

Article

# The Targetedness of English Schwa: Evidence from Schwa-Initial Minimal Pairs

Emily R. Napoli \* and Cynthia G. Clopper \*

Department of Linguistics, The Ohio State University, Columbus, OH 43210, USA

\* Correspondence: [napoli.44@osu.edu](mailto:napoli.44@osu.edu) (E.R.N.); [clopper.1@osu.edu](mailto:clopper.1@osu.edu) (C.G.C.)

**Abstract:** Schwa in English shows a considerable amount of contextual variation, to the extent that previous work has proposed that it is acoustically targetless. Although the consensus of previous research seems to suggest that schwa is targeted, the sources of schwa's contextual variation have yet to be fully explained. We explored a potential source of variation in English schwa, namely, whether schwa occurs in a content word (word-initial schwa, e.g., *accompany*) or is a function word (phrase-initial schwa, e.g., *a company*). We sought to determine whether English speakers distinguish word- and phrase-initial schwas in production, as well as whether word- and phrase-initial schwas differ in their level of targetedness. To elicit hyperarticulation of word- and phrase-initial schwas and thereby facilitate our ability to observe their targets, participants produced ambiguous and unambiguous word- and phrase-initial schwa pairs in neutral and biased sentence contexts. The first and second formant trajectories of the schwas were analyzed using growth curve analysis, allowing us to demonstrate that word-initial and phrase-initial schwas are both targeted and have different targets. Ultimately, our results suggest different underlying representations for schwas in function and content words.

**Keywords:** schwa; American English; formant trajectory; growth curve analysis



**Citation:** Napoli, Emily R., and Cynthia G. Clopper. 2024. The Targetedness of English Schwa: Evidence from Schwa-Initial Minimal Pairs. *Languages* 9: 130. <https://doi.org/10.3390/languages9040130>

Academic Editor: Jaydene Elvin

Received: 29 November 2023

Revised: 22 February 2024

Accepted: 22 March 2024

Published: 2 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Variability in English Schwa

In the International Phonetic Alphabet, the schwa symbol represents a mid-central vowel. However, schwa is used to transcribe unstressed neutral vowels in languages such as English, Dutch, and German (Wiese 1986; Booji 1995) and, therefore, represents a wide range of sounds showing considerable variation, depending on their phonemic context (e.g., Bates 1995; van Bergem 1994; Koopmans-Van Benium 1994). This contextual variation has led to a debate in the literature about whether schwa is targetless and unspecified acoustically and/or articulatorily, such that its quality is entirely dependent on the surrounding context (e.g., Bates 1995), or targeted and specified acoustically and/or articulatorily, such that its variability is a consequence of its position in unstressed, shorter syllables (e.g., Flemming 2009). According to Keating (1988), a targetless sound is one that is completely phonetically unspecified. When phonetic rules build trajectories between segments, a phonetically unspecified sound will contribute no trajectory of its own, and its acoustic quality, as a result, will be exclusively determined by the surrounding sounds. Unspecified sounds may, therefore, be modeled acoustically using “the straight-line interpolation between adjacent context segments” (Bates 1995, p. 3). The targetedness of schwa can, therefore, be assessed by considering the extent to which its formants are straight-line interpolations between the adjacent segments. A targetless schwa is phonetically unspecified and will, therefore, have formants that are straight-line interpolations between the surrounding segments. A targeted schwa is phonetically specified and will, therefore, have formants that deviate from the straight-line interpolation towards the phonetic target.

Acoustic evidence for targetless schwa comes from observations that schwa has high levels of contextual dependency along F1 and F2 in comparison to other vowels (Bates

1995, for English schwa; van Bergem 1994, for Dutch schwa). However, a greater amount of both acoustic evidence (e.g., Bakst and Niziolek 2021; Barry 1998; Cohen Priva and Strand 2023; Flemming 2009; Flemming and Johnson 2007; Kondo 1994) and articulatory evidence (Browman and Goldstein 1992; Gick 2002) supports the idea of targeted schwa. In terms of acoustic evidence, the quality of schwa is affected by speaking rate (Barry 1998; Cohen Priva and Strand 2023), which is inconsistent with the targetless account. Targetless accounts predict that schwa, already maximally affected by coarticulation, should not be affected by changes in speaking rate. The quality of schwa is also modulated during sensorimotor adaptation, another observation that is inconsistent with the targetless account, which would predict that altered feedback would not produce a mismatch that speakers must correct for (Bakst and Niziolek 2021). Moreover, contrary to a targetless account, which predicts that the formant trajectories of targetless sounds would be the linear interpolation between surrounding sounds (Bates 1995; Keating 1988), the F1 and F2 trajectories of schwa have also been found to be different from a linear interpolation of surrounding sounds (Kondo 1994). In terms of articulatory evidence, the production of schwa is best modeled using a separate schwa parameter (i.e., target; Browman and Goldstein 1992) and is produced with a retracted tongue root (Gick 2002). Both of these findings are inconsistent with a targetless account which posits that schwa is produced in an articulatorily neutral position.

Although the consensus of previous research seems to be that schwa is indeed targeted, schwa still shows a considerable amount of contextual variation, and the sources of this variation have yet to be fully explained. One possibility is that the variation in schwa quality in English is due to the existence of multiple sounds labeled as schwa, which differ in their targetedness and/or their targets. Flemming and Johnson (2007) and Flemming (2009) identified two such schwa variants: word-final schwa (e.g., *Rosa*), which is more consistent in its quality, and word-medial schwa (e.g., *suggest*), which is more variable. Flemming and Johnson (2007) and Flemming (2009) hypothesized that this difference may be because word-final schwa can contrast with other vowels (e.g., *Rosa* vs. *Rosie* and *Rosa's* vs. *Roses*), which means variation in schwa quality is constrained to avoid producing schwa in a way that overlaps perceptually with other vowels. Flemming and Johnson (2007) and Flemming (2009) further hypothesized that word-medial schwa is not contrastive with any other vowels and is, therefore, more free to vary. Moreover, word-medial schwa is also shorter in duration compared to other vowels, including word-final schwa, meaning it is naturally more susceptible to coarticulatory pressure. While these factors may be the source of word-medial schwa's variability, the existence of two schwas, one that is variable and one that is more consistent in its quality, suggests the possibility that two sounds, both transcribed as schwa, may differ in their targetedness. Specifically, whereas the more variable word-medial schwa may be targetless, the more consistent word-final schwa may be targeted.

Two other potentially different schwa variants are schwa in function words (e.g., *a*) and schwa in content words (e.g., *accompany*). Previous work (Gick 2002; Lilley 2012) has shown acoustic and articulatory differences between schwas in these two positions. One of the four speakers in Gick's (2002) study produced schwas in function words with their tongue in a resting position and schwas in content words with a retracted tongue root. While this difference was produced by only one speaker, Gick's (2002) results provide a basis for a potential articulatory distinction between schwas in function and content words. Specifically, Gick's (2002) results suggest that schwas in content words are targeted and schwas in function words are targetless. Large-scale acoustic analyses of schwa have also revealed that schwas in function words are produced with a higher F1 and a lower F2 than schwas in content words (Lilley 2012), suggesting that function word and content word schwas have different acoustic targets. Thus, together, these studies suggest that, articulatorily, schwas in function words are targetless (Gick 2002) and produced lower and further back, as indicated acoustically by a higher F1 and a lower F2 (Lilley 2012) than schwas in content words. Schwas in content words are targeted (Gick 2002), and this

articulatory target is higher and further forward than the targetless schwas in function words, as indicated acoustically by a lower F1 and a higher F2 (Lilley 2012).

Furthermore, schwas in function words can occur in minimal pairs with schwas in content words, as exemplified by the contrast between phrase-initial schwa (i.e., the schwa in the indefinite article *a* before a noun, as in *a company*) and word-initial schwa (i.e., a schwa in word-initial position, in its own syllable, in a multi-syllable word, as in *accompany*). The current study examined these phrase-initial and word-initial schwas, which comprise subsets of function word and content word schwas, respectively. According to the CELEX2 database (Baayen et al. 1995), out of 1762 schwa-initial words, approximately 18% of them form a minimal pair with a schwa-initial phrase. Given that phrase-initial schwas comprise a function word and word-initial schwas occur in content words, and that previous research (Gick 2002; Lilley 2012) suggests that schwas in function and content words are articulatorily and acoustically distinct, the same articulatory and acoustic distinctions are predicted for phrase- and word-initial schwas as for function and content word schwas. However, Kim et al. (2012) conducted acoustic analyses of the F1 and F2, the duration, and the amplitude of word- and phrase-initial schwas and found no differences between them. Their analyses were limited to the F1 and F2 at the schwa midpoint and they did not consider the possibility of a difference between word- and phrase-initial schwas in their formant trajectories (Browman and Goldstein 1992; Kondo 1994). Therefore, in the current study, we sought to determine not only whether word- and phrase-initial schwas differ in their duration and their F1 and F2 at temporal midpoint but also whether they differ in their F1 and F2 trajectories. Our ultimate goals were to determine whether speakers distinguish word- and phrase-initial schwas acoustically and, if so, if this distinction reflects differences in the underlying targetedness of word- and phrase-initial schwas.

We predicted that Kim et al.'s (2012) results would replicate and that word- and phrase-initial schwas would not differ in their durations or their midpoint F1 or F2. We then compared formant trajectories of word- and phrase-initial schwas using growth curve analysis (GCA; Mirman et al. 2008). GCA has been used in previous research to model acoustic trajectories (e.g., Chen and Mok 2019, for /r/; Tang et al. 2019, for tone). A benefit to GCA is that it can model data using higher-order polynomials and models with higher-order polynomials can be compared to those with lower-order polynomials. Targetless sounds are characterized by linear interpolation between surrounding sounds (Bates 1995; Keating 1988). Thus, if a model with only linear predictors fits as well as one with quadratic predictors, we would have evidence for targetlessness, whereas if a model with only linear predictors fits less well than one with quadratic predictors, we would have evidence for targetedness. In the current study, we used GCA to determine whether a model with quadratic predictors was significantly different from a model with only linear predictors, allowing us to interpret acoustic targetedness directly.

