: (a) 48 ambiguous (e.g. *made-made*, cf. *maid*), and (b) 80 unambiguous (e.g. *mud-mud*). The lexically ambiguous and unambiguous groups were matched to have comparable mean neighborhood density. Due to lack of agreement in the literature about how to measure lexical frequency of homophones, frequency matching between the two categories was a compromise between having the mean frequency of the lexically unambiguous words match the mean combined frequency of the homophones (e.g. the frequency of /sʌn/ as the frequency of *sun* + the frequency of *son*), having it match the mean individual frequency of the homophones (the frequency of /sʌn/ as the mean of *sun*'s frequency and *son*'s frequency), or having it match the frequency of the higher frequency homophone mate (e.g. the frequency of /sʌn/ as the frequency of *son*). In analysis, the frequency of lexically ambiguous words was treated as being the frequency of the higher frequency homophone mate, based on evidence that acoustic stimuli elicit faster retrieval of a higher frequency homophone mate than its lower frequency mates (Simpson & Burgess 1985; Binder & Rayner 1998).

There were two conditions of acoustic manipulations in the phonologically matching pairs. Each participant was assigned to a single acoustic manipulation condition: (a) F0: half of pairs were manipulated to have equal mean F0 and half differed by 70 Hz, and (b) vowel duration: half of paired items had equal vowel duration and half differed by 100 ms.

There were an equal number of phonologically matching pairs and phonologically distinct filler pairs, which had a single segmental contrast. For half of the participants, the difference was in onsets (e.g. *pile-file*), and for the other half of the participants, the difference was in codas (e.g. *leaf-leave*). None of the words that appeared in these filler pairs also appeared in a

phonologically matching pair. Results are reported only for the matching pairs.

Participants heard the word pairs presented over headphones, separated by 200 ms of silence. Instructions on a computer screen asked listeners to identify each pair as either being the same or different. Responses were given with the left and right arrow keys on the keyboard; which side corresponded to ‘same’ and ‘different’ was balanced across listeners. The experiment was self-timed; the next trial began 500 ms after a response was given. Response times were measured from the beginning of the second word. Trials with response times greater than 5 s or less than 250 ms were excluded from analysis (1.9% of the data).

The experiment was run in PsychoPy (Pierce 2007). All statistical results are from mixed effects models calculated with the lme4 package in R (Bates et al. 2015). p-values were calculated by the lmerTest package (Kuznetsova et al. 2015).

**3. Results.** Two aspects of responses are analyzed: The proportion of ‘same’ responses and the log response time. The former reflects perception of the two forms as phonologically the same or distinct, while the latter can capture processing factors that might not be reflected in listeners’ ultimate decisions about the stimuli.

Table 1 presents a logistic mixed effects model for ‘same’ responses to the phonologically matching stimuli. The random effects were participant and word pair. The fixed effects were manipulation type (duration, F0), manipulation distance (close, further), lexical ambiguity (ambiguous, unambiguous), neighborhood density, the interaction between lexical ambiguity and neighborhood density, the interaction between lexical ambiguity and manipulation distance, and the interaction between manipulation type and manipulation distance.

	Estimate	Std. Error	z value	p value
(Intercept)	2.91	0.27	10.8	< 0.001
Manipulation Duration	0.774	0.213	3.64	< 0.001
Distance Further	0.0328	0.139	0.237	0.813
LexicalAmbiguity Ambiguous	-0.685	0.372	-1.84	0.0656
NeighborhoodDensity	-0.0183	0.00736	-2.49	0.0127
LexicalAmbig Ambiguous * NeighbDens	0.0198	0.0103	1.91	0.0559
Distance Further * LexicalAmbig Ambiguous	0.0544	0.186	0.293	0.769
Manip Duration * Distance Further	-0.688	0.184	-3.73	< 0.001

Table 1: Mixed effects logistic regression model for responses. *Intercept: LexicalAmbiguity = unambiguous, Manipulation = F0, Distance = close.*

There was a significant effect of manipulation type on responses. Participants in the duration manipulation condition were more likely to identify phonologically matching pairs as being the same word. There was also a significant interaction between manipulation type and manipulated distance, which was the primary source of the main effect of manipulation type. Listeners were more likely to identify the phonologically matching pairs as the same when they matched in vowel duration than when they had distinct vowel durations, but having matching or distinct F0 mean did not influence responses. The effects of manipulation type and manipulated distance are illustrated in Figure 1. Adding a three-way interaction between manipulation type, manipulation distance, and lexical ambiguity did not significantly improve the model ( $\chi^2 = 4.27$ ,  $df = 2$ ,  $p = 0.118$ ).

