Linguistic Portfolios

Volume 6 Article 7

2017

The Acoustics of Coda Devoicing in a Central Minnesota English Idiolect

Alex Hennen
St. Cloud State University, arhennen@gmail.com

Ettien Koffi
St. Cloud State University, enkoffi@stcloudstate.edu

Linguistic Portfolios–ISSN 2472-5102 –Volume 6, 2017 |

THE ACOUSTICS OF CODA DEVOICING IN A CENTRAL MINNESOTA ENGLISH IDIOLECT

ALEX HENNEN AND ETTIEN KOFFI

ABSTRACT

In many languages, voiced segments that occur in syllable codas are systematically devoiced. This article examines coda devoicing in an idiolect of a Central Minnesota speaker of American English. The consonants under investigation are voiced fricatives, voiced stops, and the voiced affricate /dʒ/. The 40/60 threshold proposed by Gradoville (2011) is used as the limen to discriminate between voiced and devoiced consonants in syllable codas.

1.0 Introduction

The main purpose of this paper is to investigate and describe devoicing in my idiolect.[1] Before that, we will first take a quick look at two IPA transcriptions of a slightly modified version of the well-known Speech Accent Archive passage (Weinberger, 2015). The first is a purely impressionistic transcription, while the second is a modified transcription after several acoustic measurements. All spectrograms and measurements were collected by using Praat (Boersma & Weenink, 2016).

1.1 Impressionistic Transcription

At the beginning of Dr. Koffi's phonetics course, students were asked to record themselves reading a slightly modified version of the Speech Accent text.

Please call Stella. Ask her to bring these things with her from the store: Six good spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a foot long sandwich as a snack for her brother Bob. We also need a small plastic snake, a yellow book, a rubber duck, a paper I-pad, the dog video game, a big toy frog for the kids, but not the faked gun. She can scoop these things into three red bags, and two old backpacks, and we will go meet her, Jake, and Jenny Wednesday at the very last train station at the edge of the zoo near York's Treasure Bank.

After the lectures on how to use the International Phonetic Alphabet (IPA) transcription, the students were asked to transcribe their speech based on their recordings. The following is my transcription:

[pʰliːz kal stɛlə. æsk həˠ tʰɪ bɹ̃ɲ̃ ðiz θɪ̃ŋ̃s wɪθ həˠ fɹ̃ãm ðə stoɹ. s̪ɪks gʊd spʰũ̃nz ʌv fɹ̃ɛʃ sno piːz, faɪv θɪk slæbz ʌv blu ʧiːz, ɛ̃n mebi a fʊ̄t̄ lãŋ̃ sæ̃nwɪʧ æz a snæk̄ foɹ həˠ bʰɹ̃ʌðəˠ bʰabʰ. wi ʌlso niːd a smaɬ plæstɪk s̪nek, a jɛlo bʰʊkʰ, a ɹ̃ʌbəˠ dʌkʰ, a pʰepʰəˠ aɪ pʰæd, ðə dag vɪdio gɛ̃m, a̪ bɪg tʰɔɪ fɹ̃ag foɹ ðə kɪdz, bʌ̄t̄ nɑtʰ ðə fekt gɅ̃n. ʃi kæ̃n s̪kupʰ ðiz θɪ̃ŋ̃z ĩntə θɹ̃i ɹ̃ɛd̄ begz, ɛ̃n tu ol bæk̄pæks, ɛ̃n wi wɪl go mit həˠ, dʒek, ɛ̃n dʒɛni wɛ̃nzde ǣt̄ ðə veɹ̃i læs tɹ̃ɛ̃n steʃɜ̃n ǣt̄ ði ɛdʒ ʌv ðə zu, niɹ̃ joɹ̃ks ʧɹ̃ɪɛzəˠ bɛ̃nk.]

[1] The "Is" and "Mys" in the paper refer to the first author. He provided the data and the preliminary analysis for this paper. The second author has complemented the first author's initial analysis and interpreted his measurements in order to turn his findings into a suitable acoustic phonetic paper.

1.2 Acoustically Informed Transcription

Students were asked to revise their initial impressionistic transcription every time a group of segments were analyzed acoustically. This exercise gave us an appreciation of the enormous differences that exist between impressionistic transcriptions and an acoustically informed transcription. These differences are reflected in the transcription below:

