Linguistic Portfolios

Volume 11 Article 8

2022

# DO ACOUSTIC PHONETIC CORRELATES VARY IN RELATION TO GRAMMATICAL FUNCTIONS? EXEMPLIFICATION WITH "Her"

Ettien Koffi
*St. Cloud State University*

Matthew Beckstrand

# DO ACOUSTIC PHONETIC CORRELATES VARY IN RELATION TO GRAMMATICAL FUNCTIONS? EXEMPLIFICATION WITH <HER>

ETTIEN KOFFI WITH MATTHEW BECKSTRAND[1]

ABSTRACT

*Do speakers produce the same word differently if its grammatical function changes? The word <her> is used to provide some answers. This word is optimal because it is one of a few English words whose orthography and pronunciation remain the same across three grammatical functions. <Her> is spelled and pronounced the same when it functions as a direct object, an indirect object, or a possessive adjective. This makes it ideal for investigating any putative correlation between grammatical functions and acoustic phonetic correlates. Twenty (10 females and 10 males) speakers of American English from 10 different states recorded themselves reading the Speech Accent Archive (SAA) elicitation paragraph in which <her> occurs four times surrounded by 31 different words. We extracted F0/pitch, F1, F2, F3, F4, intensity, and duration correlates from <her> and the surrounding words, for a total of 4,340 measured tokens. For this paper, we focus exclusively on F0/pitch, intensity, and duration to test the existence of a putative correlation. Arithmetic means, standard deviations, and interspeaker variability analyses are provided to answer the research question. The findings help to posit the existence of the proximity and the declination principles, as a way of accounting for why correlations exist in some cases but not in others.*

**Keywords:** Acoustic Phonetics and Grammatical Function, Proximity Principle, Declination Principle, Demarcative Pause, Terminal Lengthening

## 1.0 Introduction

This paper seeks to determine any correlation that might exist between the acoustic phonetic features of a word and its grammatical functions. When this issue is brought up, people immediately think of Fry (1955) and Fry (1958). However, the issue being investigated here is altogether different. In the cases investigated by Fry, even a casual observer knows from experience that words such as **per**vert (noun) vs. perv**ert** (verb), **re**cord (noun) vs. re**cord** (verb), etc. are pronounced differently because the location of primary stress is different. Our inquiry is different because we seek to understand if speakers produce the same word differently when its grammatical function changes. A case in point is the word <her> whose spelling remains the same regardless of when it functions as a direct object, an indirect object, or a possessive adjective. In such cases, do speakers pronounce <her> differently or the same? What acoustic phonetic evidence is there to prove that <her> is produced identically or differently across these grammatical functions?

[1]**Authorship responsibilities:** The comitative preposition "with" is used instead of the coordinating conjunction "and" because Author 2 did not contribute to the writing of this paper, except for the measurements that he extracted at the request of Author 1. Author 2 did so because he needed credits to complete his BA in linguistics. Author 1 asked him to extract these measurements in an independent study under his supervision. Author 2 is recognized as such for the measurements he provided. Author 1 bears full responsibility for any analytical or interpretive errors in this publication.

We try to answer to these research questions in five installments. The first introduces the 20 participants and the material. The second provides a succinct review of the literature. The third discusses the Just Noticeable Difference (JND) thresholds that are used to interpret the extracted measurements. The fourth sheds some light on the spectrographic behavior of <her>. The fifth analyzes the three acoustic correlates, that is, F0, intensity, and duration. Under each correlate, we discuss six pairwise comparisons. This allows us to examine the issues as exhaustively as possible.

## 2.0 Material, Participants, and Methodology

The material on which the investigation is based is the elicitation paragraph below, which is available at the Speech Accent Archive (SAA):

> *Please call Stella. Ask **her[1]** to bring these things with **her[2]** from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for **her[3]** brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet **her[4]** Wednesday at the train station.*

As of November 21, 2021, 2,982 participants, among whom there are 649 speakers, have recorded themselves reading this passage out loud.[2] This text is used because it is easily accessible, and claims made about it can be easily refuted or validated by other researchers. The four occurrences of <her> are boldfaced in the text. The superscripts on each occurrence are meant to keep them distinct from each other. Their grammatical functions are listed as follows:

1. <Her[1]> is a direct object pronoun
2. <Her[2]> is an indirect object pronoun
3. <Her[3]> is a possessive adjective
4. <Her[4]> is a direct object pronoun

For each occurrence of <her>, F0/pitch, intensity, and duration measurements are extracted. The 20 participants whose pronunciations are investigated are displayed in Table 1. They come from 10 different states. They are evenly divided by gender: 10 males and 10 females. Additional sociometric information about the speakers is available at https://bit.ly/3CxeluP.

| N0 | Male Participants | Female Participants |
|---|---|---|
| 1. | CA 33M | CA 477F |
| 2. | GA 290M | GA 278F |
| 3. | IL 524M | IL 441F |
| 4. | MN 81M | MN 622F |
| 5. | NY 124M | NY 6F |
| 6. | OR 369M | OR 184F |
| 7. | PA 90M | PA 99F |
| 8. | TX 70M | TX 286F |
| 9. | VA 601M | VA 588F |
| 10. | WA 175M | WA 333F |

Table 1: List of Participants

[2] https://accent.gmu.edu/.

Figure 1 displays the annotation procedures used to extract the correlates. We listened to the audio files and can attest that the participants pronounced <her> normally, that is, it was not given contrastive stress in any of the utterances where it occurred.