We predicted that word- and phrase-initial schwas would be distinct in their acoustic realization based on previous work on function vs. content word schwas (Flemming 2009; Flemming and Johnson 2007; Gick 2002; Lilley 2012), given that the word- and phrase-initial schwas in the current study are a subset of content and function word schwas, respectively. Specifically, we predicted that phrase-initial schwa (i.e., in a function word) would be produced lower and further back than word-initial schwa (i.e., in a content word; Lilley 2012). Acoustically, this result would manifest as phrase-initial schwa having a higher F1 and a lower F2 than word-initial schwa. GCA can indicate whether two trajectories differ from one another; if word- and phrase-initial schwas have different targets, then the GCA would indicate that the trajectories are significantly different from one another. We also predicted that word-initial schwa would be targeted and phrase-initial schwa would be targetless (e.g., Gick 2002). If word-initial schwa is targeted then its F1 and F2 trajectories would be significantly different from a linear interpolation between surrounding sounds (i.e., a model with only linear predictors should fit less well than one with quadratic predictors). In addition, if phrase-initial schwa is targetless then its trajectories would not be significantly different from a linear interpolation between surrounding sounds (i.e., a

model with only linear predictors should fit as well as one with quadratic predictors; Bates 1995; Keating 1988).

### 1.2. Eliciting Hyperarticulation of Schwa Targets

The realization of speech sounds is affected by both contextual and lexical factors, such that sounds in some contexts are more likely to reach their phonetic targets than others. For example, words that have low frequencies, high neighborhood densities, and are unpredictable from the sentence context are produced with longer durations and greater articulatory detail than words with high frequencies, low neighborhood densities, and in predictable contexts (for review, see Clopper and Turnbull 2018). Productions that are longer in duration and have greater articulatory detail are relatively hyperarticulated<sup>1</sup> and, therefore, more likely to achieve their phonetic target than productions that are shorter in duration and have less articulatory detail. We, therefore, manipulated two factors in our materials to elicit different degrees of hyperarticulation of word- and phrase-initial schwas and, thereby, facilitate our ability to observe their targets, if any. In particular, we expected to observe greater evidence of targetedness in contexts that elicit greater hyperarticulation relative to those that elicit hypoarticulation and, by comparing contexts with different degrees of hyperarticulation, to obtain stronger evidence for the acoustic targets of word- and phrase-initial schwas, if any.

The first factor we manipulated was whether or not the schwa-initial pair was ambiguous (i.e., a minimal pair) or unambiguous (i.e., not a minimal pair). Ambiguous schwa-initial words and phrases are those that differ only in the placement of the word boundary. For example, the pair *accompany* and *a company* is a minimal pair and, therefore, potentially ambiguous for listeners. Unambiguous schwa-initial words and phrases are those which only have one possible segmentation and are, therefore, not potentially ambiguous for listeners. For example, the schwa-initial word *#accomplish* cannot be segmented as *\*a#complish* and the schwa-initial phrase *a#comic* cannot be segmented as *\*#acomic*. We predicted that schwas in ambiguous pairs would be hyperarticulated (i.e., produced with longer durations and more articulatory detail) compared to unambiguous schwa-initial pairs (Baese-Berk and Goldrick 2009; Buz et al. 2016; Gahl 2015; Munson and Solomon 2004; Wright 2004). Since we predicted that word-initial schwa is targeted, we expected word-initial schwa's target to be enhanced when produced in an ambiguous pair and reduced when produced in an unambiguous pair. Because we predicted that word-initial schwas would have a higher and further forward target compared to phrase-initial schwa, we specifically predicted that word-initial schwa would be produced higher and further forward in ambiguous pairs relative to unambiguous pairs. Since we predicted that phrase-initial schwa is targetless, we expected phrase-initial schwa's quality to remain the same, regardless of whether it is produced in an ambiguous or unambiguous pair. That is, if phrase-initial schwa is targetless, it does not have a target that can be enhanced in an ambiguous pair or reduced in an unambiguous pair and its quality should, therefore, not vary across ambiguous and unambiguous pairs.

The second factor we manipulated was sentence bias; schwa-initial words and phrases were produced in neutral and biased sentences. In neutral sentences, a word with word-initial schwa and a phrase with phrase-initial schwa are equally plausible given the content prior to the critical schwa-initial word/phrase. For example, the phrase *The man was sent to* may be equally likely to continue with either the word *accompany* or the phrase *a company*. Biased sentences are biased towards one of the two segmentations. For example, *The employee was sent to* is biased towards the phrase-initial *a company* relative to the word-initial *accompany*. We predicted that schwas in neutral sentences would be hyperarticulated (i.e., produced with longer duration and more articulatory detail) compared to schwas in biased sentences (Aylett and Turk 2004; Burdin et al. 2015; Calhoun 2010; Clopper and Pierrehumbert 2008; Gahl and Garnsey 2004; Lieberman 1963; Seyfarth 2014). Since we predicted that word-initial schwa is targeted, we expected word-initial schwa's target to be enhanced in neutral contexts relative to biased contexts. Since we predicted that

phrase-initial schwa is targetless, we expected that its acoustic quality would remain the same, regardless of the sentential context. We also predicted that the effects of ambiguity and sentence bias would not interact, based on previous studies that have suggested that factors associated with hyperarticulation are independent in production (Baker and Bradlow 2009; Munson and Solomon 2004).

## 2. Materials and Methods

### 2.1. Participants

Seventy-six undergraduates at The Ohio State University participated in the experiment for partial course credit. Data from non-native speakers of English ( $n = 9$ ) and individuals with a self-reported history of speech or hearing disorders ( $n = 1$ ) were excluded from the analysis. An additional two speakers were excluded from the study due to experimenter error. Thus, data from 64 speakers (44 female, 19 male, and 1 nonbinary) were included in the analysis.

### 2.2. Materials

Ten groups of ambiguous and unambiguous schwa-initial words and phrases were selected for this study and are shown in Table 1. Pairs of ambiguous schwa-initial words and phrases were selected using the stimulus list from Kim et al. (2012) and the Hoosier Mental Lexicon (Nusbaum et al. 1984). Of the ambiguous pairs, nine out of the ten sets were drawn from Kim et al.’s (2012) stimulus materials. Because the schwas in this study were not produced in isolation, they were subject to coarticulatory effects, such as formant transitions between stops and vowels (e.g., Dorman et al. 1977; Liberman et al. 1954). Coarticulatory effects were, therefore, controlled within each stimulus group of ambiguous and unambiguous schwa-initial words and phrases. Specifically, for each ambiguous pair, an unambiguous schwa-initial word and phrase were selected based on how close the following CV sequence was to the ambiguous schwa-initial pair. An unambiguous word or phrase with an identical following CV sequence was preferred. For example, the unambiguous schwa-initial word *adorn* was chosen for the fourth set in Table 1 as it is phonemically similar to the ambiguous schwa-initial pair *adore—a door*. If there was no unambiguous schwa-initial word or phrase with the same following CV sequence as the ambiguous schwa-initial word or phrase, an unambiguous schwa-initial word or phrase with a following consonant with the same place of articulation as the ambiguous pair was selected. If there was no unambiguous schwa-initial word or phrase that contained the same following vowel as the ambiguous schwa-initial word or phrase, an unambiguous schwa-initial word or phrase with the closest following vowel based on distance in the vowel space was selected. Within each stimulus group, the sound preceding schwa was always the same. For example, in the stimulus set consisting of *accompany—a company—accomplish—a comic*, the sound preceding schwa was always /u/ (i.e., *to accompany*).

**Table 1.** Stimulus groups of ambiguous and unambiguous schwa-initial words and phrases.

Ambiguous		Unambiguous	
Word-Initial	Phrase-Initial	Word-Initial	Phrase-Initial
accompany	a company	accomplish	a comic
acquire	a choir	aquatic	a quiet
acute	a cute	acuity	a cube
adore	a door	adorn	a tour
affair	a fair	effect	a fake
allowed	a loud	alarms	a lounge
attuning	a tuning	assuming	a tubing
attacks	a tax	attachment	a tap
aside	a side	asylum	a size
arose	a rose	aroma	a road

A neutral carrier phrase was created for each stimulus group with a preceding sentence context that was identical up until the schwa-initial word/phrase. For example, the stimulus group shown in the first row of Table 1 had the carrier phrase *A man was sent to* for all four words/phrases in the set. A unique biased sentence was created for each item. In the case of the ambiguous schwa-initial words/phrases, the biased sentence was meant to bias the reader towards one segmentation of the ambiguous pair. For example, the carrier phrase *The bodyguard was sent to* would bias towards the #*accompany* segmentation rather than the *a#company* segmentation. For the schwa-initial words and phrases drawn from Kim et al.'s (2012) materials, the neutral and biased sentences were also drawn from Kim et al.'s (2012) materials and modified slightly to ensure that each schwa-initial word or phrase was in the same general location in the sentence and that the neutral sentence frame would be grammatical for the unambiguous schwa-initial words and phrases in the same stimulus group. For unambiguous schwa-initial words and phrases, biased sentences were meant to prime the target words/phrases. For example, the biased equivalent to the neutral *The man was sent to* was *The nerd loved to go to* for the unambiguous phrase *a#comic*. A total of 80 stimulus sentences were created (10 sets  $\times$  4 targets  $\times$  2 sentence contexts). The full sentence list is in Appendix A in Table A1.

