Spring 5-2-2014

The New Normal: Goodness Judgments of Non-Invariant Speech

Julia R. Drouin
University of Connecticut, julia.drouin@uconn.edu

The New Normal: Goodness Judgments of Non-Invariant Speech

Julia Drouin

University of Connecticut

Honors Thesis

May 2, 2014

Acknowledgements

Gratitude is extended to Dr. Emily B. Myers and Dr. Rachel M. Theodore at the University of Connecticut for their help and guidance as advisors and mentors to this project. This research was supported by the Office of Undergraduate Research at the University of Connecticut, in the form of a Social Sciences, Humanities, and Arts Research Experience (SHARE) Award.

Abstract

Previous research has found that perceptual learning, or normalizing the idiosyncratic phonemes of speech, causes a shift in speech sound category boundaries. The present study examined if perceptual learning was limited to the boundary or if also caused a shift in internal category structure. Seventeen individuals participated in three behavioral tasks to explicate this question. In the Lexical Decision task, participants were trained in either /s/-biasing or /□/- biasing context. In the Goodness Judgment task, participants rated a continuum of sounds on perceived /s/ goodness using a designated scale. Finally, in the Phoneme Identification task, participants listened to the same continuum previously heard but were asked to classify the token as /s/ or /□/. Results suggest a shift in internal category structure, which is consistent with the view that top down processing results in a shift in the perception of within category variation. Future studies seek to further explicate this question by examining the perceptual learning mechanism in language-impaired populations.

The New Normal: Goodness Judgments of Non-Invariant Speech

In a dynamic world where relating to others is an important aspect of life, being able to effectively communicate is essential. A breakdown between communication partners can sometimes occur when the talkers encounter variation in the incoming speech signal. Examples of such speech signal variation can be accounted for in a variety of ways. Talkers with an articulatory speech impediment such as a lisp, tend to have difficulty pronouncing /s/ phonemes, such that words like “sing” sound closer to “thing.” Additionally, talkers who tend to speak faster, like native speakers on the East Coast of the United States, may also induce speech signal variation. Finally, one of the most apparent sources of speech signal variation is talkers with accented speech. These speakers may be second language learners, or may be native to a particular region with which the listener is unfamiliar with the accent. For example, in the native Boston community, speakers tend to show an /r/ dropping pattern, where words like “car” or “park” have omitted /r/ phonemes. Despite there being many pronunciations of the same word, listeners are still able to take variation in the incoming acoustic speech signal and maintain a stable representation of the talker’s speech. What accounts for this process?

Recent research has proposed two theories for learning a new talker (Kraljic & Samuel, 2005). One idea suggests that a listener hold multiple representations of every speaker they encounter. In this way, listeners would hold speaker- specific adaptations and use cues with regard to the talker. If this is the case, listeners may not actually be adjusting phonemic boundaries, but rather creating multiple representations for each speaker. The second theory suggests a change in phonemic boundaries after exposure to a new talker. If this is the case, listeners actually “retune” or adjust perceptual representations in a general way without regard to the speaker. Such perceptual learning would give the listener the ability to modify a perceptual representation based on exposure to systematic variation in the acoustic speech signal.

While initially, listeners may find a talker’s speech difficult to comprehend due to one of these speech signal variations, after a period of exposure the listener often “retunes” or shifts their speech representations to accommodate the new speaker (Norris et al., 2003). This retuning, also called perceptual learning, allows listeners to normalize the idiosyncratic details of the incoming speech signal. Amazingly, listeners are able to adjust their phonemic boundaries with minimal exposure to a new talker. Norris, McQueen, and Cutler (2003) established that listeners need only 20 ambiguous phonemes to enable such perceptual learning.

One important aspect of perceptual learning is the role of lexically guided information on perceptual adjustment. Norris, McQueen, & Cutler (2003) described such use of lexical information in perceptual learning. In their paradigm, participants were exposed to a lexical decision task where words varying in /f/ and /s/ final Dutch words were manipulated. The tokens were manipulated such that 20 of the words contained an ambiguous speech sound where the listener was required to rely on lexical information on the token they were presented with to make a decision on the word status of the token. For example, one group of Dutch participants heard /f/ final non-Dutch words like “witlos” (where “witlof is a word, but “witlos” is not) where /f/- final words ended with a fricative that was more /s/ like than normal Dutch. Similarly, another group of Dutch participants heard /s/-final non-Dutch words like “naaldbof” (where “naaldbos” is a word and “naaldbof” is not) where /s/-final words ended with a sound that was more /f/-like.

Directly after this, participants categorized an /f/-/s/ continuum and their results suggested that listeners who heard ambiguous tokens in a context which was consistent with an /f/ reading of that phoneme, later categorized more /f/ sounds in the identification task, and conversely, participants who were biased into hearing more /s/ sounds during the training were likely to categorize more sounds as /s/ in identification. This kind of top down processing allows lexical information to directly constrain the listeners perceptual adjustments, such that ambiguous tokens sounded more normalized because of the lexical cues (Norris, McQueen, &

Cutler, 2003). Additionally, this type of lexically guided retuning must be long lasting, so that future encounters with the adjusted output would be processed in a similar fashion.