[pʰliːs kʰɑl stɛˈlə. ʔæsk hə˞ tə bɹĩŋ ðis θĩŋs wɪθ hə˞ fɹɑ̃m ðə stoɹ. sɪks gʊ̰d spũns ʌv fɹɛʃ no pʰiːs, fɑɪ̰ θɪk slæps ʌv blu tʃiːs, ʔɛ̃n meˈbi ʔɑ fʊʔ lɑ̃ŋ sæ̃ˈnwɪtʃ ʔæz ʔɑ snæk̰ fɔɹ hə˞ bɹʌˈðə˞ bɑb. wi ʔɑˈlso niːd ʔɑ smɑl pʰlæˈstɪk snek, ʔɑ̰ jɛloˈ bʊk, ʔɑ ɹʌˈbə˞ dʌk, ʔɑ pʰepə˞ˈ ʔɑɪˈpʰæd̰, ðə dɑg ɣɪˈdio gẽm, ʔɑ bɪg tʰɔɪ fɹɑg fɔɹ ðə kʰɪd̰s, bʌt̰ nɑt ðə fekt gʌ̃n. ʃi kʰæ̃n skup ðis θĩŋs ĩntə θɹi ɹɛd begs, ʔɛ̃n̰ tʰu ʔol bæˈkʔpʰæks, ʔɛ̃n wi wɪl go mit hə˞, dʒek, ʔɛ̃n dʒɛni wɛ̃ˈnd̰e æt̰ ðə ɣeɹi læs tʰɹɛ̃n steˈ̰ʃə̃n æt̰ ði ʔɛdʒ̰ ʌv ðə zu, niɹ joɹks tʃɹɛzə˞ bẽnk.]

At the end of the semester, students are asked to focus on one acoustic phonetic feature that is of interest to them and write their paper on it. Several topics were of interest to me, but given the course requirement, I was forced to choose one topic. Three features of my speech caught my attention: the devoicing of the initial [v] in <very> [veɹi] and <video> [vɪdio]. I also noticed that I produced a lot of vowel initial glottalization. However, I made up my mind to write about coda devoicing in the following words: *<things>, [θĩŋs] <spoons> [spũns], <Wednesday> [wɛ̃ˈnd̰e], <slabs> [slæps], <kids> [kʰɪd̰s], <bags> [begs], <please> [pʰliːs], <these> [ðis], <peas> [pʰiːs], <cheese> [tʃiːs], <edge> [ʔɛdʒ̰], <five>̰, [fɑɪ̰], and <i-pad> [ʔɑɪˈpʰæd̰]*. We can see from the acoustically informed transcription that coda devoicing occurs frequently in my idiolect. As such, this paper will focus on my findings regarding devoicing of both single segments in codas as well as two segment coda clusters. This will be done by investigating several different types of coda clusters. Segments such as fricatives that are normally devoiced in English will not be described in detail. Yet, I give them some attention because I devoice them rather systematically. Not once did I not devoice them. This is a surprising feature of my speech that I did not notice before. The bulk of the paper focuses on how I devoice some stops and not others.

1.3 A Brief Review of the Literature on Coda Devoicing

When binary features are used, segments are perceived either as voiced or devoiced. However, the reality is more complicated than this. Between these two polar opposites, there are varying degrees of voicing and devoicing. Smith (1997) has done one of the most sophisticated studies of the devoicing of the sibilant fricative /z/ in American English. She combined airflow measurements and electroglottographic (EGG) data to pinpoint various degrees of devoicing. The following degrees of voicing and devoicing are noted on pages 478:

1. Segments are fully voiced if 90 to 100% of their duration is voiced
2. Segments are partially voiced if 25 to 90% of their duration is voiced
3. Segments are devoiced if 0 to 25% of their duration is voiced[2]

Smith's system is rather cumbersome, especially her distinction between "partially voiced" and "devoiced" categories. She says so herself on page 479, "There was not a clear boundary

[2] Smith (1997) does not distinguish between voiceless and devoiced fricatives. However, it can be assumed that segments in which voicing is 0 to less than 10% of are voiceless, especially if they occur in postvocalic positions.

between devoiced and partially voiced categories for any speaker group, but that the 0-25% division grouped together most of the tokens with less voicing." We will leave Smith's categories aside and turn to Gradoville's (2011:68) simpler yet efficient limen for differentiating between voiced and devoiced segments. We refer to it simply as the 40/60 Threshold. He explains it as follows:

> Tokens with as little as 40% voicing (60% voiceless) are perceived as voiced. A token … does not even have to be mostly voiced for it to be perceived as voiced by the participants.

In his concluding remarks, he makes the following statement on page 71:

> According to the present analysis Praat's internal pulse-based voice report and the low-frequency-to-total intensities ratio provide the best match for what can be observed in the spectrogram and auditorily. The voice report most closely matches what the linguistically-trained participants perceived, but it makes no distinction regarding the intensity of voicing at any point in time.

Since the voice report in Praat "most closely matches what linguistically-trained participants perceived," we will use it to calculate the amount of devoicing in the first author's pronunciation. Praat offers a very easy way to calculate voicing ratio. To do so, one selects the <Pulse> tab, then one goes to <Voice report>, and then <Voicing>. Praat calculates automatically the amount of voicing in any selected segment. This is the method used to calculate all voicing percentages in the codas reported in Tables 1 and 2 in this paper.

1.4 The Devoicing of Fricatives

The 40/60 Threshold is used to determine whether or not I devoice the fricatives that occur in syllable codas when I speak. To find out I measured the fricatives in the following words: <*please, these, these, spoons, Wednesday, five, edge*>, as shown in Figures 1, 2, and 3.