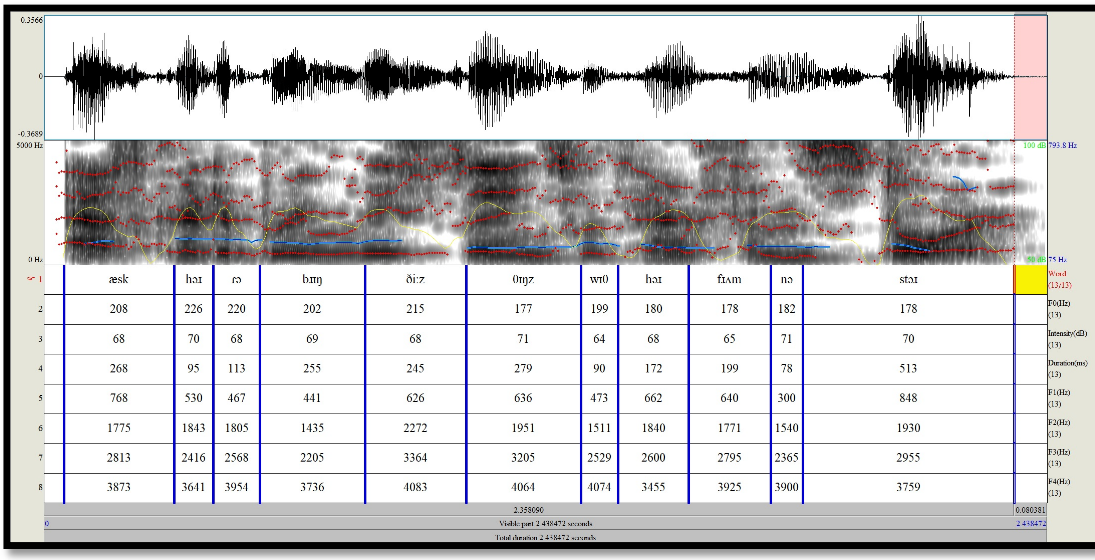


Figure 1: Annotation System <ask her> by PA 99F

All the words in the utterances where <her> occurs were annotated. Boundaries were drawn around each word and relevant measurements were extracted manually. Automatic feature extraction methods were not used because Author 1 does not deem bulk extraction to be reliable, especially because some difficult cases of <her> required visual inspection. These difficult cases brought about rampant resyllabification (see 2.3). The data that was collected is massive, involving 60 spectrographs and 4,340 measured tokens (20 participants x 31 words x 7 correlates) because we extracted seven correlates: F0, F1, F2, F3, F4, intensity, and duration. However, for this paper, we focus only on 240 tokens (20 participants x 4 her x F0 x intensity x duration).

## 2.1 Succinct Literature Review

There is a severe paucity of published data correlating acoustic phonetic measurements and the grammatical functions that words perform in utterances. As far as we can tell, the cases that have been investigated are homographs such as <**sur**vey> (noun) vs. <sur**vey>** (verb). This is not particularly informative because even a casual speaker knows that primary stress falls differently on different syllables in such homographic pairs. Yet, if the same speaker is asked if he/she pronounces <her> the same or differently, as listed above, he/she may not be able to answer this question because, if a difference exists at all, it lies below the threshold of consciousness (Kenstowicz 1994:237). However, for the purposes of enabling smart devices to “speak” as naturally as human beings and to “hear” as effortlessly as them, we must attend to allophonic variations that lie below the threshold of speakers’ consciousness.

## 2.2 The Psychoacoustic Interpretive Framework

Nearly 100 years of acoustic phonetic experimentations have led to the discovery of important Just Noticeable Difference (JND) thresholds at which people with normal hearing can perceive acoustic signals with their naked ears. The ones that are relevant for this paper are listed below. Many of the experiments that led to the discovery of these thresholds are discussed in Koffi (2021a). The JNDs are listed here with no explanation, except for the first one.

**Auditory Discrimination in F0/pitch**

> Of two speech signals **A** and **B** in running speech, **A** is perceived auditorily as having a higher pitch than **B** if there is a difference of 5 Hz or more between them.

This threshold needs a little explanation because there are two JNDs used in the perception of F0/pitch. When words are heard in citation form, i.e., in isolation, the JND for perceiving one segment as having a higher pitch than another is ≥ 1 Hz. However, the JND threshold changes to ≥ 5 Hz when words occur in running speech. The discovery of this threshold goes back to Fry's (1958) seminal paper. He controlled pitch levels at various increments. He found that when two words differed by only 3 Hz, the participants did not reach a consensus as to which word had a high pitch and which had a low pitch. However, when he increased the signals to 5 Hz, a clear consensus emerged among the participants. This JND has been confirmed by other researchers, including Liu (2013:3018). The two remaining JNDs are stated as follows:

**Auditory Discrimination in Intensity**

> Of two speech signals **A** and **B**, **A** is perceived auditorily as louder than **B** if there is a difference of 3 dB or more between them.

**Auditory Discrimination in Duration**[3]

> Of two speech signals **A** and **B** lasting less than 200 msec, **A** is perceived auditorily as longer than **B** if there is a difference of 10 msec or more between them.

These are the JND thresholds that will be used to interpret the acoustic phonetic data on the different occurrences of <her>. When JNDs are used to interpret acoustic measurements, the findings are statistically significant because, according to Stevens (2000:225), for an acoustic correlate to qualify as a JND, 75% or higher correct responses are required.

### 2.3 Spectrographic Evidence of Pronunciation Patterns

The participants did not pronounce <her> exactly the same way all the time. Yet, many of them pronounced them similarly. The spectrographs below represent common pronunciation patterns. We discuss them briefly before attempting to answer the research questions. In <ask her> and <meet her>, where <her> functions as a direct object, a resyllabification rule applies which converts the phrase into a disyllabic word. In both cases, the verbs preceding <her> end in a stop consonant. Most of the participants pronounced <ask her> and <meet her> respectively as [æs$_\sigma$ k$^h$ɚ $_\sigma$] and [mi $_\sigma$ ʔ$^h$ɚ]$_\sigma$. The symbol "$_\sigma$" is used to indicate syllable boundaries. In other words, in every case, the last consonant of the verb was linked to the beginning of <her>. This is exemplified in the pronunciation of NY 6F, as displayed in Figure 2A:

[3] For segments longer than 200 ms, see discussions in the text in section 4.4.

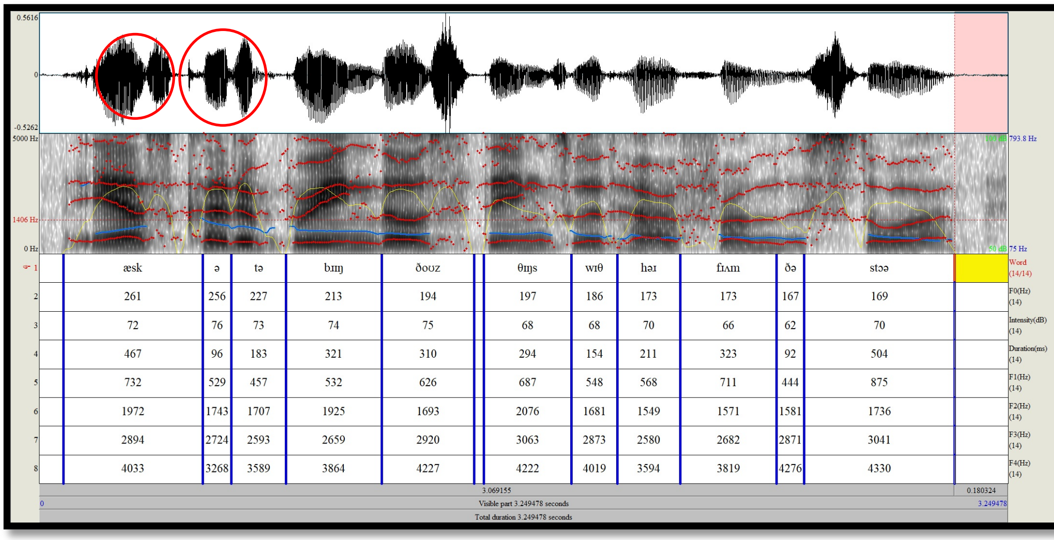


Figure 2A: Spectrographic Evidence of Resyllabification of <ask her> by NY 6F

Here, we see that in <ask her>, the first circle in the spectrograph is around the syllable [æs]$_{\sigma}$. The striations following the vowel show that this syllable ends in an [s]. The second circle contains the second syllable. We see that it begins with the velar stop [k], which is followed by some frication noise.