The perceptual system is able to maintain changes, well after exposure. Kraljic & Samuel (2005) demonstrated how listeners are able to hold perceptual adjustments for a long period of time. They exposed listeners to a new talker with ambiguous tokens midway between /s/ or /□/ in lexical contexts that were consistent with either /s/ or /□/. Immediately after exposure, they tested one group of participants on two ambiguous continua, one in the voice they heard during training, and the other in a different voice. After a silent 25-minute interval, they then tested the other group of participants on the same two continua. Their results indicate that perceptual learning was just as apparent in participants tested 25 minutes after exposure compared to participants tested immediately after. These results address the reset issue in perceptual learning, or the listener's ability to reset speech sound categories after being trained. The results of the previously described experiments propose that the perceptual system is relatively stable, and does not reset back to previous perceptual settings after a period of time has elapsed.

Previous studies have shown that listeners who are exposed to a new talker adjust their phonemic boundaries to correspond to the new speaker. By doing this, listeners make perceptual adjustments very early in processing, such that listeners adjust the boundary between phonemic categories (e.g. between the /s/ and /□/ category), which form the building blocks of higher linguistic units such as words (Kraljic & Samuel, 2005). Speech sound categories, like other cognitive/perceptual categories, have a graded internal structure, with some members of the category considered more representative than others (Theodore, 2013). In one paradigm, Theodore and colleagues (2011) manipulated voice onset time (VOT) such that one group of listeners heard a talker produce /k/ with long VOT's and one group of listeners heard a talker produce /k/ with short VOT's. The listeners were then presented with a continuum from "gain" to “cane” or “goal” to “coal” where they were asked to rate the goodness of /k/. Results indicated that the category members considered to be the best exemplars of a /k/ depended on the previous type of exposure to a talkers voice, so participants exposed to the short VOT talker had a best exemplar range that spanned the shorter VOT range compared to those exposed to the long VOT. These results suggest that internal category structure is in fact shifted when accommodating well-defined category members.

While this graded internal speech sound category offers various advantages to processing the incoming acoustic speech signal, there is also the potential for some disadvantages. On one hand, listeners may be able to process information more efficiently, in the case where there is an inaccurate or unstable representation presented. However, listeners may also be at a disadvantage, and may be too willing to shift tokens unnecessarily. Consider an example where the talker is sick with a cold, or makes a simple articulation error, that is not representative of the actual talker’s pronunciation. In this case, it would be a disadvantage to have a system that is too willing to shift for any ambiguous pronunciation. In this way, listeners must selectively choose the meaningful speech signal variations for adjustment. Thus one important aspect of this mechanism is that a listener is able to maintain a relatively stable, yet flexible perceptual representation.

Previous research suggests that when listeners are exposed to such odd pronunciations that can be attributed to something, listeners do not show perceptual learning. Kraljic & Samuel (2008) used an audiovisual presentation of stimuli, such that participants watched a video of the ambiguous phoneme being spoken. However, one key manipulation in this study was that for some participants the speaker had a pen in her mouth. Their results suggested that participants who watched the speaker with the pen in her mouth, attributed the ambiguity to the pen, and not the talker. In this way, these listeners had an “excuse” for the mispronunciation. However, the participants who saw the woman pronounce the ambiguous tokens without a pen in her mouth, did show the perceptual learning effect, and thereby attributed the ambiguity to the talker. Because of this, the system needs to be selective on what tokens to shift for, otherwise normal talker mispronunciations would lead to inaccurate perceptual adjustments. Kraljic & Samuel (2007) predicted that perceptual learning would only occur when the system has no alternative explanation for the speech signal variation. Therefore, a well-constructed system would not take radical steps of restructuring internal speech categories for alternative explanations to the speech signal variation.

Previous findings (Theodore, 2011) suggest internal category structure is modified when adjusting to clearly defined category members. However, what is currently unknown is whether the entire speech sound category moves when accommodating to ambiguous productions. In the present study, we asked if perceptual learning to ambiguous phonemes causes a shift in internal category structure, or if the perceptual adjustment was limited to the boundary region. We expect participants trained in a talker specific domain to classify some members of the speech sound category to be rated higher, and in this way better representative of the speech sound category than others. However, if this is not the case, and listeners of ambiguous phonemes do not judge all variants of a phoneme to be the same in goodness level, then we expect that listeners characterize all variants within a speech sound category to be judged equally in representativeness of the speech sound category.