There was a marginal effect of lexical ambiguity on responses when the interaction with

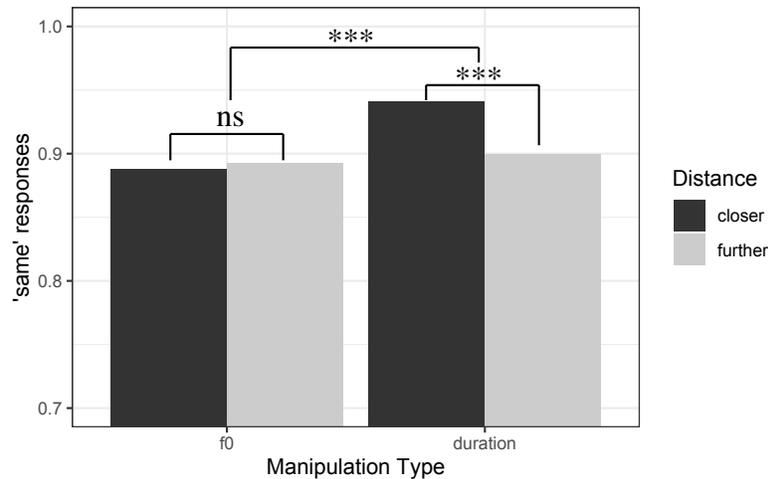


Figure 1: 'Same' responses, by manipulation type and manipulation distance

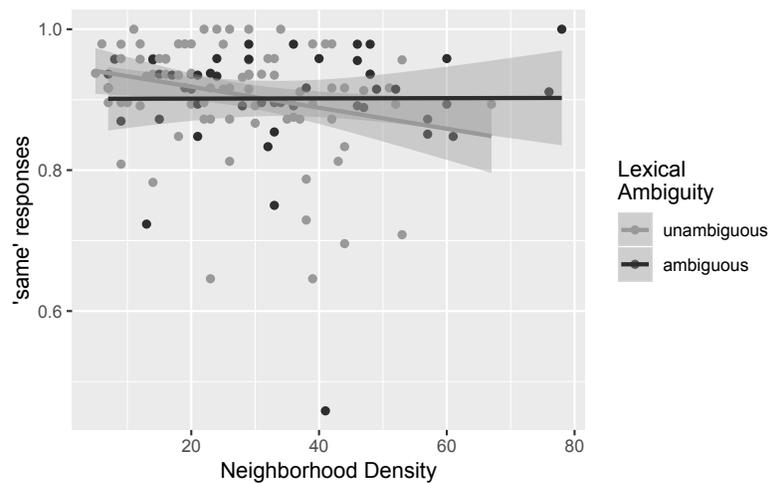


Figure 2: 'Same' responses, by lexical ambiguity and neighborhood density

neighborhood density was included, because the effect of neighborhood density differed between lexically unambiguous and lexically ambiguous words. Within lexically unambiguous words, listeners were less likely to identify words with higher neighborhood density as being the same. Among lexically ambiguous words, the effect was absent, indicated in the opposite and nearly equal coefficient for the interaction between lexical ambiguity and neighborhood density. The effects of neighborhood density and lexical ambiguity are illustrated in Figure 2. Figure 3 demonstrates the small overall effect of lexical ambiguity.

Including the position of the contrast (onsets or codas) that appeared in the phonologically distinct filler pairs did not improve the model ( $\chi^2 = 0.0176$ ,  $df = 1$ ,  $p = 0.894$ ). That is, the position of the phonological contrasts that listeners heard in phonologically distinct pairs did not influence their responses to phonologically matching pairs; listeners did not seem to develop expectations about the position where contrasts would appear, or at least did not develop expectations that influenced their decisions.