The phrase <meet her> is pronounced similarly as <ask her>. In Figure 2B, we see that the syllable [mi]$_{\sigma}$ is pronounced as one syllable and [ʔ$^{h}$ɚ]$_{\sigma}$ as another syllable. Furthermore, in [ʔ$^{h}$ɚ]$_{\sigma}$, the voiceless alveolar stop /t/ turns into a glottal stop, followed by frication noise preceding the vowel.

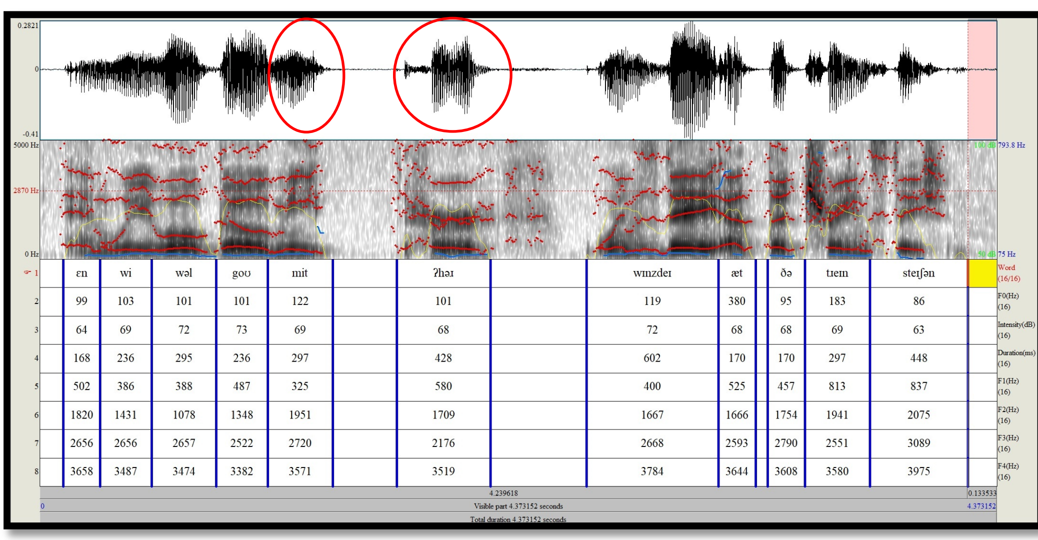


Figure 2B: Spectrographic Evidence of Resyllabification of <meet her> by CA 33M

The pronunciation of <with her> is discussed in Figure 2C. The preposition <with> normally ends with the voiceless interdental fricative [θ]. However, it is commonly produced as a voiceless dental stop [t̪] by some speakers of American English (Koffi 2015). We see this [t̪] in pronunciation NY 6F. The waveform begins with a weak vertical stroke. On the spectrograph, we see a diffused and weak frication noise that corresponds to [h].

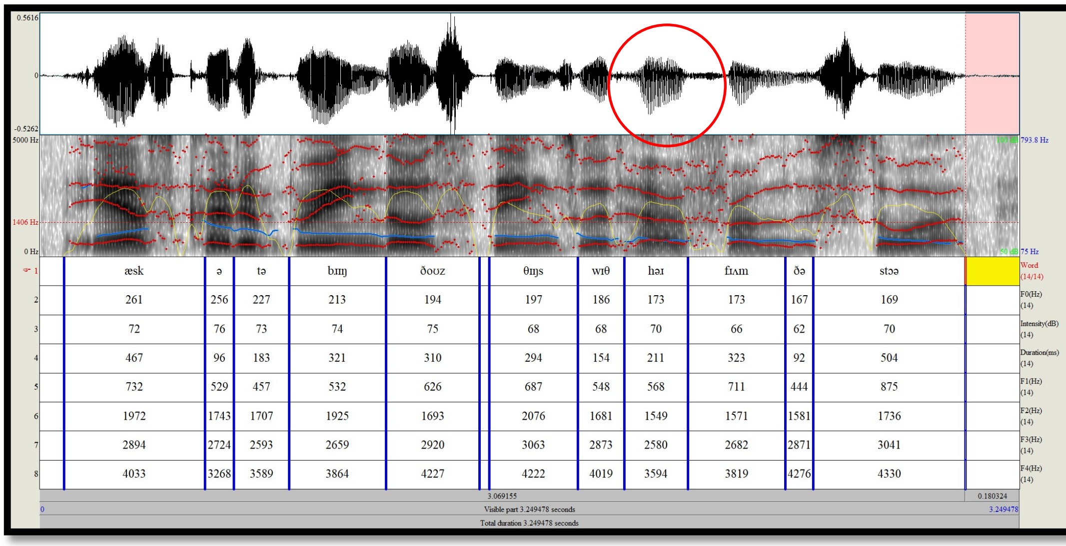


Figure 2C: Spectrographic Evidence of Resyllabification of <with her> by NY 6F

Finally, <her brother> is pronounced by NY 124M as two separate words in Figure 2D. The end of [ɚ] in <her> is different from the beginning of [bɹɑθɚ] because we see clearly the stop gap that shows that the speaker's lips came into positive contact in producing the voiced bilabial stop [b].