Method

Participants*:*

The participants (n=17) in this study were seventeen undergraduate psychology students from the University of Connecticut. All of the participants received course credit for their participation in the study. All participants were at least 18 years old and were native English speakers with normal hearing. The study enrolled eight male participants and nine female participants with a mean age of 18.86 years old. All participants indicated that English was their primary language growing up and that it was the only language spoken prior to the age of 13.

Stimulus Creation and Selection

The lists created for the Lexical Decision task were selected based on previously established training paradigm (Kraljic & Samuel, 2005), see Appendix A for a full list of stimuli. Within the word category, the items were further subdivided into filler words, unambiguous words, and ambiguous (critical) words. 60 filler words were chosen that had no instance of /s/, /□/, /z/, or /□/ phonemes anywhere in the course of the word. The fillers were matched to the critical words in the aspects of stress pattern, number of syllables, and word frequency. 100 non-words were chosen which had no instance of /s/, /□/, /z/, or /□/ in any of the word positions. In order to create equal numbers of “word” and “non-word” responses, a non-word was created from each of the filler words. Non-words were created by manipulating one phoneme, with the same manner of articulation, in each filler word.

Design & Procedure:

All seventeen participants partook in three experimental tasks from an established training paradigm (Kraljic & Samuel, 2005). In the first task, Lexical Decision task, the seventeen participants were randomly assigned to one of two between-subject exposure conditions in which they performed an auditory lexical decision task in a female voice. Each participant heard a series of the same 100 words (e.g. “bullying,” “document,” “parakeet”) and 100 non-words (e.g. “klogodar,” “ryligal,” “wonimtic”) repeated once. Their task was to determine if the item they listened to was a word or a non-word and to indicate their response using the specified computer keys. A key manipulation of this task was that 20 of the selected words contained an ambiguous (50-50% blend of /s/ and /□/) word medially. The key difference between the two training conditions was the context of ambiguous token that the listener was exposed to. Participants in the /s/-biasing training groups were exposed to 20 critical words that would normally contain an /s/ word medially (e.g. “eraser” & “episode”), such that the lexical constraints of the particular word only allowed for an interpretation of the ambiguous sound as an /s/ phoneme. Conversely, participants in the /□/-biasing training groups were exposed to 20 critical words that would normally contain an /□/ word medially (e.g. “reassure” & “publisher”), such that the lexical constraints of the particular word only allowed for an interpretation of the ambiguous sound as an /□/ phoneme.

In the second task, Goodness Judgment task, participants heard a continuum of sounds in the same female voice that they heard during exposure. The continuum ranged in a carefully constructed /s/ and /□/ percentage ratio. At one end of the continuum, participants heard a 30-70% blend (i.e. 30% /s/, 70% /□/), while the highest end of the ratio was a 80-20% blend (i.e. 80% /s/, 20% /□/). Participants heard all 6 continuum points (i.e. 30-70, 40-60, 50-50, 60-40,70-30, 80-20) played randomly 10 times each, over the course of 60 trials. Each item was inserted into an “a?i” token (where ? is the item), where the participant was asked to focus on the middle sound and rate how good the item was of an /s/ sound. Participants in both training condition (/s/ and /□/) were told the same directions for this task. Participants were asked to respond using a 1-7 scale where “1” represented the poorest exemplar of an /s/ sound, and “7” represented the best /s/ sound exemplar. Participants were encouraged to use the entire range of the scale during this task.

In the third task, participants listened to the same continuum of sounds inserted into a “a?i” context, previously described in the Goodness Judgment task, and categorized phonemes as either “asi” or “ashi” (i.e. more /s/-like or more /□/-like). Each participant (in both training groups) heard each of the 6 tokens played randomly 10 times each, over the course of 60 trials.

Results

Lexical Decision Task. Accuracy during training was analyzed with a break-down in each token category as previously described (i.e. “critical,” “filler,” “unambiguous,” “non-word”). This breakdown of stimuli is outlined in Appendix A. Figure 1 outlines mean percent correct in each of the token categories. In the /□/- biasing group, mean accuracy for critical was 95.00%, with a standard deviation 7.07; mean accuracy for filler words was 97.41%, with a standard deviation of 1.88; mean accuracy for unambiguous words was 98.89%, with a standard deviation of 2.20; and mean accuracy for non-words was 94.00%, with a standard deviation of 6.76. In the /s/-biasing group, mean accuracy for critical words was 96.88%, with a standard deviation of 4.58; mean accuracy for filler words was 96.25%, with a standard deviation of 2.14; mean accuracy for unambiguous words was 99.38%, with a standard deviation of 1.77; and mean accuracy for non-words was 92.38%, with a standard deviation of 9.11.