The model was not improved by adding lexical frequency ( $\chi^2 = 0.557$ ,  $df = 1$ ,  $p = 0.455$ )

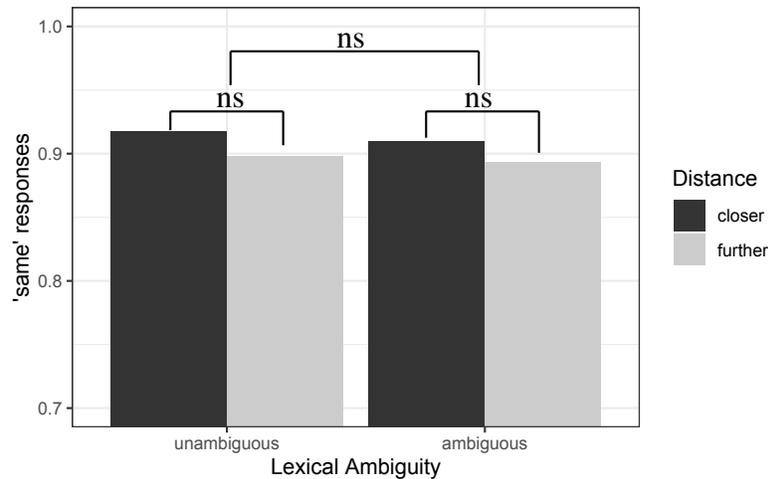


Figure 3: 'Same' responses, by lexical ambiguity and manipulation distance

or lexical frequency and an interaction between lexical frequency and lexical ambiguity ( $\chi^2 = 2.82$ ,  $df = 2$ ,  $p = 0.244$ ), even when neighborhood density was excluded from the model.

Table 2 presents a linear mixed effects model for log response times of responses to the phonologically matching stimuli. The random effects were participant and word pair. The fixed effects were manipulation type (duration, F0), manipulation distance (close, further), lexical ambiguity (ambiguous, unambiguous), neighborhood density, log lexical frequency, the interaction between lexical ambiguity and neighborhood density, the interaction between lexical ambiguity and lexical frequency, the interaction between lexical ambiguity and manipulation type, the interaction between lexical ambiguity and manipulation distance, the interaction between manipulation type and manipulation distance, and the three-way interaction between lexical ambiguity, manipulation type, and manipulation distance.

	Estimate	Std. Error	t value	p value
(Intercept)	0.829	0.0466	17.8	< 0.001
Manipulation Duration	0.00418	0.0321	0.13	0.897
Distance Further	-0.00142	0.01	-0.142	0.887
LexicalAmbiguity Ambiguous	-0.0983	0.0726	-1.35	0.178
NeighborhoodDensity	0.000319	0.000573	0.557	0.579
LexicalFrequency	-0.0187	0.00398	-4.7	< 0.001
LexAmbig Ambig * NeighbDens	0.000135	0.000814	0.166	0.869
LexAmbig Ambig * LexFreq	0.0129	0.00714	1.81	0.0734
Manip Dur * LexAmbig Ambig	-0.0588	0.0164	-3.59	< 0.001
Dist Further * LexAmbig Ambig	-0.0244	0.0164	-1.49	0.136
Manip Dur * Dist Further	0.0133	0.0142	0.939	0.348
Manip Dur * Dist Further * LexAmbig Ambig	0.0549	0.0232	2.36	0.0182

Table 2: Mixed effects linear regression model for log response time. *Intercept: LexicalAmbig = unambiguous, Manipulation = F0, Distance = close.*

There was no main effect of manipulation type, manipulation distance, or lexical ambiguity on response time. However, there were interactions between them. The three-way inter-

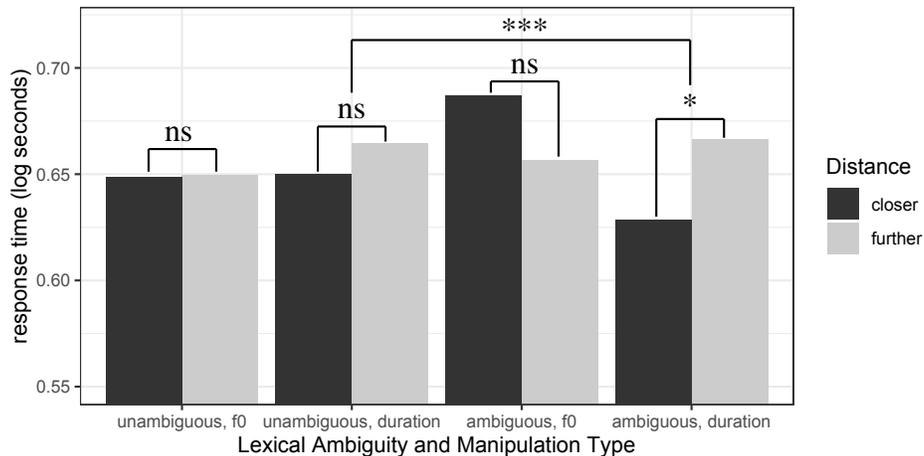


Figure 4: Log response times, by lexical ambiguity, manipulation type, and manipulation distance