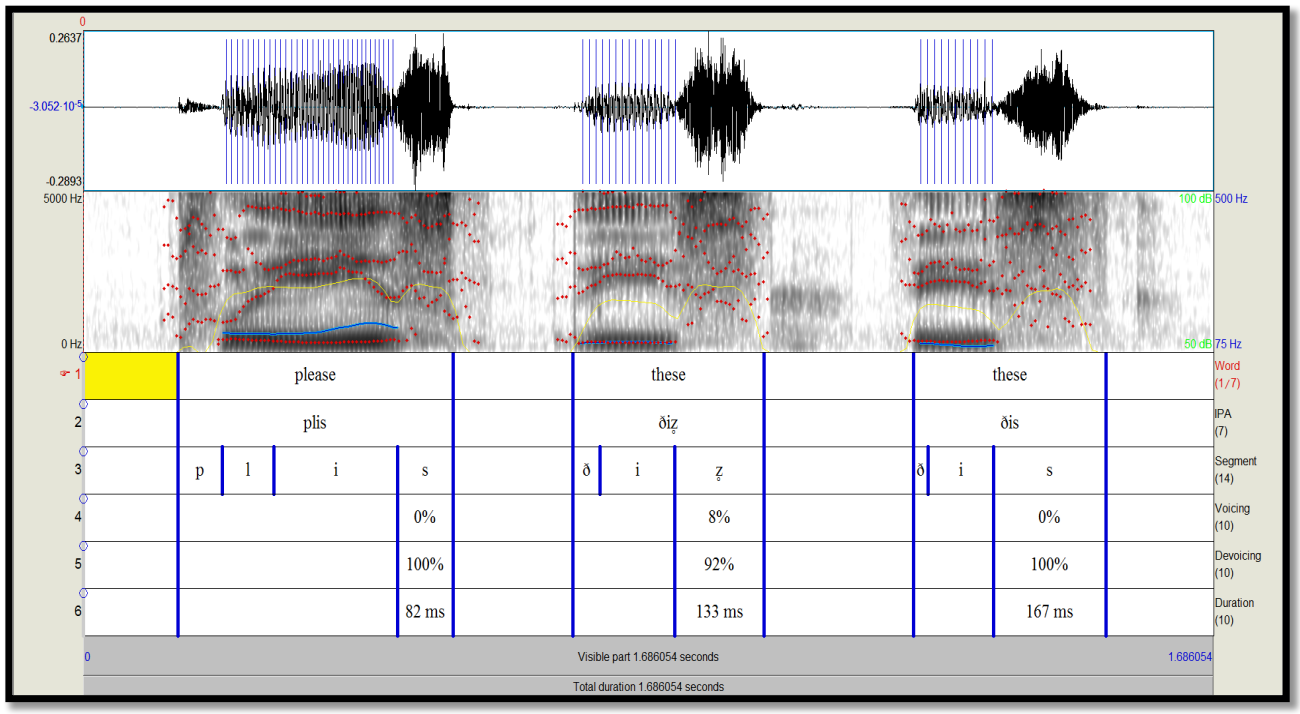

Figure 1: Devoicing of [z] and [ð]