Goodness Judgment task. Figure 2 and Table 1 outline mean /s/ goodness ratings for both /s/-biasing and /□/- biasing groups as a function of continuum point. Additionally, a two way mixed ANOVA analysis with training group (either /s/ biasing or /□/- biasing) as a between subjects factor, the continuum token (30% /s/, 40% /s/, 50% /s/, 60% /s/, 70% /s/, 80% /s/) as a within-subjects factor, and the dependent variable as goodness as “s” was performed. Perhaps unsurprisingly, tests for between subjects showed that on average, there were similar goodness ratings between /s/-biasing and /□/- biasing training groups, with no significant differences, [$F(1,15)= .950$, $p=.345$]. Additionally, a significant continuum point by training group interaction was found, such that there were significant differences between training groups, [$F(5,75)=2.53$, $p=.036$]. Post-hoc tests were performed to further explicate the nature of the interaction. Six independent sample t-tests were performed, and of these two continuum points approached significance. In the 30% /s/ continuum point a p value of .103 was found, and in the 40% /s/

continuum point a p-value of .059 was found. The results showed a marginal effect for the tokens with the least amount of /s/ (i.e. 30% /s/ and 40% /s/ tokens), with a trend for higher ratings to be given to those who were in the /s/-biasing condition. Additionally, visual inspection of Figure 3 reveals that this interaction is likely due to interactions at the lower end of the continuum.

Phoneme Identification task. Mean /s/ identification responses for both the /s/ and /□/-biasing groups as a function of percent /s/ continuum point are outlined in Figure 3 and Table 2. Additionally, a two- way mixed ANOVA analysis with training groups (/s/-biasing and /□/-biasing group) as a between- subjects factor, continuum point (30% /s/, 40% /s/, 50% /s/, 60% /s/, 70% /s/, 80% /s/) as a within-subjects factor, and percent "s" responses as the dependent variable was performed. A significant between subjects effect was found between training groups was found, such that /s/-biasing and /□/- biasing groups showed significant differences in calling a token an /s/, [F (1,15)= 7.331, p=.016]. Perhaps unsurprisingly, a significant main effect of target was found, such that there were greater /s/ judgments as the proportion of /s/ increased, [F(5,75)= 54.974, p= <.01]. There was no significant target by training group interaction found in the present study, [F(5,75)= 1.298, p=.274].

Discussion

The present experiment was designed to address the question: does perceptual learning of ambiguous phonemes cause a shift in internal speech sound category structure or is the perceptual adjustment limited to the boundary region? The results of the experiment suggest that indeed it may be the case that internal category shifted is shifted. Results for the Goodness Judgment task showed higher /s/ goodness ratings at the lower end of the continuum for participants trained in the /s/-biasing condition compared to those in the /□/- biasing group. Interestingly, participants who had heard ambiguous tokens inserted into the s-consistent words were more likely to rate all tokens as sounding "good" on the scale, and showed higher ratings than the participants trained in the /□/ context. With a proportion that was significantly higher in the /□/ phoneme, these /s/-biasing participants were already giving ratings well above the middle of the scale, which further support this shift within the category. These findings suggest that adaptation to the ambiguous /s/-/□/ blend shifted overall judgments of goodness such that non-standard /s/ tokens were rated more highly. The results from the Goodness Judgment task were consistent with the original hypothesis which stated that listeners trained in an /s/-biasing condition, would show higher goodness ratings than those trained in /□/- biasing groups.

Additionally, results from the Phoneme Identification task, replicate the predicted results, with significant differences in calling a token an /s/, between training groups. The results from the present study are consistent with previous findings, where participants who are exposed to /s/-training are more likely to call the ambiguous phoneme an /s/ and participants who are exposed to /□/- training are more likely to call the ambiguous phoneme an /□/ (Kraljic & Samuel, 2005). The /s/-biasing training group showed shifted judgments closer towards /s/ at the 50% continuum point, and the /□/- biasing training group also shifted judgments closer towards /s/ at the 60% continuum point, where mean /s/ responses jumped up to 70%.

The results from the present experiment are consistent with top- down processing of the acoustic speech signal. The perceptual learning effect demonstrated are consistent with lexically guided adjustments, and these adjustments are stored for future use (Norris et al., 2003). It would make sense that listeners are able to do these very quick perceptual adjustments to talkers. Consistent with the title of this paper, what once was characterized as a bad token may sound like a good token and thus phonetic categories may be shifting to create this new perception of normal. This kind of top down processing would also increase a listener's efficiency of processing the incoming acoustic speech signal. If listeners are able to use these quick on the fly processing abilities then it would make sense that the speech signal would be efficiently analyzed without conscious effort. This processing also gives listeners a new way of understanding tokens by allowing for this flexible speech sound category structure.