action between manipulation type, manipulation distance, and lexical ambiguity was included because it significantly improves the fit, as compared to a model that parallels the one in Table 1 ( $\chi^2 = 12.9$ ,  $df = 2$ ,  $p = 0.00159$ ). Among the lexically ambiguous items, response times were faster with vowel duration contrasts than with F0 contrasts, and pairs with matching duration were identified more quickly than items with distinct durations. The fast responses to lexically ambiguous words with matching vowel duration produced both the significant interaction between manipulation type and lexical ambiguity and the significant interaction between manipulation type, manipulation distance, and lexical ambiguity. No other pair type exhibited a significant effect of manipulation distance on response time, as illustrated in Figure 4.

There was no effect of neighborhood density on response time. It is included to parallel the model for ‘same’ responses presented in Table 1. Notably, the effect was similarly absent if lexical frequency was excluded from the model, so the lack of effect here cannot be attributed to a correlation between neighborhood density and lexical frequency.

Log lexical frequency was a significant predictor of response time. Responses were faster for higher frequency words, at least among lexically unambiguous words. For lexically ambiguous words, the effect was weaker, though the difference between the two was only marginally significant. The effect of lexical frequency on response time is presented in Figure 5. The result may indicate that lexical frequency is not a major influence on processing time for lexically ambiguous items. However, it is also possible that frequency was a weak predictor of response time for these items simply because the relevant measure of their frequency is something different from the measure used here. As described above, lexical frequency for lexically ambiguous items was treated as the frequency of the higher frequency homophone mate.

Including the position of the contrast (onsets or codas) that appeared in the phonologically distinct filler pairs did not improve the model ( $\chi^2 = 0.813$ ,  $df = 1$ ,  $p = 0.367$ ). The lack of effect of contrast position suggests that listeners did not develop expectations about the positions in which contrasts would appear; expecting contrasts earlier in the word would produce shorter response times for listeners in the onset-contrasts condition. It would be reasonable to expect the position of the contrast to interact with the manipulation type, as vowel duration is a cue to coda voicing contrasts; however, including an interaction between contrast position and manipulation type also did not improve the model ( $\chi^2 = 0.337$ ,  $df = 2$ ,  $p = 0.845$ ).

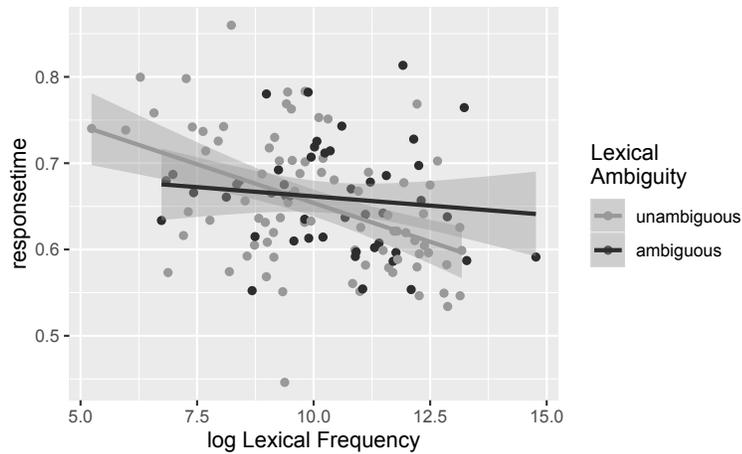


Figure 5: log response times, by lexical ambiguity and lexical frequency

**4. Discussion.** The results demonstrate that effects of acoustic distance between phonologically matching stimuli are particular to the characteristic that is manipulated. They additionally suggest that lexical ambiguity influences processing in discrimination tasks, producing effects of neighborhood density and acoustic distance that are distinct from the effects with lexically unambiguous items.

Listeners were less likely to identify the phonologically matching pairs as the same when they had larger differences in vowel duration, but having matching or distinct F0 mean did not influence responses. This suggests that at least in the context set up by this experiment, English speakers were attending to vowel duration as a potential cue contributing to phonological contrasts, but did not similarly perceive F0 mean as a potential element of contrast. This result might be due to F0 functioning as a phonological cue in English only as part of particular F0 contours, or depending on co-occurrence with other cues such as differences in VOT.