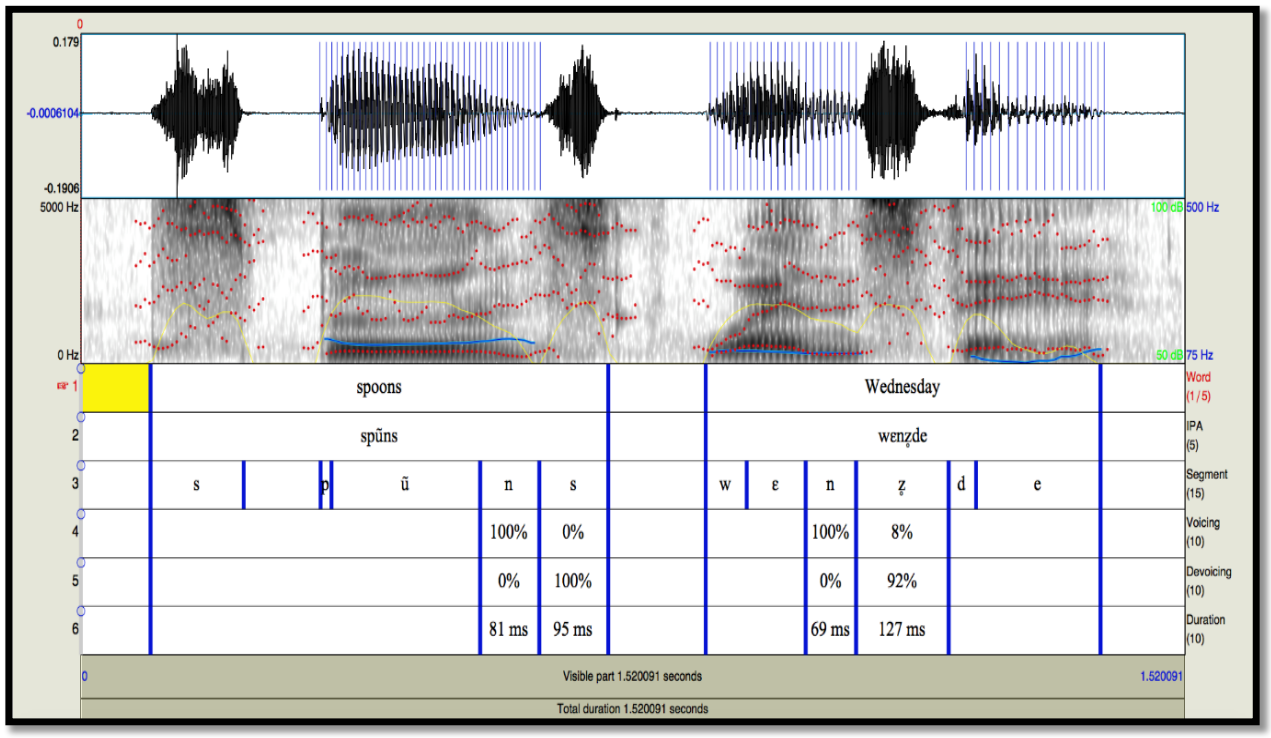

Figure 2: Devoicing of [nz] Coda Clusters

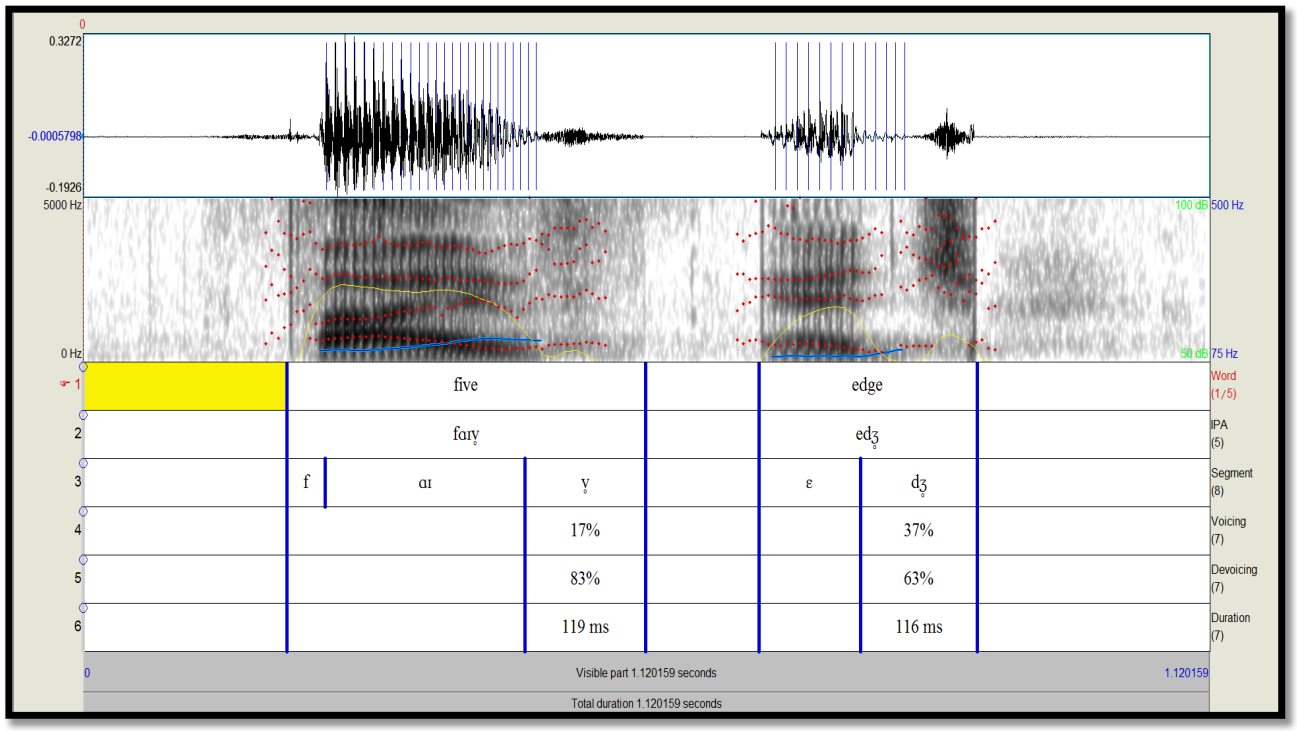

Figure 3: Devoicing of [v] and [dʒ]

The relevant measurements are summarized in Table 1:

Words	pleas**e**	th**e**se	th**e**se	spoon**s**	Wednes**d**ay	fi**v**e	e**dg**e
Segment	/z/	/z/	/z/	/z/	/z/	/v/	/dʒ/
%voicing	0%	8%	0%	0%	8%	17%	37%
%devoicing	100%	92%	100%	100%	92%	83%	63%
Duration	82 ms	133 ms	167 ms	95 ms	127 ms	119 ms	116 ms