The significant differences in goodness ratings found between /s/- biasing and /□/-biasing training groups suggests that perceptual learning may indeed have an effect on internal category structure. Further studies seek to expand the current paradigm with an increased sample size. In this way, we expect to see increase the reliability and validity of the results that were found in the present study. Additionally, we expect an increased sample size to increase the statistical power, and in that way approach significance in target by training group interactions in the Phoneme Identification task. In addition, future studies seek to include an expansion on the present paradigm by including a control training group, which would have unaltered blends of the critical words. This would allow training on good tokens to ensure that these good tokens did not induce perceptual learning, the way we see with the idiosyncratic blends. It could also be of use to include a differently designed control group that did not receive any lexical decision training at all, and rather participants would just be asked to participate in the Goodness Judgment task and Phoneme Identification task. This would allow a future study to use this control group to get baseline goodness ratings and identification data.

In addition, future studies seek to examine this phenomenon in different language impaired populations to observe similarities and differences in the way these individuals process the acoustic speech signal. One interesting population to study the perceptual learning mechanism is those with aphasia. Aphasia can present with various deficits in speech production (non-fluent aphasia), language comprehension (fluent aphasia) or both (mixed aphasia). Dunton et al. (2011) conducted an experiment where participants with aphasia heard familiar and unfamiliar accented speech and tested sentence comprehension. They found that the individuals with aphasia were lower in accuracy in both unfamiliar and familiar accents than those without a language impairment. These results may suggest that individuals with aphasia could show a deficit in this perceptual learning mechanism, and future studies seek to examine this question. Additionally, another interesting population to observe the mechanism of perceptual learning is individuals with dyslexia. Steffens et al. (1991) examined speech perception in those with dyslexia using three synthetically created speech continua and found dyslexic subjects were able to label and discriminate the continua, however they did not use the acoustic cues from the speech signal in the same way as a normal reader. This difference in processing of the acoustic speech signal may provide a basis for future studies in perceptual learning in this population. Future research is aimed to explicate these relationships.

References

Dunton, J., Bruce, C., Newton, C. (2011). Investigating the impact of unfamiliar speaker accent on auditory comprehension in adults with aphasia. *International Journal of Language & Communication Disorders*, 46 (1), 63-73.

Eisner, F., & McQueen, J.M. (2005). The specificity of perceptual learning in speech processing. *Perception and Psychophysics*, 67, 224-238.

Jesse, A, & McQueen, J. Positional effects in the lexical retuning of speech perception. Psychon Bull Rev (2011) 1-8. Print. DOI: 10.3758/s13423-011-0129-2

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal?. *Cognitive psychology*, *51*(2), 141-178.

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*(1), 1-15.

Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008). First Impressions and Last Resorts How Listeners Adjust to Speaker Variability. *Psychological science*, *19*(4), 332-338.

Norris, D., McQueen, J.M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23, 299-325.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive psychology*, *47*(2), 204-238.

Myers, E, & L. Mesite. (2012). Neural systems underlying lexically-biased perceptual learning in speech. *Paper presented at the Neurobiology of Language Meeting*, San Sebastian, Spain.

Steffens, M. L., Eilers, R. E., Gross-Glenn, K., & Jallad, B. (1992). Speech perception in adult subjects with familial dyslexia. *Journal of Speech and Hearing Research*, *35*(1), 192.

Theodore, R.M. (2013). Talker-specific perceptual processing: Influences on internal category structure. Pr*oceedings of the 14th Annual Conference of the International Speech Communication Association* (Interspeech 2013), 2802-2806.