Cloze data were collected to confirm that the biased sentence contexts were indeed biased towards the intended word. During a cloze task, participants are presented with a sentence frame and are asked to fill in the remaining part of the sentence. In our cloze task, participants were given the full sentence frame (e.g., *The man was sent to a \_\_\_\_ in New Orleans*) and were asked to report what they thought best completed the sentence. The percentage of responses containing the intended word is the cloze probability of the stimulus word, given the sentence context. We expected the biased sentences to have a higher cloze probability than the neutral sentences. Cloze data were collected from 62 Ohio State University undergraduate students via an online survey. Data from non-native English speakers ( $n = 14$ ) and individuals with a self-reported history of speech or hearing disorders ( $n = 3$ ) were excluded from the analysis. Since each schwa-initial word and phrase occurred twice in our materials, 2 lists were created of 40 sentences each and each participant saw 1 list. Cloze probabilities were calculated by counting the number of participants who responded with the intended word and dividing this count by the number of total responses for that item. A paired sample *t*-test of the cloze data revealed that biased sentences had a significantly higher cloze probability ( $M = 0.16$ ) than neutral sentences ( $M = 0.07$ ,  $t(39) = 1.99$ ,  $p = 0.05$ ), indicating that biased sentences were indeed biased towards the target schwa-initial word or phrase relative to the neutral sentences.

Analyses were conducted to confirm that no significant differences existed between the number of syllables before and after the critical words in neutral/biased sentences, ambiguous/unambiguous sentences, and for word- and phrase-initial schwas. A three-way ANOVA on the number of syllables before the critical schwa-initial word with ambiguity (unambiguous or ambiguous), schwa position (word-initial or phrase-initial), and sentence bias (biased or neutral) as factors revealed no significant main effects or interactions. A three-way ANOVA on the number of syllables after the schwa-initial word with ambiguity, schwa position, and sentence bias as factors revealed a significant main effect of ambiguity ( $F(1, 72) = 8.12$ ,  $p < 0.01$ ). The mean number of syllables after ambiguous schwa-initial pairs ( $M = 4.75$ ,  $SD = 2.00$ ) was significantly smaller than the mean number of syllables after unambiguous schwa-initial pairs ( $M = 6.13$ ,  $SD = 2.26$ ). Differences in the number of syllables before or after the schwa may result in durational differences in the schwa: Participants had a limited amount of time to read the sentences and sentences with a higher syllable count (i.e., longer sentences), therefore, had to be read faster. Thus, the duration of the target schwa in high syllable-count sentences may be shorter than the duration of the target schwa in low syllable-count sentences, leading to differences in vowel quality (Moon and Lindblom 1994). The difference in the number of syllables in each sentence group was handled in the midpoint analyses and GCAs by including schwa duration as a covariate. In the duration analyses, this difference was accounted for using by-item random effects.

Analyses were also conducted to determine whether any differences existed in the word or bigram frequency of the experimental materials. Frequency was obtained using the Corpus of Contemporary American English (COCA), a corpus containing over one billion words from a variety of sources (Davies 2008). A three-way ANOVA on mean log-transformed COCA frequency with ambiguity, schwa position, and sentence bias factors revealed a significant main effect of schwa position ( $F(1, 72) = 7.63, p < 0.01$ ). The mean log-transformed COCA frequency of word-initial schwas ( $M = 8.82, SD = 2.01$ ) was significantly greater than the mean log-transformed COCA bigram frequency of phrase-initial schwas ( $M = 7.63, SD = 1.74$ ). To account for these frequency differences, log-transformed COCA frequency was included as a covariate in all models.

### 2.3. Procedure

Participants were tested individually in a sound-attenuated booth in front of a computer. Participants were told that they would be completing a linguistic experiment studying the production of variation in speech. They were instructed to read aloud the sentences that appeared on the computer screen as if they were reading them to a friend as quickly and as accurately as possible. Each sentence appeared on the screen for 3 s. To signal the end of one trial and the onset of the next trial, participants saw a fixation cross on the screen for 3 s. Participants read each sentence aloud once. The entire experiment took approximately 15 min. Productions were digitally recorded to a Dell Optiplex 7060 Windows computer running Audacity (version 3.2.3) at a sampling rate of 44,100 Hz and 16 bit quantization using a Shure SM58 table-top microphone, positioned approximately 12 inches from the participant's mouth. The sentences were presented using E-Prime experimental software (version 3; Psychology Software Tools, Pittsburgh, PA, USA).

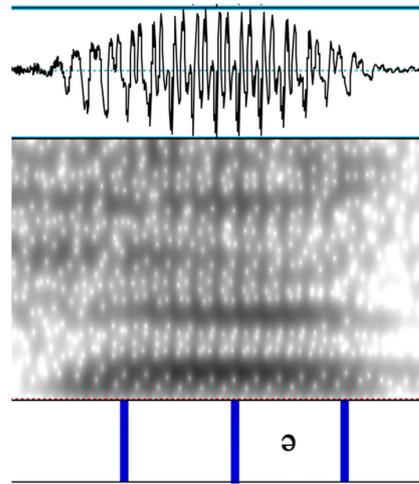
### 2.4. Data Processing

Trials were removed if the participant failed to produce the target schwa (i.e., a schwa was completely absent or the indefinite article *a* was substituted with another word) or the correct word or phrase, leading to the exclusion of 237 trials (5%) across all participants.

Each participant's recording was segmented by hand by the first author in Praat (Boersma and Weenink 2014). In cases where the sound preceding the target schwa was a consonant, the onset of the schwa was identified based on the guidelines set by Peterson and Lehiste (1960). In cases where the sound preceding the schwa was a vowel, the onset of the schwa was marked at the release of a glottal stop, defined as a period of silence preceding a glottal pulse lasting longer than 0.02 s, when present. If no glottal stop was present, the onset of the schwa was identified using the visual appearance of the waveform. The onset of schwa was placed at the point where the amplitude, periodicity (e.g., creaky voice), or shape of the waveform changed. If this point could not be identified, then the boundary was placed at the approximate midpoint of the vowel–schwa sequence. Figure 1 shows an instance where the sound preceding schwa was a vowel but there was no visual change in the waveform, indicating a boundary between the preceding vowel and schwa. In these instances, we assumed that the schwa and the preceding vowel were equivalent in duration and, therefore, placed a boundary at the midpoint of the sound. All schwas were followed by a consonant and the offset of the schwa was identified based on the guidelines set by Peterson and Lehiste (1960).

After the boundaries of all schwas had been identified, a Praat script was used to automatically estimate the first and second formant values at schwa onset, offset, and at 10% increments in between. To account for outliers and formant tracking errors, all formant estimates that were two standard deviations above or below a participants' mean were identified and corrected by hand. Corrections were made by manipulating the formant settings in Praat until the formant tracker was visually aligned with the formants on the spectrogram. The cursor was moved to the temporal point at which the outlier value was detected and the formant value at that point was checked by hand. If the formant value reported by the Praat script was within 10 Hz of the formant value checked by hand, the

original value reported by the Praat script was retained for analysis. If the formant value reported by the Praat script was more than 10 Hz greater or 10 Hz less than the value checked by hand, then the formant value was corrected and replaced with the value checked by hand. All formant estimates were converted to Bark for analysis (Traunmüller 1990).



**Figure 1.** Example of a schwa preceded by a vowel, with no clear delineation between the preceding vowel and schwa. The boundary was placed at the midpoint of the sequence.

### 2.5. Model Building

To replicate the analyses by Kim et al. (2012) and to determine whether word- and phrase-initial schwas differed in their duration or first and second formant frequencies at temporal midpoint, we ran three linear mixed effect regression models. Models were constructed to predict midpoint F1 and F2 in Bark based on schwa position (word- or phrase-initial), ambiguity (ambiguous or unambiguous), and sentence bias (biased or neutral). Schwa duration and the log-transformed COCA frequency of the word (for word-initial schwas) or bigram (for phrase-initial schwas) were included as covariates. A model was also constructed to predict schwa duration based on schwa position (word- or phrase-initial), ambiguity (ambiguous or unambiguous), and sentence bias (biased or neutral). Log-transformed COCA frequency was included as a covariate. All categorical predictors were sum contrast coded. Continuous predictors were centered. The maximal random effect structure included random intercepts for participants and for words/phrases; random by-participant slopes for sentence bias, ambiguity, and schwa position; and random by-word slopes for sentence bias. To avoid overfitting, random slopes were removed based on which element had the lowest variance until the model converged and the maximal data-driven random effects structure had been reached (Bates et al. 2015). Statistical significance was assessed using the *lmerTest* package in R (Kuznetsova et al. 2017).

GCA models were built following the recommendations of Mirman et al. (2008). Orthogonal polynomial values were estimated using the `code_poly()` function of the *gazeR* package in R, which creates orthogonal polynomial-transformed data for use in GCA models (Geller et al. 2020). The time course of formant estimates was modeled using a second-order polynomial with fixed effects of sentence bias (biased or neutral), ambiguity (ambiguous or unambiguous), and schwa position (word- or phrase-initial). Second order polynomials were chosen after visualizing the raw data and seeing that the trajectory of the schwas most resembled the shape of a second-order polynomial (i.e., a parabola). All categorical variables were sum contrast coded. The model also included participant random effects on all time terms; participant by sentence bias, ambiguity, and schwa position random interaction effects; item random effects on all time terms; and item by sentence bias random interaction effects. Word frequency and schwa duration were included as covariates. To avoid overfitting, random slopes were removed based on which element had the lowest variance until the model converged and the maximal data-driven random effect

structure had been reached (Bates et al. 2015). Statistical significance was assessed using the *lmerTest* package in R (Kuznetsova et al. 2017).