Table 1: Fricative Coda Devoicing Measurements

Voiced fricatives were, without exception, devoiced in codas, regardless of the segments preceding them. This devoicing was heavy, as shown by some examples. The segment [z] lost all voicing to become [s] with 0% voicing in all cases except two. Only in one instance of <*these*> and in <*Wednesday*> did the normally expected [z] retain any voicing at all, but even in these instances they had only 8% of their voicing left and became [z̥]. This devoicing occurred regardless if the segment occurred by itself in the coda or as part of a cluster. When in a cluster after a nasal segment, the nasal was voiced, but the fricative was unvoiced. We see this in <*spoons*> and <*Wednesday*> where the preceding [n]s were 100% voiced, but the following [z]s were devoiced. The segment [v] in <*five*> retains 17% voicing, and [dʒ] in <*edge*> had 37% voicing. The latter is included here as it is the combination of a stop and a fricative. We will revisit [dʒ] later when we discuss devoicing of stops in the coda.

1.5 Fricative Devoicing and Intelligibility

Does the devoicing of fricatives in syllable codas affect the intelligibility of the first author's speech? The question can be answered by referring to a relative functional load (RFL) table of phonemes such as that provided by Koffi (2016: 46-47). Page 48, he ranks levels of unintelligibility, from slight to severe, which we will also refer to. Essentially, RFL percentages between segments are a measure of how often the two segments contrast phonemically between words in a language (Koffi, 2016: 48). The word final RFL between [s] and [z] is 38%, which would produce low unintelligibility. The one between [f] and [v] is 9%, so we would expect only slight unintelligibility. The one between [tʃ] and [dʒ] is 8%, which would also produce slight unintelligibility. Since we are contrasting codas, we will not discuss word initial relative functional loads.

1.6 The Articulatory and Aerodynamic Explanations of Fricative Devoicing

As shown from the RFL data above, devoicing fricatives in the coda, by itself, only has a small impact on intelligibility. This could help explain why I always devoice my coda fricatives. If a pronunciation of this kind does not cause confusion, I'm unlikely to notice it as an integral part of my idiolect. However, this does not explain how this pronunciation comes about in the first place. The *Cambridge Handbook of Second Language Acquisition* describes devoicing as an "…aerodynamic difficulty in producing voicing in final fricatives; while voicing requires adduction of the glottis, frication requires a sufficient airflow through the glottis" (Cambridge, 2013: 542). Basically, this means that the frication that characterizes fricatives requires an open passage to escape through the glottis, while voicing requires the glottis to close to create the vocal fold vibration characteristic of voicing. So, it is difficult to produce frication and voicing at the same time, and it is easier to produce voiceless fricatives. This combined with the fact that devoicing coda fricatives will not cause much confusion seems to be a likely explanation for why my coda fricatives are mostly devoiced. Even though coda devoicing is characteristic of my speech, I'm at the mercy of an articulatory phenomenon that is widely spread. The pronunciation of fricatives in syllable coda pits aerodynamic and articulatory features against each other. This may explain why, according to Johnson (2012:156), "Voiced fricatives are relatively unusual in the languages of the world, [and] undergo a variety of motivated alternations, and are surprisingly difficult to produce. … Because a certain degree of airflow is necessary in order to produce turbulence, voiced fricatives lose their frication, …"

1.7 Devoicing of Voiced Stops in the Coda

Devoicing of voiced stops in the coda in my idiolect is not nearly as straightforward as the behavior of my fricatives in codas. We refer to the following table:

Words	goo**d**	bi**g**	re**d**	fro**g**	Bo**b**	I-pa**d**	ki**ds**	ba**gs**	sla**bs**	e**dg**e
Segments	[d]	[g]	[d]	[g]	[b]	[d]	[d]	[g]	[b]	[dʒ]
%voicing	100%	100%	82%	43%	42%	24%	25%	13%	0%	37%
%devoicing	0%	0%	18%	57%	58%	76%	75%	87%	100%	63%

Table 2: Stop Coda Devoicing Measurements

We can discover some insights by investigating the data in Table 1. Voiced stops were voiced or devoiced in codas in my speech to varying degrees. The amount of voicing did not seem to depend on place of articulation of the stop itself: [d] was 100% voiced in <*good*>, 82% voiced

in <*red*>, 24% voicing in <*I-pad*>, and 25% voiced in <kids>. The segment [g] was 100% voiced in <*big*>, 43% voiced in <*frog*>, and only 13% voiced in <*bags*>. The segment [b] has comparatively less data, but still has a difference of 42% voicing in <Bob>, and 0% voicing in <*slabs*>. The syntactic context in which the coda occurs may have something to do with the degree of devoicing. Take <*i-pad*> for example, it occurs before a pause indicated orthographically by a comma. However, the <d>s in <*good*> and <*big*> were not devoiced because they were immediately followed by the nouns that they modify. A syntactic explanation is plausible.