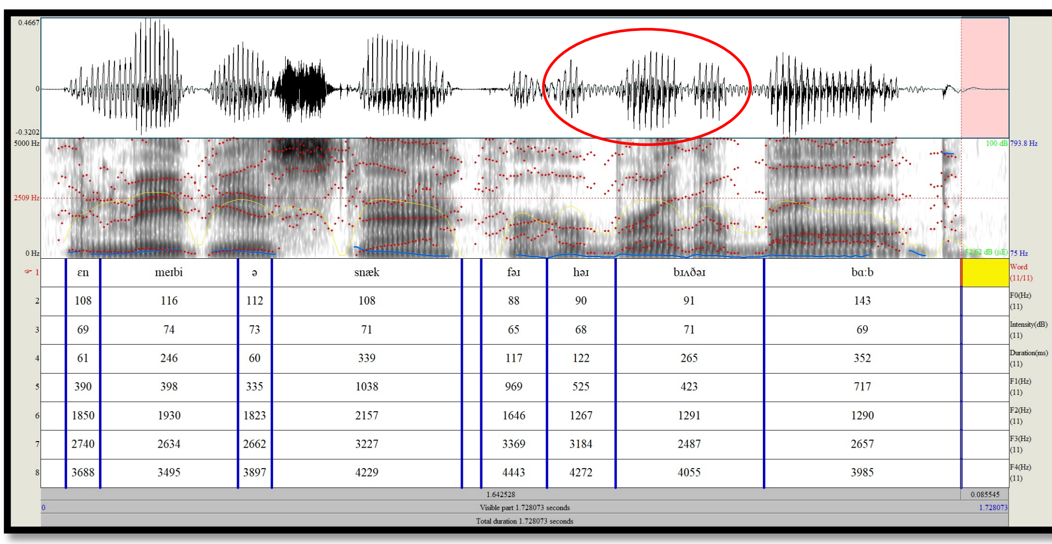


Figure 2D: Spectrographic Evidence of Resyllabification of <with her> by NY 124M

The spectrographic analyses in the preceding sections are important for understanding how the participants pronounced <her>s in the various positions in which they occurred. We have scrutinized the pronunciation patterns of the participants in the minutest of details to derive insights about interspeaker variability. We followed each individual speaker, extracted, and measured how he/she pronounced every occurrence of <her> to see if his/her acoustic features are similar or different from that of other speakers. The **interspeaker** variability analysis affords us as much insights as the arithmetic means and the standard deviations (SDs) that are displayed under each type of <her>. We note in passing that we do not go through the trouble of discussing SDs because all occurrences of <her> are less than 1 SD away from the corresponding mean. This underscores the fact that the speakers pronounced their <her>s almost identically.[4] All the measurements are listed for individual speakers as well as for the entire group. This presentation makes it easy to examine interspeaker variability at a glance. In the discussion sections, reference is made to the superscript on <her> without necessarily pointing out the grammatical function. Male and females' data are displayed separately, but in the analysis, we focus on both genders, unless gender differences are worth pointing out.

### 3.0 F0/Pitch Analysis

F0/pitch measurements indicate how fast the vocal folds vibrate. A higher pitch means that the vocal folds vibrate faster. In the context of this study, a positive correlation between F0/pitch and grammatical function would indicate that when certain lexical items perform certain grammatical functions in an utterance, speakers would match the grammatical function with a corresponding increase or decrease in pitch. We do not claim that the speaker is consciously aware of what he/she is doing. In fact, we contend that much of what goes on phonetically lies below the speaker's threshold of consciousness. Let's scrutinize the measurements in Tables 2A and 2B to see what we can learn from the linguistic behavior of the 20 participants.

| **F0/Male** | **Her[1] Dir. Object** | **Her[2] Ind. Object** | **Her[3] Poss. Adjective** | **Her[4] Di. Object** |
|---|---|---|---|---|
| CA 33M | 112 | 108 | 98 | 101 |
| GA 290M | 137 | 107 | 110 | 220 |
| IL 524M | 176 | 149 | 131 | 119 |
| MN 81M | 110 | 224 | 76 | 184 |
| NY 124M | 138 | 95 | 90 | 106 |
| OR 369M | 166 | 80 | 83 | 125 |
| PA 90M | 136 | 91 | 95 | 97 |
| TX 70M | 179 | 119 | 109 | 124 |
| VA 601M | 77 | 111 | 90 | 105 |
| WA 175M | 203 | 112 | 114 | 121 |
| **Mean** | **143** | **120** | **100** | **130** |
| St. deviation | 38 | 41 | 16 | 40 |

Table 2A: F0 and Grammatical Function by Males

[4] A rule of thumb for knowing what a standard deviation (SD) means is to divide the SD by the arithmetic mean. If the quotient is **less than 1**, it means that the speakers have a similar pronunciation. If the quotient is **higher than 1**, this is an indication that the speakers do not have a similar pronunciation, (see bit.ly/3GRXLZi for more details).

| F0/Female | Her[1] Dir. Object | Her[2] Ind. Object | Her[3] Poss. Adjective | Her[4] Di. Object |
|---|---|---|---|---|
| CA 477F | 213 | 162 | 159 | 180 |
| GA 278F | 290 | 226 | 215 | 248 |
| IL 441F | 328 | 166 | 201 | 196 |
| MN 622F | 269 | 192 | 181 | 190 |
| NY 6F | 256 | 173 | 174 | 208 |
| OR 184F | 257 | 189 | 173 | 198 |
| PA 99F | 226 | 180 | 178 | 150 |
| TX 286F | 259 | 174 | 163 | 194 |
| VA 588F | 352 | 218 | 217 | 225 |
| WA 333F | 300 | 188 | 224 | 206 |
| **Mean** | **275** | **187** | **189** | **200** |
| St. deviation | 43 | 21 | 24 | 26 |

Table 2B: F0 and Grammatical Function by Females

A quick reminder about F0/pitch is in order before we proceed with the various pairwise comparisons of <her>. For running speech, a minimum difference of 5 Hz is needed before it can be said that one speech signal has a higher pitch than another.

### 3.1 <Her[1]> as a Direct Object vs. <Her[2]> as an Indirect Object

Nineteen of the 20 participants (95%) produced <her[1]> differently from <her[2]>. The only exception is CA 33M who produced them the same, i.e., the F0 difference between his two <her>s is 4 Hz, which is below the audibility threshold of 5 Hz. Moreover, of the 19 participants who produced them differently, <her[1]> (143 Hz) has a higher pitch than <her[2]> (120 Hz) among males. The same is true for females, where <her[1]> (275 Hz) is higher than <her[2]> (187 Hz) by 88 Hz. Only MN 81M produced his <her[2]> (224 Hz) higher than his <her[1]> (110 Hz). The remaining participants produced <her[1]> higher than <her[2]>. These measurements indicate that when <her> is a direct object, it has a higher F0 than when it functions as an indirect object.

### 3.2 <Her[1]> as a Direct Object vs. <Her[3]> as a Possessive Adjective

All 20 participants, without exception (100%), produced <her[1]> with a higher pitch than <her[3]>. The arithmetic mean bears this out. For males, the difference between <her[1]> (143 Hz) and <her[3]> (100 Hz) is 43 Hz. For females, <her[1]> (275 Hz) is 86 Hz higher than <her[2]> (189 Hz). In other words, when <her> functions as a direct object, its F0 is higher than when it functions as a possessive adjective.