The results of the GCA are interpreted as follows: If the intercept term is significant, the average formant frequency is significantly different from 0. If the linear term is significant, the overall slope of the formant trajectory is significantly different from 0 (i.e., horizontal). If the quadratic term is significant, the best-fitting curve is significantly different from a straight line or, in other words, the curve of best fit is significantly more parabolic than linear. A larger absolute quadratic coefficient corresponds to a narrower parabola (i.e., a deeper and more extreme formant trajectory); as the quadratic coefficient approaches zero, the wider and more straight-line-like the curve becomes (i.e., a shallower formant trajectory). The sign of the quadratic coefficient corresponds to whether the curve is u-shaped (positive coefficient) or inverted-u-shaped (negative coefficient). The GCA results are further interpreted based on interactions between the intercept and the polynomial terms (linear and quadratic), on the one hand, and the three independent variables (schwa position, ambiguity, and sentence bias), on the other hand. An interaction between an independent variable and the intercept term indicates a significant difference in the overall mean formant frequency across the levels of the independent variable. An interaction between an independent variable and the linear term indicates a significant difference in the overall slope of the formant trajectory across the levels of the independent variable. Finally, an interaction between an independent variable and the quadratic term indicates a significant difference in the shape of the formant trajectory across the levels of the independent variable.

### 3. Results

#### 3.1. Midpoint and Duration Analyses

The first set of analyses were designed to replicate Kim et al.'s (2012) analysis and involved predicting midpoint F1 and F2 as well as duration from schwa position, ambiguity, and sentence bias. Full summaries of the midpoint F1 and F2 and duration models can be found in Appendix B. Sentence bias was a significant predictor of F1 of schwa at the midpoint [Estimate =  $-0.03$ , SE =  $0.01$ ,  $t(39.79) = -2.899$ ,  $p < 0.01$ ]. The F1 value at the midpoint of schwas in biased sentences was approximately 0.03 Bark lower than the F1 value at the midpoint of schwas in neutral sentences. The covariate of schwa duration was a significant predictor of schwa F1 at the midpoint [Estimate =  $6.634$ , SE =  $0.328$ ,  $t(4709.92) = 20.253$ ,  $p < 0.01$ ] and F2 at the midpoint [Estimate =  $2.63$ , SE =  $0.437$ ,  $t(4744.18) = 6.019$ ,  $p < 0.01$ ]. These results suggest that, as the duration of schwa increases, F1 and F2 at the midpoint of schwa also increase (i.e., schwas are produced lower and further forward with longer duration). No other significant predictors were found for F1 or F2 of schwa at the midpoint. The lack of an effect of schwa position on midpoint F1 and F2 suggests no difference in quality between word- and phrase-initial schwas, as observed by Kim et al. (2012).

Sentence bias was a significant predictor of schwa duration [Estimate =  $-0.002$ , SE =  $0.001$ ,  $t(40.07) = -2.329$ ,  $p < 0.05$ ], indicating that schwas produced in neutral sentences were significantly longer than schwas produced in biased sentences. This duration effect, as well as the main effect of sentence bias on F1 at the midpoint, is consistent with previous work showing that schwas in biased contexts are reduced relative to schwas in neutral contexts (e.g., Aylett and Turk 2004; Burdin et al. 2015; Calhoun 2010; Clopper and Pierrehumbert 2008; Gahl and Garnsey 2004; Lieberman 1963; Seyfarth 2014), although the magnitude of these effects of sentence bias in the current study are small.

#### 3.2. F1 Trajectory

The second set of analyses was designed to explore the effects of schwa position, ambiguity, and sentence bias on the F1 and F2 trajectories of schwa. A summary of the model<sup>2</sup> output for the GCA of the F1 trajectory is shown in Table 2.

**Table 2.** Summary of the GCA model predicting the F1 trajectory from schwa position (word-initial or phrase-initial), sentence bias (neutral or biased), and ambiguity (ambiguous or unambiguous).

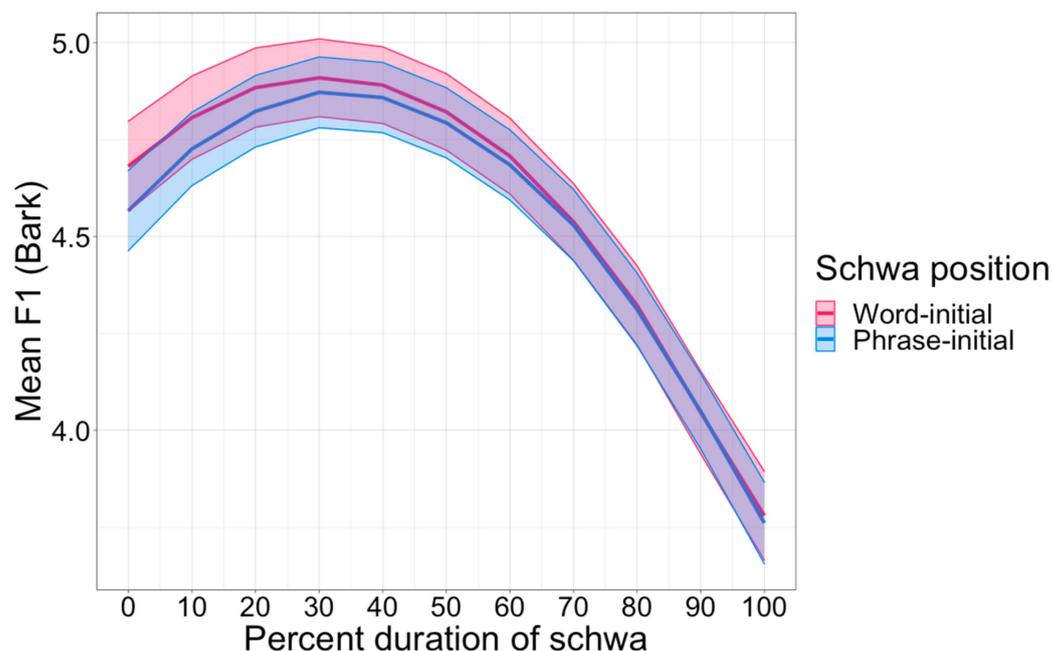
Model Term	Estimate	SE	t-Value	p-Value
Intercept	4.574	0.069	66.728	<0.001
Poly 1 (linear term)	−0.927	0.074	−12.505	<0.001
Poly 2 (quadratic term)	−0.724	0.008	−90.261	<0.001
Bias	−0.029	0.012	−2.477	0.01
Ambiguity	−0.025	0.027	−0.935	0.350
Schwa position	−0.001	0.021	−0.052	0.958
Frequency	−0.011	0.012	−0.882	0.378
Duration	4.366	0.114	38.249	<0.001
Linear × bias	0.031	0.008	3.878	<0.001
Quadratic × bias	0.023	0.008	2.848	<0.01
Linear × ambiguity	−0.010	0.049	−0.212	0.833
Quadratic × ambiguity	−0.008	0.008	−0.944	0.345
Bias × ambiguity	0.008	0.010	0.803	0.426
Linear × schwa position	0.012	0.039	0.304	0.762
Quadratic × schwa position	0.023	0.008	2.861	<0.01
Bias × schwa position	−0.008	0.009	−0.899	0.372
Ambiguity × schwa position	−0.020	0.016	−1.205	0.229
Linear × bias × ambiguity	0.002	0.008	0.288	0.773
Quadratic × bias × ambiguity	0.020	0.008	2.491	0.012
Linear × bias × schwa position	−0.009	0.008	−1.214	0.225
Quadratic × bias × schwa position	0.006	0.008	0.727	0.468
Linear × ambiguity × schwa position	0.004	0.037	0.103	0.918
Quadratic × ambiguity × schwa position	−0.008	0.008	−1.005	0.315
Bias × ambiguity × schwa position	−0.002	0.008	−0.238	0.813
Linear × bias × ambiguity × schwa position	0.018	0.008	2.265	0.023
Quadratic × bias × ambiguity × schwa position	0.003	0.008	0.377	0.706

The quadratic term was significant overall, suggesting targeted schwa, although this effect was mediated by several interactions. To confirm that both schwa variants were targeted, schwa position was treatment contrast coded to test the significance of the quadratic term for each schwa position separately. The quadratic term of the word-initial schwa was significant (Estimate = −0.701, SE = 0.011,  $t(52,834.926) = -61.6, p < 0.001$ ), as was the quadratic term of the phrase-initial schwa (Estimate = −0.747, SE = 0.011,  $t(52,834.942) = -66.063, p < 0.001$ ). These results suggest that both word- and phrase-initial schwas have an F1 target, contrary to our prediction that word-initial schwa would be targeted and phrase-initial schwa would be targetless. At the same time, the significant interaction between schwa position and the quadratic term suggests that their targets differ. Specifically, the F1 trajectory for word-initial schwa is significantly wider than the F1 trajectory for phrase-initial schwa, as shown in Figure 2.

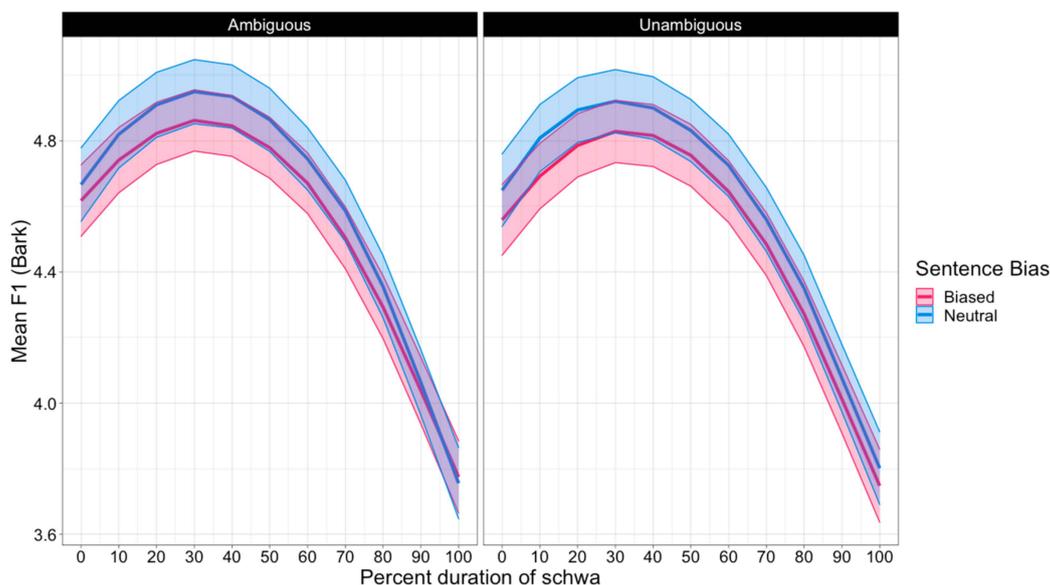
In addition to the primary quadratic effects of interest, there was a significant four-way interaction between schwa position, ambiguity, bias, and the linear term. To unpack this interaction, schwa position, ambiguity, and bias were treatment contrast coded and relevelled. The linear term of the word- and phrase-initial schwas were not significantly different for ambiguous or unambiguous pairs in neutral or biasing sentences. The significant interaction in the full model, therefore, likely results from a modest crossover interaction of schwa position as a function of ambiguity and bias and will not be interpreted further.