On average, voiced stops that occurred in coda clusters had much less voicing than voiced stops that occurred in the coda individually, except the case of <*i-pad*> explained earlier. The [d] in <*I-pad*> was actually devoiced enough to be perceptually devoiced according to the 40/60 Threshold. Still, it is notable that there were varying amounts of voicing in these segments, especially when the voicing of [g] in <frog> and [b] in <Bob> were so close to the 40% threshold.

In coda clusters, the voiced stops were always followed immediately by a fricative, which as we saw in the last section, are usually devoiced. Phonologically, the devoicing of these voiced segments makes sense. Since the fricatives at the end of the clusters are devoiced, a rule of regressive devoicing assimilation operates to cause the preceding segment to be also devoiced. There is, as it were, a devoicing harmony rule that applies in the speech of the first author. However, this is mere speculation for the moment because we would need far more data to support such a contention.

1.8 Stop Devoicing and Intelligibility

What impact does the devoicing of stops in the coda have on the intelligibility of the first author's speech? In single codas, answering the question of unintelligibility is simple. The Relative Functional Load (RFL) in word final for [p] and [b] is 14% which means slight unintelligibility. That for [k] and [g] is 29%, which means low unintelligibility. And that for [t] and [d] is 72%, which would mean high unintelligibility. Thus, the devoicing of [d] is the only one that could produce a serious problem because in a case such as <*i-pad*>, the first author actually produced <*i-pat*>. This specifically could cause a lexical competition between <*i-pad*> and <*i-pat*>. However, since there is not yet a product on the market called <*i-pat*>, hearers will most likely reinterpret [d̥] as [d] even though in actuality no voiced [d] was produced.

Unintelligibility when double coda devoicing occurs can be measured as follows. First, the RFL between the last segments in the coda are calculated. The words with double codas in the data are <*slabs*>, <*kids*>, and <*bags*>. The RFL for [s] and [z] is 38%. Secondly, we calculate the RFL of the segments that immediately precede [s] and [z]. The RFL between [p] and [b] is 14%, that of [t] and [d] is 72%, and that of [k] and [g] is 29%. The devoicing of [b] in <*slabs*> produces <*slaps*>, that of [d] in <*kids*> yields <*kits*>, and that of [g] in <*bags*> leads to <*backs*>. Figure 4 shows that in all three cases double coda devoicing takes place:

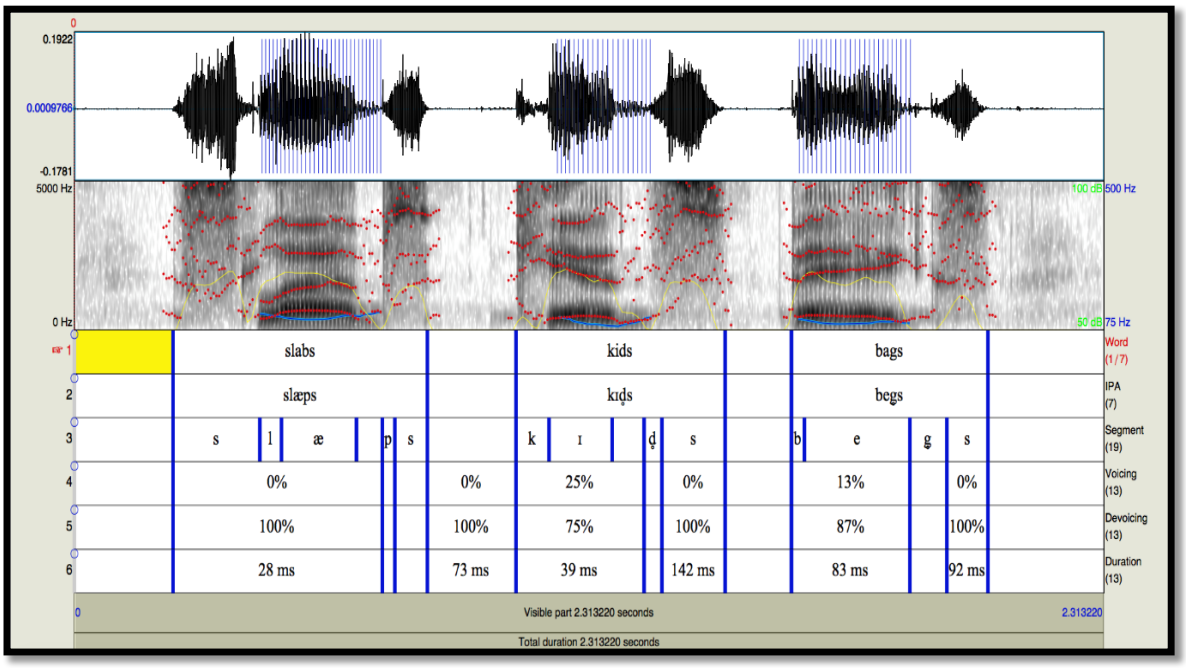

Figure 4: Devoicing of Coda Clusters Involving Stops

All three of these instances create specific lexical competition. The pronunciation of <*slabs*> as [slæps] can be confused with <*slaps*>, <*kids*> as [k^{h}ɪts] can be mistaken for <*kits*>, and <*bags*> [bæks] can be misunderstood as <*backs*>. These examples show that double coda devoicing exacerbates unintelligibility issues more than single coda devoicing. As a rule of thumb, we can say that when a lexical item undergoes two changes or more, recoverability is challenging, and lexical recognition becomes problematic. This is true, irrespective of the unintelligibility ratings of the RFLs of the individual segments involved in the coda cluster.