Lexical ambiguity did not have an overall effect on response time or the proportion of ‘same’ responses. However, lexical ambiguity interacted with manipulation type and manipulation distance as predictors of response time. Among the lexically ambiguous items, response times were faster with duration contrasts than with F0 contrasts, and pairs with matching duration were identified more quickly than items with more distinct duration. Among lexically unambiguous items and items with an F0 contrast, the manipulation distance did not have a significant effect on response time. This result suggests that lexical ambiguity influences listeners’ attention to detail; consistent with the interaction between manipulation type and manipulation distance as predictors of ‘same’ responses, increased attention to acoustic detail only impacted vowel duration. Greater attention to detail could result from the expectation that these items might be phonologically distinct based on having two competing lexical entries. With lexically ambiguous items, listeners are likely to activate two possible lexical entries that are consistent with the stimuli, and require more careful processing to evaluate whether or not they are phonologically distinct. However, the lack of effect on ‘same’ responses indicates that this additional processing does not change listeners’ ultimate decisions based on the phonetic forms.

Neighborhood density interacted with lexical ambiguity in predicting ‘same’ responses. Among lexically ambiguous words, there was no evidence for an effect of neighborhood density on ‘same’ responses. In contrast, the proportion of ‘same’ responses with lexically un-

ambiguous words decreased with neighborhood density. The result is consistent with lexical ambiguity forcing listeners to use a different process of evaluation. With lexically unambiguous items, listeners can identify paired items as matching based on narrowing down the lexical identity of both items to the same entry. On the other hand, with lexically ambiguous items, the lexical identity cannot be narrowed down to a single item, so listeners must evaluate phonological status without relying on how it aligns with the lexicon. This phonological evaluation strategy makes the existence of lexical neighbors less relevant. Given that there is actually a higher proportion of 'same' responses for lexically ambiguous items than for lexically unambiguous items at high neighborhood density, lexical information might not just aid decisions in sparse neighborhoods but also impede decisions in dense neighborhoods. However, there were not enough items in very high density neighborhoods to conclusively demonstrate the strength of this pattern at this end of the continuum.

It is notable that the interaction between neighborhood density and lexical ambiguity could create apparent effects of lexical ambiguity based on the neighborhood density of the words used in an experiment, even if there are no effects of lexical ambiguity itself. If a word set primarily contains words of high neighborhood density, the proportion of 'same' responses would be higher for lexically ambiguous words. If a word set primarily contains words of low neighborhood density, the proportion of 'same' responses for lexically ambiguous words would be lower, particularly if the neighborhood density in each group is not controlled, because lexically unambiguous words are likely to have lower mean neighborhood density than lexically ambiguous words.

However, neighborhood density was not a significant predictor of response time, in contrast to many previous studies that have found that response times increase with neighborhood density in discrimination tasks (e.g. Vitevitch & Luce 1999; Luce & Large 2001). It is possible that the set of stimuli in this experiment did not include a wide enough distribution to capture effects of neighborhood density on response time; studies on neighborhood density often select stimuli of very high neighborhood density and very low neighborhood density and test it as a binary effect rather than a continuous effect.

Listeners gave faster responses for higher frequency words. The effect of lexical frequency was primarily apparent among lexically unambiguous words, and much weaker among lexically ambiguous words. It is possible that frequency was a weak predictor of response time due to how frequency was measured. In this study, lexical frequency for lexically ambiguous items was treated as the frequency of the higher frequency homophone mate, based on previous work demonstrating that a higher frequency homophone mate is more rapidly retrieved than a lower frequency mate when listeners hear ambiguous acoustic stimuli (Simpson & Burgess 1985; Binder & Rayner 1998). However, the best way of measuring the frequency of homophones is debated in the literature. When orthographic stimuli or semantic context make it possible to distinguish between the frequency of each homophone mate, many studies support measurement of individual frequency, based on individual frequencies predicting response latency in tasks such as lexical decision (e.g. Grainger et al. 2001), gaze duration in reading (e.g. Binder & Rayner 1998), and picture naming (e.g. Caramazza et al. 2001). However, there is also evidence for combined frequency or the frequency of the higher-frequency homophone mate as the relevant measure, based on low-frequency homophones exhibiting patterns like their high-frequency homophone mates in speed of translation and picture naming (e.g. Jescheniak & Levelt 1994; Antón-Méndez et al. 2012) and low susceptibility to produc-

tion errors (e.g. Dell 1990). It is possible that none of these measures of lexically frequency would alone be sufficient to explain the frequency-related behavior of homophones.