There was also a significant three-way interaction of bias, ambiguity, and the quadratic term. To unpack this interaction, bias and ambiguity were treatment contrast coded and relevelled. These comparisons revealed that the shape of the F1 trajectories of ambiguous schwa-initial pairs were narrower in neutral sentences than in biased sentences (Estimate = −0.086, SE = 0.023,  $t(52,834.899) = -3.789, p = 0.001$ ). However, the F1 trajectories of unambiguous pairs were not different in neutral vs. biased sentences (Estimate = −0.006, SE = 0.023,  $t(52,834.918) = -0.251, n.s.$ ). This difference is consistent with hyperarticula-

tion of ambiguous schwas in neutral sentences but not biased sentences; a narrower F1 trajectory is indicative of greater deviation from a straight line and, thus, a more target-like production, as shown in Figure 3. Moreover, this interaction provides further evidence that both word- and phrase-initial schwas are targeted, since a targetless account would not predict any change in formant trajectories based on bias and ambiguity.



**Figure 2.** Grand mean of the first formant trajectories for word-initial schwa (red) and phrase-initial schwa (blue). Ribbons represent the standard error of subject means.



**Figure 3.** Grand mean of the first formant trajectories for ambiguous schwas (left) and unambiguous schwas (right) in biased (red) and neutral (blue) sentence contexts. Ribbons represent the standard error of subject means.

Finally, the significant two-way interaction of sentence bias and the intercept term indicates that schwas produced in biased sentences had a significantly lower F1 than schwas produced in neutral sentences, consistent with the results of the midpoint analysis. In

addition, the significant interaction of sentence bias and the linear term indicates that the slope of the F1 trajectory is less steep for schwas in biased sentences than in neutral sentences. This difference in slope of the F1 trajectory is also consistent with hyperarticulation in neutral sentences, as a steeper slope is indicative of greater deviation from a horizontal line and closer production to a potential target. This effect of sentence bias on hyperarticulation is consistent with previous research (e.g., Aylett and Turk 2004; Burdin et al. 2015; Calhoun 2010; Clopper and Pierrehumbert 2008; Gahl and Garnsey 2004; Lieberman 1963; Seyfarth 2014). All other significant effects and interactions were involved in higher-order interactions and will not be discussed further.

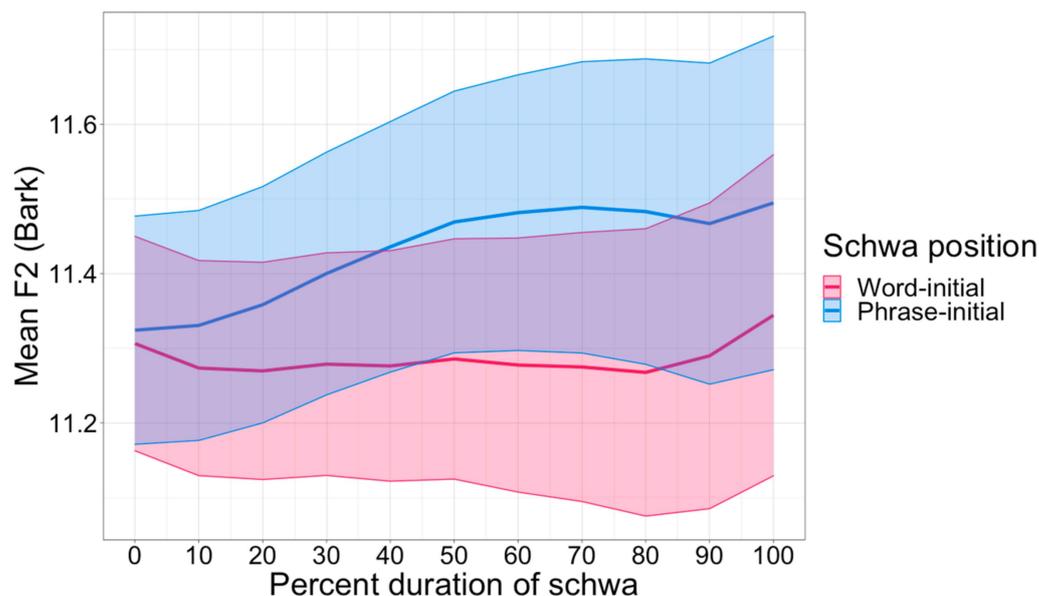
### 3.3. F2 Trajectory

A summary of the model<sup>3</sup> output for the GCA of the F2 trajectory is shown in Table 3.

**Table 3.** Summary of the GCA model predicting the F2 trajectory from schwa position (word-initial or phrase-initial), sentence bias (neutral or biased), and ambiguity (ambiguous or unambiguous).

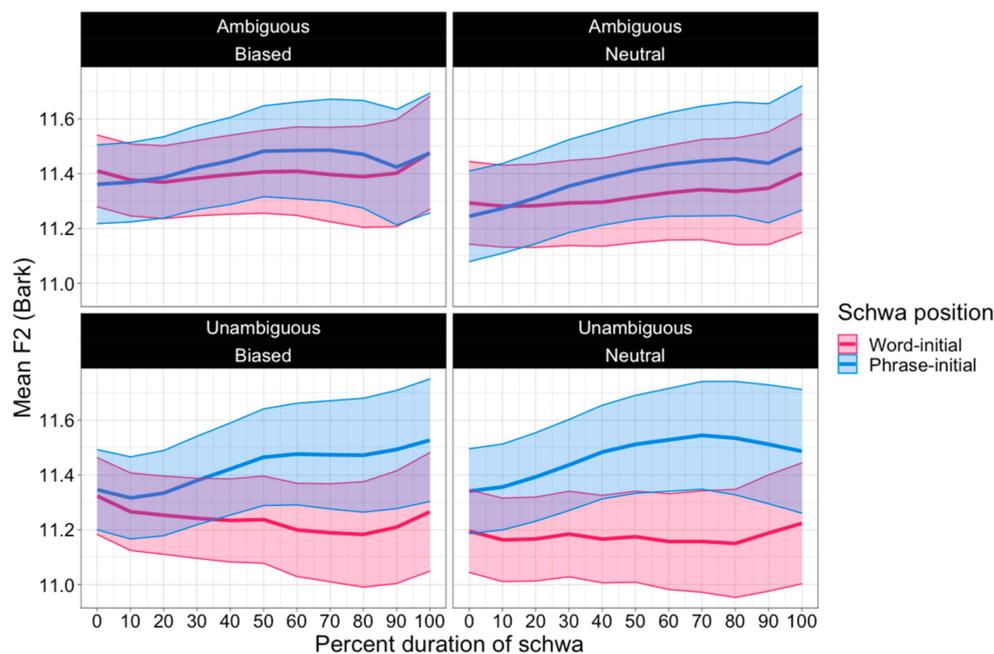
Model Term	Estimate	SE	t-Value	p-Value
Intercept	11.220	0.181	62.081	<0.001
Poly 1 (linear term)	0.106	0.011	9.815	<0.001
Poly 2 (quadratic term)	−0.012	0.011	−1.102	0.271
Bias	0.018	0.0230	0.779	0.440
Ambiguity	−0.052	0.0427	−1.224	0.221
Schwa position	−0.032	0.036	−0.892	0.373
Frequency	0.038	0.019	2.010	0.044
Duration	1.985	0.153	13.003	<0.001
Linear × bias	−0.031	0.011	−2.841	0.005
Quadratic × bias	0.014	0.011	1.300	0.194
Linear × ambiguity	0.022	0.011	2.058	0.040
Quadratic × ambiguity	−0.006	0.011	−0.545	0.586
Bias × ambiguity	0.042	0.020	2.140	0.037
Linear × schwa position	−0.084	0.011	−7.844	<0.001
Quadratic × schwa position	0.058	0.011	5.401	<0.001
Bias × schwa position	−0.013	0.016	−0.805	0.423
Ambiguity × schwa position	−0.017	0.025	−0.680	0.497
Linear × bias × ambiguity	−0.014	0.011	−1.314	0.187
Quadratic × bias × ambiguity	−0.017	0.011	−1.560	0.119
Linear × bias × schwa position	−0.004	0.011	−0.373	0.709
Quadratic × bias × schwa position	−0.006	0.011	−0.555	0.579
Linear × ambiguity × schwa position	0.033	0.011	3.103	0.002
Quadratic × ambiguity × schwa position	−0.006	0.011	−0.608	0.543
Bias × ambiguity × schwa position	−0.001	0.015	−0.046	0.963
Linear × bias × ambiguity × schwa position	0.022	0.011	2.047	0.040
Quadratic × bias × ambiguity × schwa position	0.011	0.011	1.041	0.298

Unlike for F1, the overall quadratic term was not significant for F2. However, the significant interaction between schwa position and the quadratic term indicates that the trajectories of word- and phrase-initial schwas were significantly different from one another, as shown in Figure 4. To unpack this interaction, schwa position was treatment contrast coded. The quadratic term was significant for both word-initial schwa (Estimate = 0.046, SE = 0.015,  $t(53,004.731) = 3.03$ ,  $p < 0.01$ ) and phrase-initial schwa (Estimate = −0.070, SE = 0.015,  $t(53,004.740) = −4.613$ ,  $p < 0.01$ ), indicating that the F2 trajectory of both schwas was significantly different from the linear interpolation between surrounding sounds and suggesting that both word-initial and phrase-initial schwas have an F2 target. The lack of an overall effect of the quadratic terms reflects the different targets for the two schwas: the word-initial schwa’s target is lower than that of the phrase-initial schwa and, thus, the F2 curve of word-initial schwa is u-shaped, whereas the curve of the phrase-initial schwa is inverted-u-shaped.