1.9 The Devoicing of the Affricate [dʒ]

Now we will take another look at the word <*edge*>. We will do so in two steps. First, we will investigate the devoicing that occurs in <*red*>. Secondly, we will examine the devoicing of [dʒ]. We do so because the affricate [dʒ] is made up of a stop and a fricative. Both [dʒ] and [d] are immediately preceded by the vowel [ɛ].

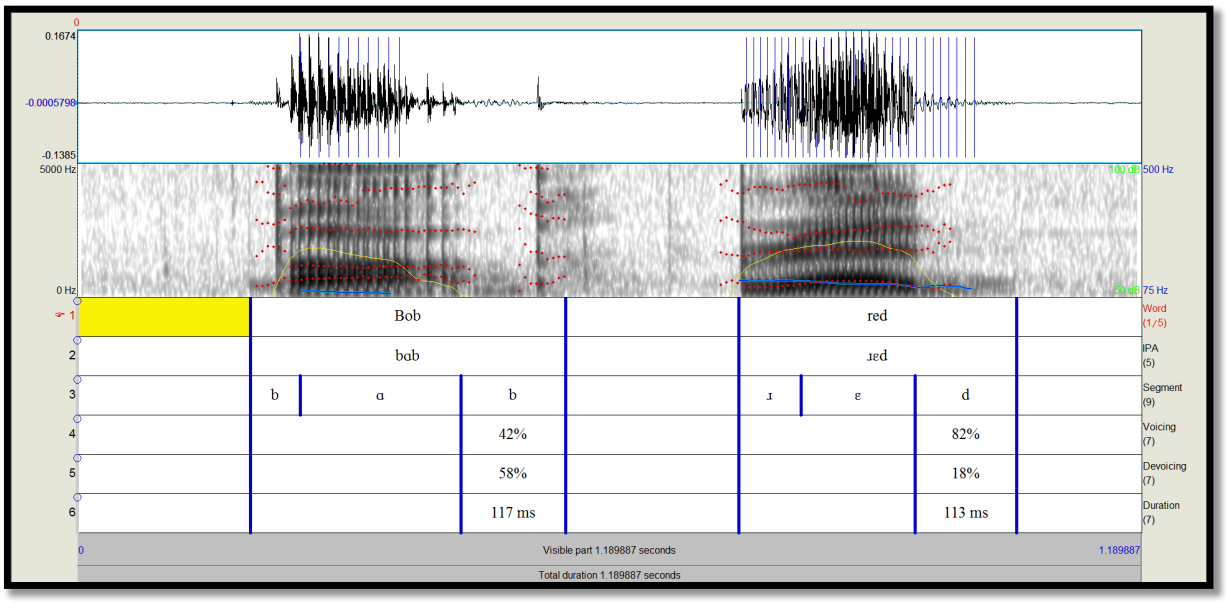

Figure 5: Devoicing of [d] in the Coda

The [d] of [ɹɛd] has the following measurements: It is voiced 82%, and devoiced 18%. Its entire duration is 113 ms.