### 3.3 <Her[2]> as an Indirect Object vs. <Her[3]> as a Possessive Adjective

The correlation is not as clear-cut between <her[2]> and <her[3]> as it is in the two previous cases. The interspeaker variability analysis shows that 6 out of 10 males and 6 out of 10 females, that is, 60% of the participants differentiated between <her[2]> and <her[3]>. Yet, there are important differences between males and females. Among males, the difference between <her[2]> (120 Hz) and <her[3]> (100 Hz) is 20 Hz. Since this is auditorily salient, we would be tempted to conclude that the indirect object is produced with a higher pitch than the possessive adjective. However, among females, the difference of 2 Hz between <her[2]> (187 Hz) and <her[3]> (189 Hz) is not auditorily salient. The correlation appears to be stronger among males than among females. For this reason, we conclude that the correlation between acoustic phonetics and grammatical function is not conclusive for the indirect object and the possessive adjective.

### 3.4 <Her$^{1}$> as a Direct Object vs. <Her$^{4}$> as a Direct Object

In this pairwise comparison, <her$^{1}$> and <her$^{4}$> are both direct objects. So, if a difference exists, it cannot be attributed to grammatical functions. A cursory look at the data shows that males produced <her$^{1}$> (143 Hz) differently from <her$^{4}$> (130 Hz) because they differ by 13 Hz. The same is true for female speakers, because <her$^{1}$> (275 Hz) and <her$^{4}$> (200 Hz) differ by 75 Hz. What accounts for these differences since <her$^{1}$> and <her$^{4}$> fulfill the same grammatical function as direct objects? The answer lies in the syntactic distribution of the two <her>s. In <her$^{1}$>, <her> is the second word in the utterance, but in <her$^{4}$>, it is the 15$^{th}$ word in the utterance. The difference in syntactic distribution is responsible for the variation in pitch. The interspeaker variability analysis bears this out. Seventeen of 20 speakers (85%) produced <her$^{1}$> with a higher pitch than <her$^{4}$>. Since both <her$^{1}$> and <her$^{4}$> function as direct objects, how can the difference in pitch between them be explained? The difference can be attributed to **declination**. It is a well-known phenomenon that pitch goes gradually down from the beginning of an utterance to its end. The specific label that is used to describe this phenomenon is **"terminal pitch fall**." Its converse is called **terminal pitch rise** or **uptalk** (Curzan and Adams 2006:124-5). Three participants, GA 290M, MN 81M, and VA 601M, produced <her$^{4}$> with a terminal pitch rise. The insights derived from <her$^{1}$> and <her$^{4}$> allow us to posit the existence of two principles in intonation. The first is the **"proximity principle"** and the second is the **"declination principle."** The proximity principle states that elements closer to the main verb receive a higher pitch than elements that are further away from it.[5] When elements are directly inside the orbit of the main verb, the pitch of the verb carries over to that element. This can manifest itself phonologically by resyllabification. By the same token, elements that lie outside of the direct orbit of the verb are less likely to receive pitch prominence.

### 3.5 <Her$^{4}$> as a Direct Object vs. <Her$^{2}$> as an Indirect Object

The proximity principle helps to explain why the pitch of <her$^{4}$> is higher than that of <her$^{2}$> among males and females. For males, the F0 of <her$^{4}$> (130 Hz) is higher by 10 Hz than <her$^{2}$> (120 Hz). Similarly, the F0 of <her$^{4}$> (200 Hz) is 13 Hz higher than <her$^{2}$> (187 Hz) among females. The interspeaker variability analysis indicates that all 10 males and 8 females, that is, 18 out of 20 participants (90%) produced <her$^{4}$> and <her$^{2}$> differently. Furthermore, 14 of 18 participants (77.77%) produced <her$^{4}$> higher than <her$^{2}$>. Given that <her$^{4}$> occurs near the end of an utterance, one would expect that its pitch would be lower than that of <her$^{2}$>. But it is not. This means that the proximity principle takes precedence over the declination principles for elements that occur directly in the orbit of the main verb. We can postulate that, all things being equal, direct objects are likely to have a higher F0/pitch than indirect objects.

### 3.6 <Her$^{4}$> as a Direct Object vs. <Her$^{3}$> as a Possessive Adjective

The final comparison has to do with <her$^{4}$> and <her$^{3}$>. Here also, males produced <her$^{4}$> (130 Hz), that is, 30 Hz higher than <her$^{3}$> (100 Hz). Females also produced <her$^{4}$> (200 Hz) higher than <her$^{3}$> (189 Hz) by 11 Hz. The interspeaker variability analysis shows that 18 of the of participants 20 (90%) produced <her$^{4}$> higher than <her$^{3}$>. Among the 18 who produced them differently, 15 (83.33%) produced <her$^{4}$> with higher pitch than <her$^{3}$>. This observation is particularly important because, given the syntactic distribution of <her$^{4}$>, the declination principle would suggest that its F0 should be lower than that of <her$^{3}$>. The fact that it is higher than <her$^{3}$>

[5] Italics are used to underscore important findings.

suggests that the proximity principle is robust. In other words, all things being equal, direct objects are likely to have a higher F0/pitch than possessive adjectives.

### 3.7 Interim Summary

The six-way comparisons in the previous sections indicate that F0/pitch differences correlate with grammatical functions in all cases but two. In <her$^1$> and <her$^4$> the differences cannot be attributed to grammatical functions because both <her>s function as direct objects. Instead, the differences are attributable to the positions in which the two <her>s occur. In <her$^1$>, it is the second word, whereas in <her$^4$>, it is the 15$^{th}$ word. Here, the declination principle helps to explain why, even though <her$^1$> and <her$^4$> have the same grammatical function, their pitch measurements are different. F0/pitch does not conclusively differentiate between the indirect object <her$^2$> and the possessive adjective <her$^3$>. Here, the proximity principle does not apply because in both cases, the elements do not occur within the orbit of the main verb. In general, where the proximity principle fails to apply, pronunciation differences emerge between different speakers.