**Figure 4.** Grand mean of the second formant trajectories for word-initial schwa (red) and phrase-initial schwa (blue). Ribbons represent the standard error of subject means.

In addition to the primary quadratic effects of interest, there was a significant four-way interaction between schwa position, ambiguity, bias, and the linear term. To unpack this interaction, schwa position, ambiguity, and bias were treatment contrast coded and revealed. These comparisons revealed that the slope of the trajectory of word-initial schwa was significantly less steep than the slope of phrase-initial schwa in unambiguous pairs in both neutral (Estimate = 0.183, SE = 0.043,  $t(53,004.706) = 2.246$ ,  $p < 0.001$ ) and biased sentences (Estimate = 0.288, SE = 0.043,  $t(53,004.722) = 6.633$ ,  $p < 0.001$ ) and in ambiguous pairs in neutral sentences (Estimate = 0.138, SE = 0.043,  $t(53,004.733) = 3.223$ ,  $p = 0.001$ ). However, the slope of the trajectory of word-initial schwas was not significantly different than the slope of phrase-initial schwas in ambiguous pairs in biased sentences (Estimate = 0.066, SE = 0.043,  $t(53,004.73) = 1.544$ , *n.s.*). These results suggest that the effect of bias goes in the expected direction in the case of ambiguous schwa-initial pairs, with hyperarticulation of schwa-initial minimal pairs for neutral, but not for biased, sentences. As shown in Figure 5, the overall slopes of the F2 trajectories are very similar for word- and phrase-initial schwas in ambiguous pairs in biased sentences, but they are more different from one another in ambiguous pairs in neutral sentences, consistent with enhancement of the targets in neutral vs. biased sentences. In the case of unambiguous schwa-initial pairs, however, word- and phrase-initial schwas are produced with significantly different overall slopes and closer to their respective targets in all sentences, regardless of sentence bias. Thus, sentence bias did not lead to target enhancement for unambiguous pairs. The effect of ambiguity is also unexpected, given that reduction of the targets is predicted for unambiguous pairs relative to ambiguous pairs. Unambiguous schwa-initial pairs are produced significantly differently from one another, regardless of whether they are in neutral or biased sentences, whereas ambiguous schwa-initial pairs are only distinguished in neutral sentences. These results do, however, suggest that both of the variants are targeted, as both are affected by the contextual manipulation.



**Figure 5.** Grand mean of the second formant trajectories for ambiguous (**top**) and unambiguous (**bottom**) word-initial schwa (red) and phrase-initial schwa (blue) in biased (**left**) and neutral (**right**) sentence contexts. Ribbons represent the standard error of subject means.

Finally, to unpack the interaction between bias, ambiguity, and the intercept term, bias and ambiguity were treatment contrast coded and relevelled. These comparisons revealed that the mean F2 of ambiguous schwa-initial words and phrases was significantly lower in neutral compared to biased sentences (Estimate = 0.06, SE = 0.03,  $t(46.135) = 1.979$ ,  $p = 0.05$ ). However, unambiguous schwa-initial words and phrases were not produced differently in neutral and biased sentences (Estimate =  $-0.024$ , SE = 0.03,  $t(46.31) = -0.795$ , *n.s.*). This difference is again consistent with hyperarticulation of ambiguous schwas in neutral sentences relative to biased sentences. All other significant effects and interactions were involved in higher-order interactions and, as such, they will not be discussed further or interpreted on their own.

#### 4. Discussion

The primary goals of the current study were to determine whether speakers distinguish word- and phrase-initial schwas acoustically and, if so, if this distinction reflects differences in the underlying targetedness of word- and phrase-initial schwas. To promote hyperarticulation of a potential schwa target, we compared ambiguous and unambiguous word- and phrase-initial schwas produced in neutral and biased sentences.

##### 4.1. Targetedness of Word-Initial and Phrase-Initial Schwas in English

Previous studies that analyzed schwa-initial minimal pairs' F1 and F2 at the temporal midpoint failed to find any differences in the quality of word- and phrase-initial schwas (Kim et al. 2012). We replicated these findings with our own midpoint analysis, which also did not find a difference in the F1 and F2 at the temporal midpoint of word- and phrase-initial schwas.

However, our results diverge from previous studies on schwa-initial minimal pairs (Kim et al. 2012) because we included an analysis of the formant trajectories (Browman and Goldstein 1992; Kondo 1994). We found that the F1 and F2 trajectories of word- and phrase-initial schwas were significantly different from the linear interpolation between surrounding sounds, suggesting that both word- and phrase-initial schwas are targeted (Bates 1995; Keating 1988). Specifically, while the overall mean F1 value was the same for word- and phrase-initial schwas, the F1 trajectories for word-initial schwas were significantly

steeper and wider than the F1 trajectories of phrase-initial schwas. These results suggest that word- and phrase-initial schwas have different targets along F1. Similarly, for F2, the coefficient of the quadratic term for the phrase-initial schwa is negative and the coefficient for the word-initial schwa is positive, indicating that the F2 target of phrase-initial schwa is higher than the F2 target of word-initial schwa. Thus, our results show that word- and phrase-initial schwas have different acoustic targets. This targetedness of both word- and phrase-initial schwas suggests that they are underlyingly different sounds, both of which have an underlying phonetic specification (i.e., Keating 1988). However, these acoustic differences between word- and phrase-initial schwas are small, along both F1 and F2. Further research is needed to determine whether these small differences between word- and phrase-initial schwas are perceptible to listeners.

The evidence for targetedness in both word- and phrase-initial schwas does not align with our predictions based on Gick (2002) that word-initial schwa would be targeted and phrase-initial schwa would be targetless. Moreover, our results are contrary to our predictions based on Gick (2002) and Lilley (2012) that the target of phrase-initial schwa would have a higher F1 and a lower F2 than that of word-initial schwa. Rather, we found word- and phrase-initial schwas had approximately equivalent F1 targets, and phrase-initial schwa's F2 target was higher than word-initial schwa's F2 target. This discrepancy may be due to our focus on phrase-initial schwa, specifically, rather than schwas in function words in general, as in previous studies (Gick 2002; Lilley 2012). Our analysis was also limited to acoustic measures, whereas Gick (2002) examined articulatory data. A replication of our study with articulatory measures is needed to confirm our findings.

Regarding duration, while the duration analysis did not reveal a difference between the durations of word- and phrase-initial schwas, the duration covariate was significant for both the F1 and F2 midpoint analyses and for the GCAs. As duration increased, the average F1 and F2 also increased, suggesting lowering and fronting with longer duration. Flemming (2009) suspected that a reason for schwa's contextual variability is its short duration, relative to other vowels. If schwa is particularly short the speaker may not have enough time to successfully articulate the target in the time given. Thus, we would expect that the duration of schwa would affect the quality of schwa. In the current study, word- and phrase-initial schwas are not distinguished by duration, although duration is one of multiple related factors that affects the acoustic quality of both word-initial and phrase-initial schwas, as predicted for targeted sounds (Barry 1998; Cohen Priva and Strand 2023; Flemming 2009).

Our interpretation of the significant quadratic terms as evidence of targetedness reflects prior claims that targetless sounds will have a linear acoustic interpolation from surrounding segments (Bates 1995; Keating 1988). However, research on vowel inherent spectral change (VISC) has demonstrated that some spectral change is inherent to the identity of vowel categories (Hillenbrand 2013). In particular, steady-state synthesized vowel tokens (i.e., tokens containing vowels synthesized to have steady formants) have been shown to be identified less accurately than naturally produced tokens, suggesting that spectral change across the duration of a vowel is important in the perceptual identification of vowels (Hillenbrand and Gayvert 1993). Moreover, including time-varying spectral information has been shown to improve the separation of vowels in discriminant analyses (Hillenbrand et al. 1995; Hillenbrand et al. 2001). Previous VISC research has not included schwa as a vowel of study and, thus, more research is needed to determine whether the spectral changes we observed in word- and phrase-initial schwas in the current study should be characterized as VISC and, if so, how VISC is related to notions of acoustic targetedness and phonetic (un)specification.