Figures and Charts

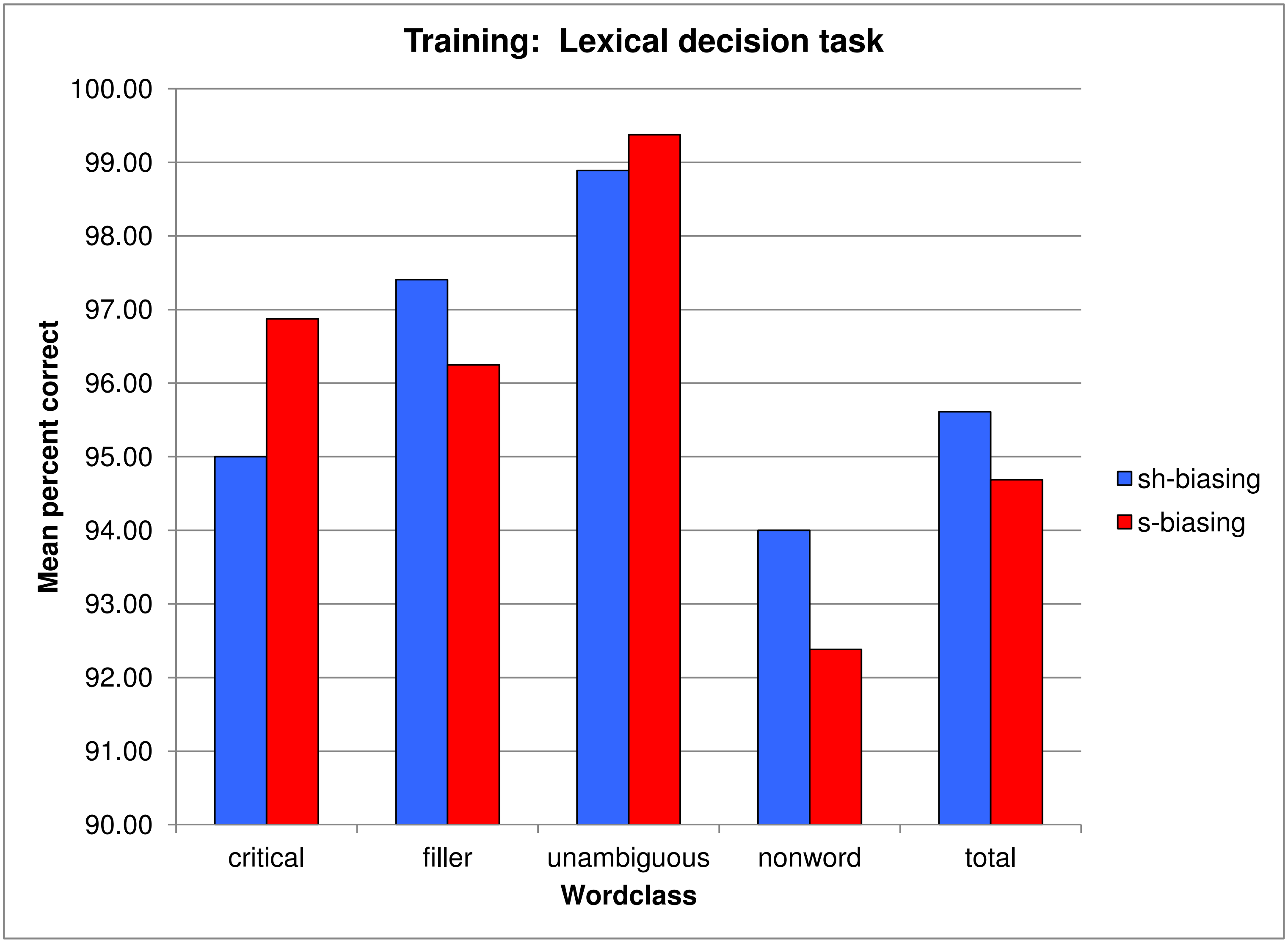

Figure 1. Accuracy during training was analyzed with a break-down in each token category.

Table 1. This table outlines mean goodness ratings at each continuum point for both /s/ and /□/-biasing groups.

	/s/- biasing condition	**/□/-biasing group**
30% /s/	4.28	3.06
40% /s/	4.90	3.40
50% /s/	5.02	4.18
60% /s/	5.42	5.08
70% /s/	5.33	5.40
80% /s/	5.42	5.66