### 4.0 Intensity Analysis

Intensity is somewhat related to loudness. If a correlation exists between intensity and grammatical function, it entails that some words in an utterance are pronounced louder than others simply because they fulfill certain grammatical functions. In the context of this paper, it means that some instances of <her> will be louder than others. Let's examine the data in Tables 3A and 3B to see if this is indeed the case.

| **Intensity** | **Her$^1$ Dir. Object** | **Her$^2$ Ind. Object** | **Her$^3$ Poss. Adjective** | **Her$^4$ Di. Object** |
|---|---|---|---|---|
| CA 33M | 69 | 72 | 65 | 68 |
| GA 290M | 80 | 76 | 73 | 71 |
| IL 524M | 67 | 67 | 61 | 65 |
| MN 81M | 77 | 71 | 69 | 70 |
| NY 124M | 72 | 71 | 68 | 71 |
| OR 369M | 82 | 76 | 73 | 77 |
| PA 90M | 72 | 67 | 66 | 70 |
| TX 70M | 76 | 73 | 68 | 75 |
| VA 601M | 70 | 72 | 66 | 72 |
| WA 175M | 76 | 68 | 65 | 66 |
| **Mean** | **74** | **71** | **67** | **71** |
| St. Dev. | 4 | 3 | 3 | 3 |

Table 3A: Intensity and Grammatical Function by Males

| **Intensity** | **Her$^1$ Dir. Object** | **Her$^2$ Ind. Object** | **Her$^3$ Poss. Adjective** | **Her$^4$ Di. Object** |
|---|---|---|---|---|
| CA 477F | 71 | 68 | 64 | 70 |
| GA 278F | 73 | 71 | 70 | 71 |
| IL 441F | 76 | 70 | 72 | 70 |
| MN 622F | 71 | 68 | 66 | 67 |
| NY 6F | 76 | 70 | 67 | 72 |
| OR 184F | 75 | 73 | 68 | 71 |
| PA 99F | 70 | 68 | 71 | 59 |
| TX 286F | 76 | 71 | 69 | 72 |
| VA 588F | 78 | 77 | 75 | 73 |
| WA 333F | 76 | 72 | 72 | 72 |
| **Mean** | **74** | **71** | **69** | **70** |
| St. Dev. | 2 | 2 | 3 | 4 |

Table 3B: Intensity and Grammatical Function by Females

In interpreting intensity measurements, it is good to be reminded once again that differences less than 3 dB between two speech signals are not perceived by the human ear in every day listening conditions.

### 4.1 <Her$^1$> as a Direct Object vs. <Her$^2$> as an Indirect Object

In both male and female speech, <her$^1$> (74 dB) and <her$^2$> (71 dB) differ in intensity by 3 dB, which means that <her$^1$> is louder than <her$^2$>. This is also confirmed by the interspeaker variability analysis. Fourteen out 20 participants (70%), 7 out 10 males and 7 out 10 females, produced <her$^1$> and <her$^2$> differently in accordance with grammatical functions. Of the 14 who differentiated between them, 13 (92.85%) produced <her$^1$> louder than <her$^2$>. When <her$^1$> is a direct object, it is louder than when it is an indirect object. Here, as in section 3.7, the proximity principle operates and accounts for the difference.

### 4.2 <Her$^1$> as a Direct Object vs. <Her$^3$> as a Possessive Adjective

The proximity principle also helps to explain why <her$^1$> (74 dB) is louder than <her$^3$> (67 dB) in male and in female (69 dB) speech. The intensity differences between <her$^1$> and <her$^3$> are auditorily very salient. For males, the difference is 5 dB; for females, it is 7 dB. This means that if <her$^1$> and <her$^3$> are played one right after the other, a person listening with their naked ear will perceive the difference between them clearly. The interspeaker variability analysis shows that 19 out of 20 participants (95%), all 10 males and 9 out of 10 females, produced <her$^1$> louder than <her$^3$>. Only PA 99F produced them similarly. This sonority difference can also be explained by the proximity principle. The direct object <her$^1$> occurs within the orbit of the transitive verb <ask>, whereas in <her$^3$> there is no verb.

### 4.3 <Her$^2$> as an Indirect Object vs. <Her$^3$> as a Possessive Adjective

Here, neither <her$^2$> nor <her$^3$> occurs within the orbit of a main verb. As a result, the proximity principle is inoperative. What do we have in such a case? There is a difference between males and females in how loud <her$^2$> and <her$^3$> are. For males, 8 out of 10 (80%) produced <her$^2$> (71 dB) louder than <her$^3$> (67 dB) by 4 dB. Among females, the there is no audible difference between <her$^2$> (71 dB) and <her$^3$> (69 dB) because the 2 dB that separates them is below the threshold of audibility. The interspeaker variability analysis shows that five females produced <her$^2$> higher than <her$^3$>, while five others did the exact opposite. The observation made in section 3.7, namely, where the proximity principle fails to apply, pronunciation differences emerge between different speakers is seen here also. Male speakers go one way, while females go the other way. Even among females, there is no congruence. Half produced <her$^{2}$> louder than <her$^{3}$>, while the other half produced <her$^{3}$> louder than <her$^{2}$>.

### 4.4 <Her$^{1}$> as a Direct Object vs. <Her$^{4}$> as a Direct Object

There is a loudness difference between <her$^{1}$> (74 dB) vs. <her$^{4}$> (71 dB) in male speech. The same is true among females, <her$^{1}$> (74 dB) vs. <her$^{4}$> (70 dB). However, this difference cannot be attributed to differences in grammatical function because both <her$^{1}$> and <her$^{4}$> are direct objects. So, why do we have an intensity difference? A difference exists because of the syntactic distributions of the two words, as noted previously. In <ask her>, <her$^{1}$> is the second word in the sentence, whereas in <meet her>, <her$^{4}$> is the 15$^{th}$ word in the utterance. As was noted before, terminal sonority fall applies to <her$^{4}$>. Because of this, a lower sonority on <her$^{4}$> is to be expected. The declination principle is responsible for why <her$^{1}$> is louder than <her$^{4}$> even though they both have the same grammatical function. The interspeaker variability analysis shows that 16 out 20 participants, 80%, (8 males and 8 females), produced <her$^{1}$> louder than <her$^{4}$>. Only CA 33M, IL 524M, IL 441F, and GA 278F did not.

### 4.5 <Her$^{4}$> as a Direct Object vs. <Her$^{2}$> as an Indirect Object

The intensity differences between <her$^{4}$> (71 dB) and <her$^{2}$> (71 dB) is neither auditorily perceptible in male pronunciation nor in female speech (70 dB vs. 71 dB). Since <her$^{4}$> is in the immediate orbit of the verb, we expect its sonority to be higher than <her$^{2}$>. However, it is not because it occurs near the end of the utterance. The proximity principle is inoperative near or at the end of utterances because it is trumped by the declination principle. The interspeaker variability analysis bears this out. Only CA 33M, GA 290M, PA 99F, and VA 588F produced <her$^{4}$> and <her$^{2}$> differently. Most of the speakers, 16 out 20 (80%), produced <her$^{4}$> and <her$^{2}$> similarly in intensity.

### 4.6 <Her$^{4}$> as a Direct Object vs. <Her$^{3}$> as a Possessive Adjective

In male speech, <her$^{4}$> (71 dB) is louder than <her$^{3}$> (67 dB). The interspeaker variability analysis shows that 7 out of 10 males (70%), produced <her$^{4}$> louder than <her$^{3}$>. Yet, GA 290M, MN 81M, and WA 175M did not differentiate between them in loudness. Among females, the arithmetic mean shows that no audible difference exists between <her$^{4}$> (70 dB) and <her$^{3}$> (69 dB). The participants are split in how they pronounced <her$^{4}$> and <her$^{3}$>. Five of them produced <her$^{4}$> louder than <her$^{3}$>, while five others did the exact opposite. This situation is analogous to the one in section 4.3 in that, when the proximity principle is inoperative, there is no clear correlation between acoustic phonetics and grammatical function. As a result, pronunciation varies between groups of speakers.