#### 4.2. Hyperarticulation of Word-Initial and Phrase-Initial Schwas in English

To promote hyperarticulation of potential schwa targets, we compared ambiguous and unambiguous word- and phrase-initial schwas produced in neutral and biased sentences. We predicted that if schwa is targeted, its target would be enhanced when it is produced in

an ambiguous pair relative to an unambiguous pair and in neutral sentences relative to biased sentences. Since we found that both word- and phrase-initial schwas are targeted along F1 and F2, we would expect to observe these patterns of target enhancement for both schwa types. Indeed, along F1 we find evidence of hyperarticulation of word- and phrase-initial schwas in ambiguous pairs relative to unambiguous pairs. When word- and phrase-initial schwas in ambiguous pairs were produced in neutral sentences, they were produced with a significantly narrower F1 trajectory relative to their productions in biased sentences. This narrower F1 trajectory indicates greater deviation from the straight-line interpolation and, thus, a more target-like production (i.e., hyperarticulation). In addition, word- and phrase-initial schwas were produced with a significantly higher F1 in neutral compared to biased sentences, and the F1 trajectories of schwas produced in neutral sentences were significantly steeper than in biased sentences. As the F1 target is high for both schwas, the higher F1 and steeper slope in neutral sentences is reflective of hyperarticulation. Along F2, we found that word- and phrase-initial schwas in ambiguous pairs were produced with a significantly lower F2 in neutral compared to biased sentences. This lower F2 is consistent with hyperarticulation of the lower word-initial schwa target, although it does not clearly reflect hyperarticulation of the higher phrase-initial schwa target. We also found that the slope of the F2 trajectories of word-initial schwas were significantly less steep than that of phrase-initial schwas in unambiguous pairs in both sentence contexts and in ambiguous pairs in neutral sentences. As the target of word-initial schwa is lower than the target of phrase-initial schwa, the decreasing slope of word-initial schwa is reflective of hyperarticulation. Thus, the contextual manipulations led to hyperarticulation of both word- and phrase-initial schwas in F1 and of word-initial schwas in F2. This hyperarticulation provides further evidence that both word-initial and phrase-initial schwas are targeted, because the quality of targetless sounds should not be affected by context.

We also predicted that the effects of ambiguity and sentence bias would not interact. This prediction is based on previous effects in production which have suggested that factors associated with hyperarticulation are independent (Baker and Bradlow 2009; Munson and Solomon 2004). Contrary to this prediction, we found a significant interaction between bias and ambiguity, such that the F1 trajectory of ambiguous schwas in neutral sentences was significantly narrower than the F1 trajectory of ambiguous schwas in biased sentences. However, we found no difference in the F1 trajectory of unambiguous schwas in neutral and biased sentences. This interaction suggests maximum enhancement of ambiguous schwa targets in neutral sentences, a super-additive effect of ambiguity and bias. This observed interaction may be a result of the factors we used to promote hyperarticulation. While previous studies (Baker and Bradlow 2009; Munson and Solomon 2004) have examined the effects of word-level, stylistic, and local factors, such as word frequency and neighborhood density, the current study manipulated ambiguity (a word-level factor) and bias (a sentence-level factor).

#### 4.3. Conclusions

Our results demonstrate that word-initial and phrase-initial schwas are both targeted. In addition, the targets for these two schwas are acoustically different, suggesting that the underlying representations of schwas in function words and content words are different from one another. The analysis of formant trajectories, rather than formant values at the vowel midpoint (e.g., Kim et al. 2012), allowed us to explicitly examine targetedness and to observe that word- and phrase-initial schwas have different targets along F1 and F2. In addition, by manipulating factors that promote hyperarticulation, we obtained converging evidence for the word- and phrase-initial schwa targets.

**Author Contributions:** Conceptualization, E.R.N. and C.G.C.; methodology, E.R.N. and C.G.C.; formal analysis, E.R.N.; investigation, E.R.N.; writing—original draft preparation, E.R.N.; writing—review and editing, E.R.N. and C.G.C.; visualization, E.R.N.; supervision, C.G.C.; project administration, E.R.N. and C.G.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and was determined to be exempt by the Institutional Review Board of The Ohio State University (2021E0986 9/21/2021).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors on request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A

**Table A1.** Full sentence list.

Ambiguity	Word/Phrase	Neutral Sentence	Biased Sentence
Ambiguous	accompany a company	The man was sent to accompany a young lady. The man was sent to a company in New Orleans.	The bodyguard had to accompany a celebrity. Workers are drawn to a company of high standards.
Unambiguous	accomplish a comic	The man was sent to accomplish a lofty goal. The man was sent to a comic shop in New York.	The goal he set out to accomplish was finally complete. The nerd loved to go to a comic store around the corner.
Ambiguous	acquire a choir	John was sent to acquire new skills. John was sent to a choir near Kentucky.	The collectors wanted to acquire new items. The singers belong to a choir near Kentucky.
Unambiguous	aquatic a quiet	John was sent to aquatic nursing school. John was sent to a quiet nation for work.	The fish were brought to aquatic nurseries to grow. The librarian walked to a quiet nearby room.
Ambiguous	acute a cute	The teenaged girl had acute kidney disease. The teenaged girl had a cute kitten in her arms.	The doctor said he had acute colon cancer. The girl had a cute collection of dolls.
Unambiguous	acuity a cube	The teenaged girl had acuity kids did not usually have. The teenaged girl had a cube kept on her dresser.	The eye doctor said he had acuity comparable to a teenager. The engineer had a cube collection on their desk.
Ambiguous	adore a door	The servant came to adore every puppy. The servant came to a door in the basement.	Lovers are meant to adore each other. The hallway leads to a door at the end.
Unambiguous	adorn a tour	The servant came to adorn the crown with jewels. The servant came to a tour of the house.	The jewels were used to adorn the queen’s crown. The band planned to do a tour of the world.
Ambiguous	affair a fair	She didn’t know what affair Dave was involved in. She didn’t know what a fair deal would be.	The adulterer asked what affair Diane was talking about. The judge should know what a fair deal should be.
Unambiguous	effect a fake	She didn’t know what effect Dean would have on the project. She didn’t know what a fake dealer may sell her.	The scientist should know what effect deer have on mice populations. A scammer must know what a fake deed looks like.
Ambiguous	allowed a loud	I think Janet might have allowed Sue to go. I think Janet might have a loud singing voice.	The government has allowed so much corruption. The rock singer has a loud sound system.
Unambiguous	alarms a lounge	I think Janet might have alarms so that she wakes up. I think Janet might have a lounge space in her house.	The firehouse has alarms sounding constantly. The night club has a lounge so clients can escape the noise.

**Table A1.** *Cont.*

Ambiguity	Word/Phrase	Neutral Sentence	Biased Sentence
Ambiguous	attuning	The young man recently heard attuning various senses would help his spy career.	The spy would need to stop attuning his senses to this scenario.
	a tuning	The young man recently heard a tuning violinist in the distance.	The musician should stop a tuning fork from playing that note.
Unambiguous	assuming	The young man recently heard assuming violence would occur is bad.	The judge would need to stop assuming both parties were being honest.
	a tubing	The young man recently heard a tubing venue was to be built nearby.	The skiers wanted to stop a tubing hill from being built.
Ambiguous	attacks	They always claim that attacks largely happen at night.	It's a vicious bear that attacks like lightning.
	a tax	They always claim that a tax levy would help.	The IRS told us that a tax law had changed.
Unambiguous	attachment	They always claim that attachment like that makes team building easier.	The mother said that attachment like the one to her son kept her going.
	a tap	They always claim that a tap lightly on the shoulder could wake her up.	The fighter said that a tap leveled on the head would knock someone out.
Ambiguous	aside	Meghan quickly took aside the children under 10.	The trainer took aside the boxer and yelled at him.
	a side	Meghan quickly took a side that the others disagreed with.	The president took a side that was popular in debates.
Unambiguous	asylum	Meghan quickly took asylum there in a different country.	The refugee took asylum that was offered in the country.
	a size	Meghan quickly took a size that was too big.	The clerk took a size that was too small off the rack.
Ambiguous	arose	Michael said the thought that arose could not have been more brilliant.	The film had zombies that arose quickly from the dead.
	a rose	Michael said the thought that a rose could bloom here is strange.	No flower says the things a rose could say.
Unambiguous	aromas	Michael said the thought that aromas could be this bad was surprising.	The old bakery had aromas that are second to none.
	a road	Michael said the thought that a road could be this bumpy is ridiculous.	The potholes that filled a road says a lot about the government.

## Appendix B

**Table A2.** F1 linear model.

Model Term	Estimate	SE	t-Value	p-Value
Intercept	4.8122	0.069	69.588	<0.001
Bias	-0.0300	0.010	-2.899	<0.01
Ambiguity	0.0051	0.040	0.130	0.897
Schwa position	-0.0156	0.035	-0.442	0.659
Duration	6.6345	0.328	20.253	<0.001
Frequency	0.0150	0.020	0.738	0.463
Bias × ambiguity	-0.0024	0.010	-0.235	0.816
Bias × schwa position	-0.0003	0.010	-0.033	0.974
Ambiguity × schwa position	-0.0214	0.031	-0.698	0.0487
Bias × ambiguity × schwa position	0.0015	0.010	0.150	0.881

Table A3. F2 linear model.

Model Term	Estimate	SE	t-Value	p-Value
Intercept	11.372	0.173	65.908	<0.001
Bias	0.023	0.019	1.196	0.239
Ambiguity	−0.054	0.100	−0.538	0.591
Schwa position	−0.025	0.073	−0.338	0.736
Duration	2.630	0.437	6.019	<0.001
Frequency	−0.033	0.045	−0.737	0.461
Bias × ambiguity	0.023	0.018	1.290	0.204
Bias × schwa position	0.011	0.017	0.674	0.504
Ambiguity × schwa position	−0.021	0.062	−0.348	0.728
Bias × ambiguity × schwa position	−0.011	0.017	−0.679	0.500

Table A4. Duration linear model.