**5. Conclusions.** This study demonstrates that listeners’ sensitivity to within-category acoustic distance is particular to the characteristic being manipulated. Only vowel duration influenced responses in this task, though both vowel duration and F0 are components of English contrasts. The effect of vowel duration on discrimination suggests that listeners attend to it as a potentially contrastive cue, while mean F0 might function as a cue only as part of differences in F0 contour or in combination with other cues.

Pairs with matching vowel duration were identified more quickly than pairs with distinct duration among lexically ambiguous items, but not among lexically unambiguous items, suggesting that lexical ambiguity mediates attention to acoustic detail. However, lexical ambiguity was not a significant predictor of responses, either as a main effect or in interaction with manipulation distance; while listeners spend longer considering lexically ambiguous items with larger duration differences, this deliberation does not change their ultimate decisions.

Lexical ambiguity also interacted with neighborhood density: Among lexically unambiguous words, the proportion of ‘same’ responses decreased with neighborhood density, but there was no effect among lexically ambiguous words. This interaction suggests that knowledge of lexical competitors influences phonological decisions in lexically unambiguous words, while processing lexically ambiguous words is less shaped by lexical competition. Given differences in the mean neighborhood density of lexically ambiguous and unambiguous words, some of the main effects of lexical ambiguity found in previous work could be due to neighborhood density rather than ambiguity itself.

add	feet	peak	sole
ail	flea	piece	stake
bee	foul	plain	steal
break	groan	plot	sun
club	heard	raise	tale
coarse	knight	right	tide
cord	loot	road	toad
creak	made	sale	waste
cruise	mail	scene	week
days	meat	sea	weight
deer	odd	sheer	whale
dough	pair	soar	whole

Table 3: Lexically ambiguous items

bad	dim	guess	move	sheet
badge	dog	haze	mud	shoes
ban	fed	heat	neat	sing
beam	feel	job	pan	soon
best	flaws	keys	peeve	soup
big	fluff	lane	pick	stem
braid	food	less	place	stop
brave	fought	life	plays	suit
brief	freeze	lit	pose	take
chin	frock	look	ran	tea
cliff	game	make	ridge	team
clove	give	men	room	tell
cooed	go	mill	same	tome
crave	good	mob	sat	toss
cup	greed	mop	shade	weave
cut	grief	moss	sheen	wig

Table 4: Lexically unambiguous items

## References

- Andruski, Jean E., Sheila E. Blumstein & Martha Burton. 1994. The effect of subphonetic differences on lexical access. *Cognition* 52(3). 163–187. [http://dx.doi.org/10.1016/0010-0277\(94\)90042-6](http://dx.doi.org/10.1016/0010-0277(94)90042-6).
- Antón-Méndez, Inés, Carson T. Schütze, Mary K. Champion & Tamar H. Gollan. 2012. What the tip-of-the-tongue (tot) says about homophone frequency inheritance. *Memory & Cognition* 40(5). 802–811. <http://dx.doi.org/10.3758/s13421-012-0189-1>.
- Babel, Molly. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39. 437–456. <http://dx.doi.org/10.1017/S0047404510000400>.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. <http://dx.doi.org/10.18637/jss.v067.i01>.
- Binder, Katherine S & Keith Rayner. 1998. Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review* 5(2). 271–276. <http://dx.doi.org/10.3758/BF03212950>.
- Bond, Zinny S. 1973. The perception of sub-phonemic phonetic differences. *Language and Speech* 16(4). 351–355. <http://dx.doi.org/10.1177/002383097301600405>.
- Borowsky, Ron & Michael E. J. Masson. 1996. Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(1). 63–85. <http://dx.doi.org/10.1037/0278-7393.22.1.63>.
- Caramazza, Alfonso, Albert Costa, Michele Miozzo & Yanchao Bi. 2001. The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27(6). 1430–1450. <http://dx.doi.org/10.1037/0278-7393.27.6.1430>.
- Carroll, John B. & Margaret N. White. 1973. Word frequency and age of acquisition as determiners of picture-naming latency. *The Quarterly Journal of Experimental Psychology* 25(1). 85–95. <http://dx.doi.org/10.1080/14640747308400325>.
- Connine, Cynthia M., Debra Titone & Jian Wang. 1993. Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(1). 81–94. <http://dx.doi.org/10.1037/0278-7393.19.1.81>.
- Conwell, Erin. 2017. Prosodic disambiguation of noun/verb homophones in child-directed speech. *Journal of Child Language* 44(3). 734–751. <http://dx.doi.org/10.1017/S030500091600009X>.
- Dell, Gary S. 1990. Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes* 5(4). 313–349. <http://dx.doi.org/10.1080/01690969008407066>.
- Gahl, Susanne. 2008. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84(3). 474–496. <http://dx.doi.org/10.1353/lan.0.0035>.
- Gernsbacher, Morton A. 1984. Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology* 113(2). 256–281. <http://dx.doi.org/10.1037/0096-3445.113.2.256>.
- Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2). 251–279. <http://dx.doi.org/10.1037/0033-295X.105.2.251>.