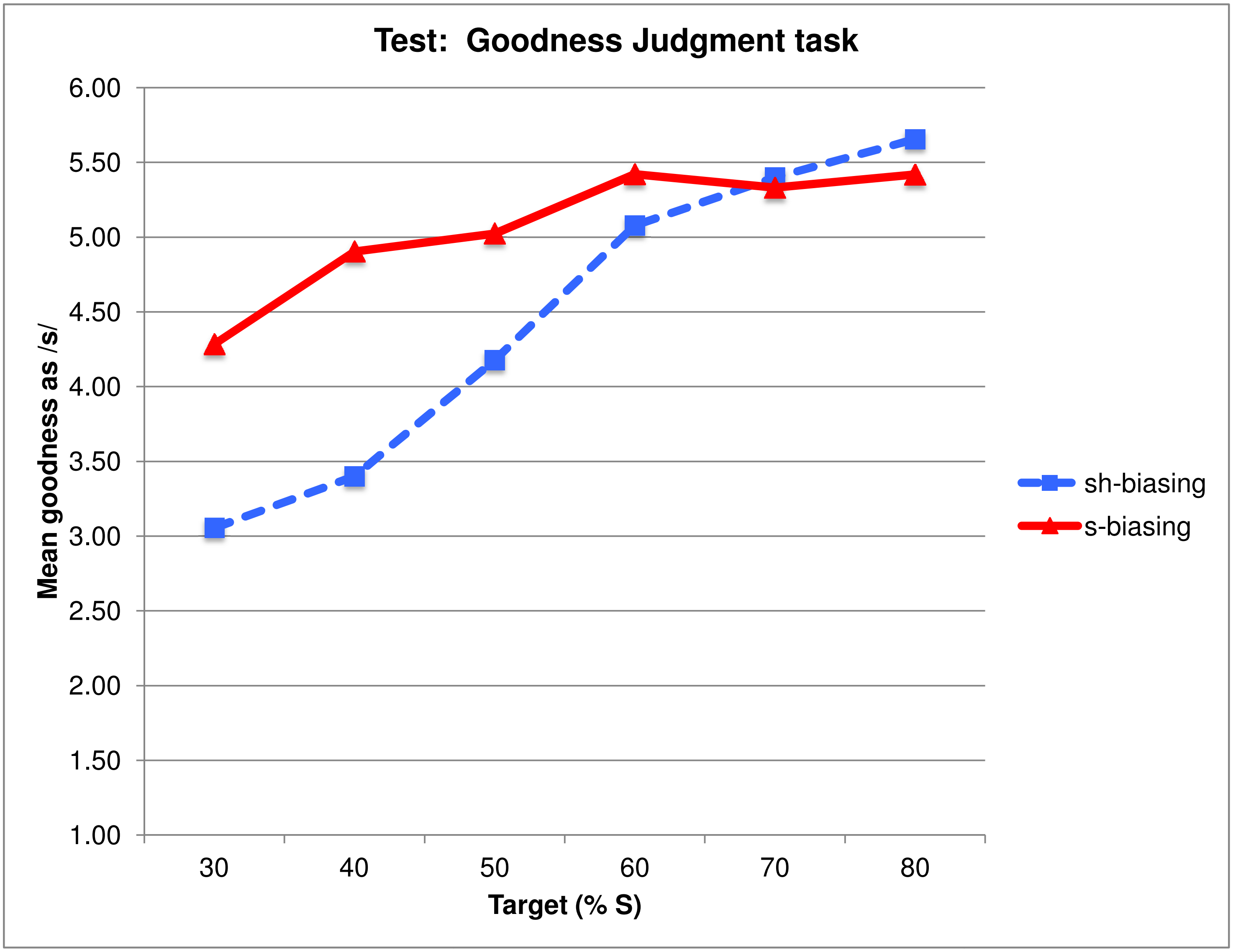

Figure 2. Mean /s/ goodness ratings for both /s/-biasing and /□/- biasing groups as a function of continuum point.

Table 2. This table outlines mean identification responses at each continuum point for both /s/ and /□/-biasing groups.

30% /s/	21.25	7.78
40% /s/	37.50	10.00
50% /s/	53.75	30.00
60% /s/	90.00	70.00
70% /s/	93.75	93.33
80% /s/	97.50	94.44