Model Term	Estimate	SE	t-Value	p-Value
Intercept	0.0636	0.002	31.889	<0.001
Bias	−0.0016	0.001	−2.329	0.025
Ambiguity	0.0004	0.001	0.276	0.784
Schwa position	−0.002	0.001	−1.571	0.119
Frequency	−0.0004	0.001	−0.485	0.629
Bias × ambiguity	−0.0001	0.001	−0.216	0.830
Bias × schwa position	−0.0002	0.001	−0.452	0.653
Ambiguity × schwa position	−0.0004	0.001	−0.315	0.753
Bias × ambiguity × schwa position	−0.0002	0.001	−0.302	0.764

## Notes

- <sup>1</sup> We assume a continuum in production from hyperarticulation (i.e., enhancement) to hypoarticulation (i.e., reduction). The manipulations in the current study were intended to affect the relative degree of hyperarticulation (more vs. less) along this continuum.
- <sup>2</sup> The model specification for the GCA for F1 was as follows:  $f1.bark \sim (\text{poly1} + \text{poly2}) * \text{bias} * \text{ambiguity} * \text{schwa.position} + \text{frequency} + \text{duration} + (\text{poly1} | \text{subject}) + (1 | \text{subject:bias}) + (1 | \text{subject:ambiguity}) + (1 | \text{subject:schwa.position}) + (\text{poly1} | \text{word}) + (1 | \text{word:bias})$ .
- <sup>3</sup> The model specification for the GCA for F2 was as follows:  $f2.bark \sim (\text{poly1} + \text{poly2}) * \text{bias} * \text{ambiguity} * \text{schwa.position} + \text{frequency} + \text{duration} + (1 | \text{subject}) + (1 | \text{subject:bias}) + (1 | \text{subject:schwa.position}) + (1 | \text{word}) + (1 | \text{word:bias})$ .

## References

- Aylett, Matthew, and Alice Turk. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47: 31–56. [CrossRef] [PubMed]
- Baayen, R. Harald, Richard Piepenbrock, and Leon Gulikers. 1995. CELEX2 LDC96L14 [Web Downloaded]. *Linguistics Data Consortium*. [CrossRef]
- Baese-Berk, Melissa, and Matthew Goldrick. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24: 527–54. [CrossRef] [PubMed]
- Baker, Rachel E., and Ann R. Bradlow. 2009. Variability in word duration as a function of probability, speech style, and prosody. *Language and Speech* 52: 391–413. [CrossRef] [PubMed]
- Bakst, Sarah, and Caroline A. Niziolek. 2021. Effects of syllable stress in adaptation to altered auditory feedback in vowels. *Journal of the Acoustical Society of America* 149: 708–19. [CrossRef] [PubMed]
- Barry, William J. 1998. Time as a factor in the acoustic variation of schwa. Paper presented at the 5th International Conference on Spoken Language Processing, Sydney, Australia, November 30–December 4; pp. 3071–74. [CrossRef]
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]
- Bates, Sally A. R. 1995. Towards a Definition of Schwa: An Acoustic Investigation of Vowel Reduction in English. Doctoral thesis, University of Edinburgh, Edinburgh, UK.
- Boersma, Paul, and David J. M. Weenink. 2014. Praat: Doing Phonetics by Computer (5.3.84) [Computer Software]. Available online: <http://www.praat.org/> (accessed on 4 October 2020).
- Booji, Geert. 1995. *The Phonology of Dutch*. Oxford: Oxford University Press.

- Browman, Catherine P., and Louis Goldstein. 1992. "Targetless" schwa: An articulatory analysis. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, pp. 26–56. [CrossRef]
- Burdin, Rachel S., Rory Turnbull, and Cynthia G. Clopper. 2015. Interactions among lexical and discourse characteristics in vowel production. *Proceedings of Meetings on Acoustics* 22: 060005. [CrossRef]
- Buz, Esteban, Michael K. Tenenhaus, and T. Florian Jaeger. 2016. Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language* 89: 68–86. [CrossRef]
- Calhoun, Sasha. 2010. How does informativeness affect prosodic prominence? *Language and Cognitive Processes* 27: 1099–140. [CrossRef]
- Chen, Shuwen, and Peggy P. K. Mok. 2019. Speech production of rhotics in highly proficient bilinguals: Acoustic and articulatory measures. Paper presented at the 19th International Congress of Phonetic Sciences, Melbourne, Australia, August 5–9; pp. 1818–22.
- Clopper, Cynthia G., and Janet B. Pierrehumbert. 2008. Effects of semantic predictability and regional dialect on vowel space reduction. *Journal of the Acoustical Society of America* 124: 1682–88. [CrossRef] [PubMed]
- Clopper, Cynthia G., and Rory Turnbull. 2018. Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors. In *Rethinking Reduction: Interdisciplinary Perspectives on Conditions, Mechanisms, and Domains for Phonetic Variation*. Berlin: De Gruyter Mouton, pp. 25–72. [CrossRef]
- Cohen Priva, Uriel, and Emily Strand. 2023. Schwa's duration and acoustic position in American English. *Journal of Phonetics* 96: 101198. [CrossRef]
- Davies, Mark. 2008. *The Corpus of Contemporary American English (COCA)*. Available online: <https://www.english-corpora.org/coca/> (accessed on 20 November 2022).
- Dorman, Michael F., Michael Studdert-Kennedy, and Lawrence J. Raphael. 1977. Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception and Psychophysics* 22: 109–22. [CrossRef]
- Flemming, Edward, and Stephanie Johnson. 2007. Rosa's roses: Reduced vowels in American English. *Journal of the International Phonetic Association* 37: 83–96. [CrossRef]
- Flemming, Edward. 2009. The phonetics of schwa vowels. *Phonological Weakness in English* 493: 78–95. [CrossRef]
- Gahl, Susanne, and Susan M. Garnsey. 2004. Knowledge of Grammar, Knowledge of Usage: Syntactic Probabilities Affect Pronunciation Variation. *Language* 80: 748–75. [CrossRef]
- Gahl, Susanne. 2015. Lexical competition in vowel articulation revisited: Vowel dispersion in the Easy/Hard database. *Journal of Phonetics* 49: 96–116. [CrossRef]
- Geller, Jason, Matthew B. Winn, Tristian Mahr, and Daniel Mirman. 2020. GazeR: A Package for Processing Gaze Position and Pupil Size Data. *Behavior Research Methods* 52: 2232–55. [CrossRef]
- Gick, Bryan. 2002. An X-ray investigation of pharyngeal constriction in American English schwa. *Phonetica* 59: 38–48. [CrossRef]
- Hillenbrand, James, and Robert T. Gayvert. 1993. Identification of steady-state vowels synthesized from the Peterson and Barney measurements. *Journal of the Acoustical Society of America* 94: 668–74. [CrossRef] [PubMed]
- Hillenbrand, James, Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97: 3099–111. [CrossRef] [PubMed]
- Hillenbrand, James, Michael J. Clark, and Terrance M. Nearey. 2001. Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America* 109: 748–63. [CrossRef] [PubMed]
- Hillenbrand, James. 2013. Static and Dynamic Approaches to Vowel Perception. In *Vowel Inherent Spectral Change*, 1st ed. Berlin: Springer, pp. 9–30.
- Keating, Patricia A. 1988. Underspecification in Phonetics. *Phonology* 5: 275–92. [CrossRef]
- Kim, Dahee, Joseph D. W. Stephens, and Mark A. Pitt. 2012. How does context play a part in splitting words apart? Production and perception of word boundaries in casual speech. *Journal of Memory and Language* 66: 509–29. [CrossRef] [PubMed]
- Kondo, Yuko. 1994. Targetless schwa: Is that how we get the impression of stress-timing in English? Paper presented at the Edinburgh Linguistics Department Conference '94, Edinburgh, UK, May 26–27; pp. 63–76.
- Koopmans-Van Benium, Florian J. 1994. What's in a schwa? Durational and spectral analysis of natural continuous speech and diphones in Dutch. *Phonetica* 51: 68–79. [CrossRef]
- Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [CrossRef]
- Liberman, Alvin M., Pierre C. Delattre, Franklin S. Cooper, and Louis J. Gerstman. 1954. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied* 68: 1–13. [CrossRef]
- Lieberman, Phillip. 1963. Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech. *Language and Speech* 6: 172–87. [CrossRef]
- Lilley, Jason. 2012. The Characterization of Phonetic Variation in American English Schwa Using Hidden Markov Models. Doctoral thesis, University of Delaware, Newark, DE, USA.
- Mirman, Daniel, James A. Dixon, and James S. Magnuson. 2008. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language* 59: 475–94. [CrossRef]
- Moon, Seung-Jae, and Björn Lindblom. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* 96: 40–55. [CrossRef]
- Munson, Benjamin, and Nancy P. Solomon. 2004. The Effect of Phonological Neighborhood Density on Vowel Articulation. *Journal of Speech, Language, and Hearing Research* 47: 1048–58. [CrossRef] [PubMed]

- Nusbaum, Howard, David B. Pisoni, and Christopher K. Davis. 1984. Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. In *Research on Speech Perception Progress Report No. 10*. Bloomington: Speech Research Laboratory, Indiana University, vol. 10, pp. 357–76.
- Peterson, Gordon E., and Ilse Lehiste. 1960. Duration of Syllable Nuclei in English. *Journal of the Acoustical Society of America* 32: 693–703. [\[CrossRef\]](#)
- Seyfarth, Scott. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition* 133: 140–55. [\[CrossRef\]](#)
- Tang, Ping, Ivan Yuen, Nan X. Rattanasone, Liqun Gao, and Katherine Demuth. 2019. Acquisition of weak syllables in tonal languages: Acoustic evidence from neutral tone in Mandarin Chinese. *Journal of Child Language* 46: 24–50. [\[CrossRef\]](#) [\[PubMed\]](#)
- Traunmüller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America* 88: 97–100. [\[CrossRef\]](#)
- van Bergem, Dick R. 1994. A model of coarticulatory effects on the schwa. *Speech Communication* 14: 143–62. [\[CrossRef\]](#)
- Wiese, Richard. 1986. Schwa and the structure of words in German. *Linguistics* 24: 697–724. [\[CrossRef\]](#)
- Wright, Richard. 2004. Factors of lexical competition in vowel articulation. In *Phonetic Interpretation: Papers in Laboratory Phonology, VI*. Cambridge: Cambridge University Press, pp. 75–87. [\[CrossRef\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.