### 4.7 Interim Summary

Three important insights flow from the analyses above. The first is that when the proximity principle applies, lexical items in the orbit of the main verb are louder than those that are outside. A notable exception is when that lexical item occurs at or near the end of an utterance. In that case, the declination principle trumps the proximity principle. The second insight is that when the proximity principle fails to apply, a gender difference emerges between males and females. The following loudness hierarchy, her$^{1}$ > her$^{2}$ > her$^{3}$ > her$^{4}$>, is found among male speakers. However, in female speech, the hierarchy is limited to her$^{1}$ > her$^{2}$ > because <her$^{2}$> is not louder than <her$^{3}$>, nor is <her$^{3}$> louder than <her$^{4}$>. Finally, in both male and female pronunciations, the possessive adjective <her³> is the least sonorous. The interspeaker variability analysis confirms that 17 out 20 participants (85%) produced <her³> quieter than <her¹>, <her²>, or < her⁴>. It is premature to make a generalization about genitive constructions from this data set. Yet, the findings so far point to the fact that the possessive adjective <her³> is not very loud, unless it is used in contrastive stress utterances.

## 5.0 Duration Analysis

If a correlation exists between duration and grammatical function, it will entail that speakers are likely to pronounce such and such a word longer or shorter simply because it fulfills such and such grammatical function. Before tackling the existence of any such correlation, we will do well to review what Klatt (1976) and Koffi (2021b) have to say about segmental duration. The review here is intentionally brief. Klatt discusses various factors that affect the duration of segments. This includes speaking rate, semantic factors, discourse-level factors, syntactic factors, and demarcative pauses. Koffi (2021b:16) proposes a durational hierarchy for English segments in running speech. The take-away is that the intrinsic durational characteristics of English segments can change substantially for a wide variety of reasons. As for the case of <her>, it was noted in section 2.3 that [h] is pronounced differently according to its syntactic distributions. In <ask her> and <meet her>, resyllabification caused the [h] of <her> not to be fully aspirated because it became linked with the preceding voiceless stops [k] and [t]. However, in <with her> and <her brother>, the [h] of <her> was pronounced fully.

It goes without saying that the syntactic distributions can also affect the duration of <her>. For this reason, we need to pay closer attention to where it occurs in the four utterances. <Her¹> occurs before a subordinate clause. <Her²> is at the end of a prepositional phrase (PP) <with her> next to another PP <from the store>. At the juncture of these two PPs, a slight pause is usually observed. <Her³> is part of a noun phrase (NP) <her brother>. In contemporary syntax, <her³> is called a determiner, but in traditional grammar, it is known as a possessive adjective. There is usually not a pause between the determiner and the noun. Finally, <her⁴> in <meet her> occurs inside of a verb phrase (VP) which is immediately followed by an adverbial phrase (AdvP) <Wednesday>. In such cases, a slight demarcative pause occurs at phrasal boundaries. These four syntactic environments can affect the duration of <her>. Varying degrees of terminal lengthening are also expected to impact the duration of <her>, as shown in Tables 4A and 4B:

| **Duration** | **Her¹ Dir. Object** | **Her² Ind. Object** | **Her³ Poss. Adjective** | **Her⁴ Di. Object** |
|---|---|---|---|---|
| CA 33M | 129 | 135 | 161 | 428 |
| GA 290M | 101 | 172 | 181 | 194 |
| IL 524M | 80 | 146 | 167 | 270 |
| MN 81M | 66 | 123 | 147 | 161 |
| NY 124M | 60 | 86 | 122 | 99 |
| OR 369M | 63 | 117 | 128 | 86 |
| PA 90M | 167 | 194 | 269 | 174 |
| TX 70M | 76 | 69 | 129 | 144 |
| VA 601M | 82 | 94 | 175 | 135 |
| WA 175M | 68 | 96 | 111 | 87 |
| **Mean** | **89** | **123** | **159** | **178** |
| St. Dev. | 34 | 39 | 45 | 104 |

Table 4A: Duration and Grammatical Function by Males

| **Duration** | **Her$^{1}$ Dir. Object** | **Her$^{2}$ Ind. Object** | **Her$^{3}$ Poss. Adjective** | **Her$^{4}$ Di. Object** |
|---|---|---|---|---|
| CA 477F | 219 | 296 | 157 | 320 |
| GA 278F | 136 | 234 | 132 | 233 |
| IL 441F | 94 | 179 | 206 | 223 |
| MN 622F | 65 | 122 | 114 | 193 |
| NY 6F | 96 | 211 | 199 | 176 |
| OR 184F | 83 | 169 | 153 | 126 |
| PA 99F | 95 | 172 | 139 | 184 |
| TX 286F | 95 | 86 | 96 | 155 |
| VA 588F | 116 | 113 | 129 | 151 |
| WA 333F | 88 | 129 | 122 | 179 |
| **Mean** | **109** | **171** | **145** | **194** |
| St. Dev. | 31 | 63 | 35 | 54 |

Table 4B: Duration and Grammatical Function by Females

Since the four <her>s last less than 200 msec, the JND threshold of 10 msec is used for all the analyses below.

### 5.1 <Her$^{1}$> as a Direct Object vs. <Her$^{2}$> as an Indirect Object

<Her$^{1}$> and <her$^{2}$> occur in a complex sentence. The first is found in the main clause, while the second appears in the subordinate clause. Moreover, as noted earlier, the [h] of <her$^{1}$> is not pronounced fully because it is resyllabified with the preceding voiceless stop [k]. This explains why <her$^{1}$> (89 msec) is shorter than <her$^{2}$> (123 msec) by 34 msec in male speech. <Her$^{1}$> (109 msec) is also shorter than <her$^{2}$> (171 msec) in female speech by 62 msec. In both cases, the durational difference is auditorily salient. The interspeaker variability analysis shows that 16 of the 20 participants (80%) produced <her$^{1}$> shorter than <her$^{2}$>, except for CA 33M, TX 70M, TX 286F, and VA 588F who produced them similarly. The proximity principle has a shortening effect on <her$^{1}$>. The syntactic environment is responsible for this. As has been noted previously, in <ask her>, the direct object is in the immediate orbit of the main verb. As a result, there is no pause between <ask> and <her>. In fact, because of resyllabification, they are pronounced as a single lexical item. <Her$^{2}$> is longer than <her$^{1}$> because it lies outside of the orbit of the main verb. Moreover, <her$^{2}$> is longer than <her$^{1}$> because a demarcative pause occurs between it and the PP <from the store>. Demarcative pauses between phrases lead to slight terminal lengthening (Hyman 1988:443-470).