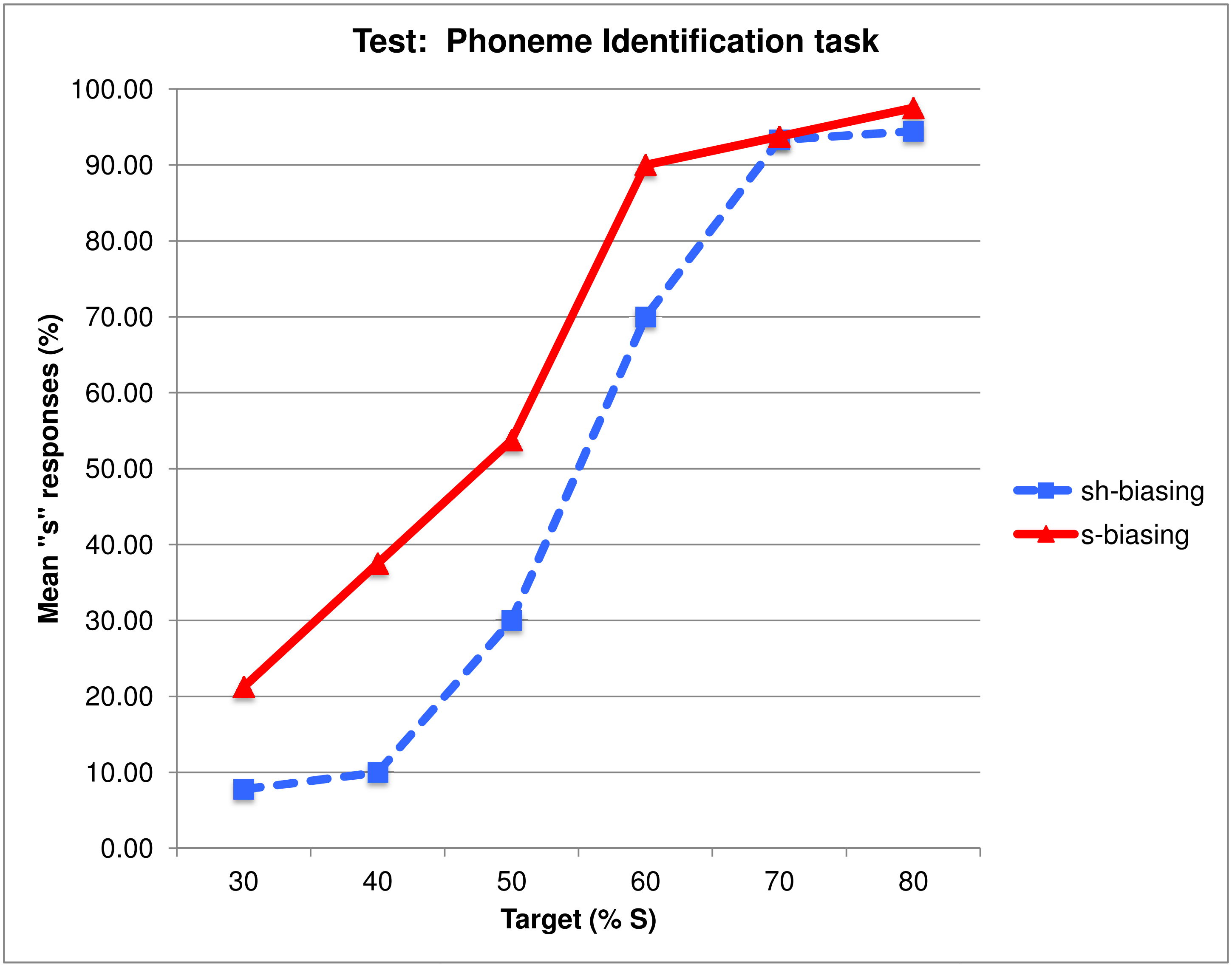

Figure 3. Mean /s/ identification responses for both the /s/-biasing and /□/- biasing groups as a function of percent /s/ continuum point.

Appendix A

Table A. The following table is a list of critical, filler, and non-words words for /s/-biasing and /□/- biasing groups training group. This list was a replication of a well-established training paradigm (Kraljic & Samuel, 2005).

/s/-biasing critical words (where /s/ phoneme was replaced with 50% /s/-50% /□/ blend)	**/□/- biasing critical words** (where /□/ phoneme was replaced with 50% /s/-50% /□/ blend)
Arkansas	Ambition
Coliseum	Beneficial
Compensate	Brochure
Democracy	Commercial
Dinosaur	Crucial
Embassy	Efficient
Episode	Flourishing
Eraser	Glacier
Hallucinate	Graduation
Legacy	Impatient
Literacy	Initial
Medicine	Machinery
Obscene	Negotiate
Parasite	Official
Peninsula	Parachute
Personal	Pediatrician
Pregnancy	Publisher
Reconcile	Reassure
Rehearsal	Refreshing
Tennessee	Vacation

Table B. The following table is a list of the filler words (60) and non-words (100) that every participant in both training conditions had exposure to. This list was a replication of a well-established training paradigm (Kraljic & Samuel, 2005).

Accordion	Document	Immoral	Membrane	Platonic
America	Domineering	Inhabit	Memory	Remedial
Annoying	Dynamite	Knowingly	Metrical	Romantic
Armadillo	Embody	Laminate	Military	Tactical
Bakery	Gardenia	Legally	Momentary	Titanium
Ballerina	Grammatical	Liability	Napkin	Turbulent
Blueberry	Gullible	Lobbying	Negate	Tutorial
Bullying	Hamburger	Lunatic	Outnumber	Umbrella
Camera	Honeymoon	Lyrical	Panicky	Warrantee
Crocodile	Hurdle	Manually	Parable	Wealthy
Darken	Identical	Marina	Parakeet	Withdrawal
Directory	Ignite	Melancholy	Pineapple	Wrinkle

Table C. The following table is a list of the non-words (100) that every participant in both training conditions had exposure to. This list was a replication of a well-established training paradigm (Kraljic & Samuel, 2005).

Igoldian	Hominaim	Lindel	Lindel	Kegimel
Anolipa	Hilder	Acominig	Aigi	Kelabidel
Imoyem	Itempider	Mibgem	Ailounam	Kermimer
Alnadiro	Aknid	Mikid	Amaler	Kerkrun
Pakelo	Emhoutic	Admunker	Anemer	Lilgrai
Galliwinou	Mowery	Bimikay	Bamtel	Logelai
Pluepelai	Wonimitic	Baliber	Bliparg	Loubel
Pourilar	Weekery	Bawaseet	Gairelom	Maidnow
Ganla	Riakirity	Bimobel	Galliwinou	Marody
Klogodar	Woppakin	Kradomet	Gerbualo	Omperog
Perkum	Rumatik	Lenediaw	Geypalg	Pirugalo
Tilegkalo	Ryligal	Wonontic	Gilday	Rakil
Pogunemd	Namuery	Dadigal	Gondimually	Rengimer
Konimeelum	Nawinow	Bikanian	Gonedial	Rimkuwar
Tymolipe	Neramgory	Durkuwomt	Hamimoc	Tamical
Enpaiki	Nempring	Datiliar	Hiliun	Tounamlemp
Kaldemia	Nomeray	Omplero	Ibirak	Umikory
Kloumidiger	Nekridal	Rawamtee	Imdalier	Ungelnin
Kuradel	Niritaly	Lirthy	Ithomel	Wojalto
Hintarber	Nomemtoly	Rikmaral	Irimelder	Youmgel