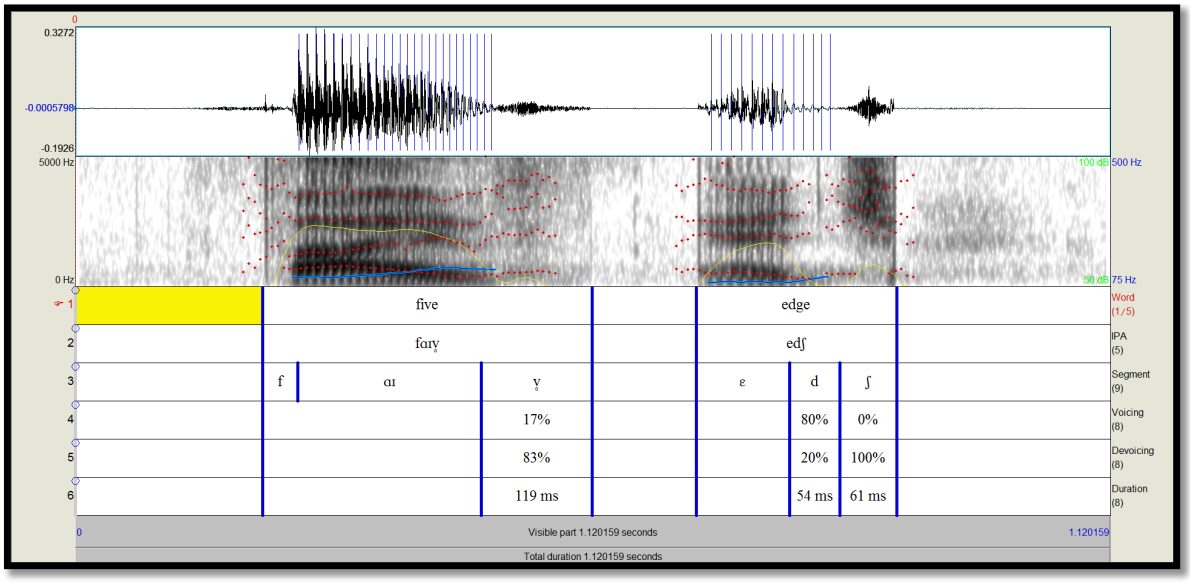

Figure 6: Revisiting the Affricate [dʒ]

The [d] portion of [dʒ] has the following measurements. It is voiced 80%, devoiced 20% and lasts 54 ms. With regard to voicing, the portions of [d] in <red> (82%) and the one in <edge> (80%) are for all practical purposes the same. The acoustic data shows that the first author does not pronounce <edge> as [ɛdʒ], but rather as [ɛdʃ]. He devoices [ʒ] without devoicing [d]. The regressive devoicing assimilation discussed in 1.4 and 1.8 does not apply to the [dʒ] portion of <edge>. This may be attributable somehow to the special characteristics of affricates. Though they consist of a sequence of two segments from the point of view of

articulation, yet they behave like singletons. Since there is no lexical item in English produced as [ɛdʃ], hearers will mostly likely perceive the intended word accurately as <*edge*>. Consequently, the way the first author pronounced <edge> is not likely to be confused with <*etch*> because in the latter, both [t] and [ʃ] are voiceless.

1.10 Summary

In this paper, we explored the devoicing of consonants in syllable codas of the first author's speech. It was found that he always devoiced fricatives in codas, regardless of if they appeared in a cluster or not. As for stops, there appears to be a trend in his speech to devoice them more heavily following low vowels, and less heavily after high vowels. It is unclear at this time what could be causing this trend. In coda clusters where a voiced stop was followed by [z], the whole coda was perceptually devoiced. The singleton coda [d] in <*I-pad*> was perceptually devoiced. The devoicing of this segment is likely to have a large effect on intelligibility given the sheer number of lexical minimal pairs between [d] and [t]. For example, <*pot*> and <*pod*> can be very easily confused in the first author's speech. Devoicing in coda clusters constitutes an obstacle for intelligibility because they increase lexical competition and/or cause lexical items not to be readily recognized.

ABOUT THE AUTHORS

Alex Hennen is an undergraduate student at Saint Cloud State University, MN. He is studying to complete a major in English with an emphasis in linguistics, and a minor in philosophy. He is currently on a study abroad program in Japan for a year where he is mainly studying Japanese and linguistics. His plans include becoming a Japanese-English translator and/or interpreter. He can be reached via his Saint Cloud State University email at: arhennen@stcloudstate.edu or via his personal email at: arhennen@gmail.com.

Ettien Koffi, Ph.D. in linguistics from Indiana University, teaches linguistics at Saint Cloud State University, MN. Author of many peer-reviewed articles on various topics in linguistics and of four books: *Language Society in Biblical Times* (1996), *Paradigm Shift in Language Planning and Policy: Game Theoretic Solutions* (2012), *Applied English Syntax* (2010, 2015), and the *New Testament in Anyi Morofu* (2017), a task which took over 25 years. Specializing in acoustic phonetics, dialect variation, and emergent orthographies, his current research centers on speech acoustics of L2 English (within the Speech Intelligibility Framework), Central Minnesota English, and Anyi. He can be reached at enkoffi@stcloudstate.edu.

References

Boersma, Paul & Weenink, David (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.17, retrieved 21 April 2016 from http://www.praat.org/

Cambridge Handbooks in Language and Linguistics: *Cambridge Handbook of Second Language Acquisition*. (2013). West Nyack, US: Cambridge University Press. Retrieved from http://www.ebrary.com.

Gradoville, Michael. Validity in Measurements of Fricative Voicing: Evidence from Argentine Spanish. Selected Proceedings of the *5th Conference on Laboratory Approaches to Romance Phonology*, ed. Scott M. Alvord, pp. 59-74. Somerville, MA: Cascadilla Proceedings Project.

Johnson, Keith. 2002. *Acoustic and Auditory Phonetics.* Third Edition. Malden, MA: Wiley-Blackwell.

Koffi, Ettien. 2016. *Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility.* Course Manuscript: St. Cloud, MN.

Smith, Caroline L. 1997. The Devoicing of /z/ in American English: Effects of Local and Prosodic Context. *Journal of Phonetics* 25, 471-500.

Weinberger, Steven. (2015). *Speech Accent Archive*. George Mason University. Retrieved from http://accent.gmu.edu