- Grainger, Jonathan, Maryline Nguyen Van Kang & Juan Segui. 2001. Cross-modal repetition priming of heterographic homophones. *Memory & Cognition* 29(1). 53–61. <http://dx.doi.org/10.3758/BF03195740>.
- Guion, Susan G. 1995. Word frequency effects among homonyms. In *Texas Linguistic Forum*, vol. 35, 103–116.
- Hino, Yasushi & Stephen J. Lupker. 1996. Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance* 22(6). 1331–1356. <http://dx.doi.org/10.1037/0096-1523.22.6.1331>.
- Hino, Yasushi, Stephen J. Lupker & Penny M. Pexman. 2002. Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28(4). 686–713. <http://dx.doi.org/10.1037/0278-7393.28.4.686>.
- Howes, Davis. 1957. On the relation between the intelligibility and frequency of occurrence of English words. *Journal of the Acoustical Society of America* 29(2). 296–305. <http://dx.doi.org/10.1121/1.1908862>.
- Jastrzembski, James E. 1981. Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology* 13(2). 278–305. [http://dx.doi.org/doi:10.1016/0010-0285\(81\)90011-6](http://dx.doi.org/doi:10.1016/0010-0285(81)90011-6).
- Jescheniak, Jörg D. & Willem J. M. Levelt. 1994. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20(4). 824–843. <http://dx.doi.org/10.1037/0278-7393.20.4.824>.
- Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (eds.), *Talker Variability in Speech Processing*, 145–165. San Diego, CA: Academic Press.
- Joordens, Steve & Derek Besner. 1994. When banking on meaning is not (yet) money in the bank: Explorations in connectionist modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20(5). 1051–1062. <http://dx.doi.org/10.1037/0278-7393.20.5.1051>.
- Ju, Min & Paul A. Luce. 2006. Representational specificity of within-category phonetic variation in the long-term mental lexicon. *Journal of Experimental Psychology: Human Perception and Performance* 32(1). 120–138. <http://dx.doi.org/10.1037/0096-1523.32.1.120>.
- Jurafsky, Daniel, Alan Bell & Cynthia Girand. 2002. The role of the lemma in form variation. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology VII*, 3–34. Berlin: Mouton de Gruyter.
- Kellas, George, F. Richard Ferraro & Greg B. Simpson. 1988. Lexical ambiguity and the time-course of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 14(4). 601–609. <http://dx.doi.org/10.1037/0096-1523.14.4.601>.
- Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2015. *lmerTest: Tests in Linear Mixed Effects Models*. <https://CRAN.R-project.org/package=lmerTest>. R package version 2.0-29.
- Lewellen, Mary J., Stephen D. Goldinger, David B. Pisoni & Beth G. Greene. 1993. Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision,

- and semantic categorization. *Journal of Experimental Psychology: General* 122(3). 316–330. <http://dx.doi.org/10.1037/0096-3445.122.3.316>.
- Liberman, Alvin M., Katherine Safford Harris, Howard S. Hoffman & Belver C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5). 358–368. <http://dx.doi.org/10.1037/h0044417>.
- Lohman, Arne. 2017. *Cut*(n) and *cut*(v) are not homophones: Lemma frequency affects the duration of noun-verb conversion pairs. *Journal of Linguistics* 1–25. <http://dx.doi.org/10.1017/S0022226717000378>.
- Luce, Paul A. & Nathan R. Large. 2001. Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes* 16(5-6). 565–581. <http://dx.doi.org/10.1080/01690960143000137>.
- McMurray, Bob, Michael K. Tanenhaus & Richard N. Aslin. 2002. Gradient effects of within-category phonetic variation on lexical access. *Cognition* 86(2). B33–B42. [http://dx.doi.org/10.1016/S0010-0277\(02\)00157-9](http://dx.doi.org/10.1016/S0010-0277(02)00157-9).
- Monsell, Stephen, Michael C. Doyle & Patrick N. Haggard. 1989. Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General* 118(1). 43–71. <http://dx.doi.org/10.1037//0096-3445.118.1.43>.
- Mulatti, Claudio, Michael G Reynolds & Derek Besner. 2006. Neighborhood effects in reading aloud: New findings and new challenges for computational models. *Journal of Experimental Psychology: Human Perception and Performance* 32(4). 799–810. <http://dx.doi.org/10.1037/0096-1523.32.4.799>.
- Murray, Wayne S. & Kenneth I. Forster. 2004. Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review* 111(3). 721–756. <http://dx.doi.org/10.1037/0033-295X.111.3.721>.
- Nielsen, Kuniko. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39. 132–142. <http://dx.doi.org/10.1016/j.wocn.2010.12.007>.
- Oldfield, Richard C & Arthur Wingfield. 1965. Response latencies in naming objects. *Quarterly Journal of Experimental Psychology* 17(4). 273–281. <http://dx.doi.org/10.1080/17470216508416445>.
- Onifer, William & David A. Swinney. 1981. Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition* 9(3). 225–236. <http://dx.doi.org/10.3758/BF03196957>.
- Pardo, Jennifer S., Rachel Gibbons, Alexandra Suppes & Robert M. Krauss. 2012. Phonetic convergence in college roommates. *Journal of Phonetics* 40. 190–197. <http://dx.doi.org/10.1016/j.wocn.2011.10.001>.
- Pexman, Penny M., Stephen J. Lupker & Debra Jared. 2001. Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27(1). 139–156. <http://dx.doi.org/10.1037/0278-7393.27.1.139>.
- Pierce, Jonathan W. 2007. PsychoPy–Psychophysics software in Python. *Journal of Neuroscience Methods* 162(1-2). 8–13. <http://dx.doi.org/10.1016/j.jneumeth.2006.11.017>.
- Pierrehumbert, Janet. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 101–140. Berlin: Mouton de Gruyter.
- Pisoni, David B. & Jeffrey Tash. 1974. Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics* 15(2). 285–290. <http://dx.doi.org/10.3758/bf03213946>.

- Pylkkänen, Liina, Rodolfo Llinás & Gregory L. Murphy. 2006. The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience* 18(1). 97–109. <http://dx.doi.org/10.1162/089892906775250003>.
- Rubenstein, Herbert, Lonnie Garfield & Jane A. Millikan. 1970. Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior* 9(5). 487–494. [http://dx.doi.org/10.1016/S0022-5371\(70\)80091-3](http://dx.doi.org/10.1016/S0022-5371(70)80091-3).
- Samuel, Arthur G. 1981. The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance* 7(5). 1124–1131. <http://dx.doi.org/10.1037//0096-1523.7.5.1124>.
- Sanker, Chelsea. 2019. Effects of lexical ambiguity, frequency, and acoustic details in auditory perception. *Attention, Perception, & Psychophysics* 81. 323–343. <http://dx.doi.org/10.3758/s13414-018-1604-x>.
- Siakaluk, Paul D., Penny M. Pexman, Christopher R. Sears & William J. Owen. 2007. Multiple meanings are not necessarily a disadvantage in semantic processing: Evidence from homophone effects in semantic categorisation. *Language and Cognitive Processes* 22(3). 453–467. <http://dx.doi.org/10.1080/01690960600834756>.
- Simpson, Greg B. & Curt Burgess. 1985. Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance* 11(1). 28–39. <http://dx.doi.org/10.1037/0096-1523.11.1.28>.
- Sorensen, John M., William E. Cooper & Jeanne M. Paccia. 1978. Speech timing of grammatical categories. *Cognition* 6(2). 135–153. [http://dx.doi.org/10.1016/0010-0277\(78\)90019-7](http://dx.doi.org/10.1016/0010-0277(78)90019-7).
- Stanners, Robert F., James E. Jastrzembski & Allen Westbrook. 1975. Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behavior* 14(3). 259–264. [http://dx.doi.org/10.1016/S0022-5371\(75\)80069-7](http://dx.doi.org/10.1016/S0022-5371(75)80069-7).
- Vitevitch, Michael S. & Paul A. Luce. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40(3). 374–408. <http://dx.doi.org/10.1006/jmla.1998.2618>.