### 5.2 <Her$^{1}$> as a Direct Object vs. <Her$^{3}$> as a Possessive Adjective

The proximity principle is at work here also and explains why <her$^{1}$> (89 msec) is shorter than <her$^{3}$> (159 msec) by 70 msec in male speech. Among females, <her$^{1}$> (109 msec) is still shorter than <her$^{3}$> (145 msec) by 36 msec. The interspeaker variability analysis shows that all 10 male and 8 female speakers out 20 participants (90%) produced <her$^{1}$> shorter than <her$^{3}$>. The exceptions are GA 278F and TX 286F. The proximity principle does not apply to <her$^{3}$> because it is not part of a VP. In fact, in <her brother>, the [h] of <her$^{3}$> is fully produced as a voiceless glottal fricative, but the [h] of <her$^{1}$> in <ask her> is linked to [k], as shown in Figure 2A. The fact that the [h] in <her$^{3}$> is outside of the orbit of the main verb and is pronounced fully explains why it is longer than <her$^{1}$>. Ninety percent of the participants produced it this way.

Linguistic Portfolios – ISSN 2472-5102 –Volume 11 – 2022| 128

### 5.3 <Her$^{2}$> as an Indirect Object vs. <Her$^{3}$> as a Possessive Adjective

<Her$^{2}$> and <her$^{3}$> both lie outside of the orbit of the main verb. In both instances, [h] is pronounced fully. Yet, we notice an important gender difference in how they are pronounced. Males pronounced <her$^{3}$> (159 ms) longer than <her$^{2}$> (123 msec) by 36 msec. The interspeaker variability analysis indicates that all 10 males produced <her$^{3}$> longer than <her$^{2}$>. Among females, there is also a difference of 26 msec between <her$^{2}$> (171 msec) and <her$^{3}$> (145 msec), but this time <her$^{3}$> is longer than <her$^{2}$>. In fact, except for IL 441F and VA 588F, the remaining females produced <her$^{3}$> longer than <her$^{2}$>. The gender difference can be interpreted to mean that females privileged terminal lengthening at the juncture between <with her> in <her$^{2}$> as opposed to their male counterparts who opted for pronouncing <her$^{3}$> in <her brother> fully. Both pronunciations are possible but result in different phonetic realizations. The lack of a consensus on the duration of <her> lends support to the claim that has been made earlier about the proximity principle. Wherever it is inoperative, we tend to have different pronunciation patterns.

### 5.4 <Her$^{1}$> as a Direct Object vs. <Her$^{4}$> as a Direct Object

The phonological and syntactic environments of <her$^{1}$> and <her$^{4}$> are identical. Furthermore, they both function as direct objects. As a result, we do not expect to see any duration difference between them. However, there are differences. Among males, <her$^{4}$> (178 msec) is longer than <her$^{1}$> (89 msec) by 89 msec. The same is true for females, where <her$^{4}$> (194 msec) is longer than <her$^{1}$> (109 msec) by 85 msec. What accounts for these significant differences between <her$^{1}$> and <her$^{4}$> given that they are similar phonologically, syntactically, and functionally? The answer is found in their distribution. In <ask her>, <her$^{1}$> occurs in the main clause followed by the infinitive marker of the subordinate clause. There is only a slight pause at the juncture between <her$^{1}$> and <to>. <Her$^{4}$>, on the other hand, occurs at the juncture of the VP and the AdvP <Wednesday>. In such an environment, the demarcative pause is longer. As a result, a terminal lengthening rule applies. This explains why <her$^{4}$> is much longer than <her$^{1}$>. This also confirms Klatt's (1974:1211) observation that a word that occurs before a pause is lengthened by as much as 60 msec compared to when the same word occurs elsewhere. The interspeaker variability analysis shows that 19 out 20 participants (95%) produced <her$^{4}$> considerably longer than <her$^{1}$>, except for PA 90M who pronounced them with the same duration. So, even when two lexical items fulfill the same grammatical function and undergo the same phonological process, demarcative pauses can cause their durations to be different, as is the case of <her$^{1}$> and <her$^{4}$>.

### 5.5 <Her$^{4}$> as a Direct Object vs. <Her$^{2}$> as an Indirect Object

The demarcative pause causes <her$^{4}$> (178 msec) to be 55 msec longer than <her$^{2}$> (123 msec) among male speakers. Among females also, <her$^{4}$> (194 msec) is 23 msec longer than <her$^{2}$> (171 msec). This is so even though the [h] of <her$^{4}$> is weakly pronounced, while that of <her$^{2}$> is fully articulated. The intraspeaker variability analysis shows that 14 out 20 participants (70%) produced <her$^{4}$> longer than <her$^{2}$>, except for OR 369M, PA 90M, NY 124M, WA 175M, NY 6F, and OR 184F. This means that the type of demarcative pause influences the degree of terminal lengthening.

### 5.6 <Her$^{4}$> as a Direct Object vs. <Her$^{3}$> as a Possessive Adjective

The effect that demarcative pauses have on duration is again underscored by <her$^{4}$> and <her$^{3}$>. Among males, <her$^{4}$> (178 msec) is 19 msec longer than <her$^{3}$> (159 msec). Females also produced <her$^{4}$> (194 msec) longer than <her$^{3}$> (145 msec) by 49 msec. Five males and 8 females (65%) produced <her$^{4}$> longer than <her$^{3}$>, while five males (NY 124M, OR 369M, PA 90M, VA 601M, and WA 175M) and two females, NY 6F and PA 99F, produced <her$^{3}$> longer than <her$^{4}$>.

### 5.7 Interim Summary

Two observations can be made from the analyses of duration. First, demarcative pauses play a more prominent role in duration than grammatical function. This explains why <her$^{4}$> is longer than <her$^{2}$> and <her$^{3}$>. Secondly, the proximity principle causes shortening because of the resyllabification rule that takes place between the main verb and the direct object. This explains why <her$^{1}$> is shorter than <her$^{2}$> and <her$^{3}$>.

## Summary

Pitch and intensity converge to bolster the view that there is some correlation between acoustic phonetic features and grammatical function. This correlation is a limited one because it applies only when the proximity principle is operative. In all the cases of <her> investigated in this paper, pitch and intensity are higher if <her> occurs immediately in the orbit of the main verb, that is, when it functions as a direct object. The further it is from the main verb, the lesser its pitch and its intensity. Exceptions to the proximity principle exist and can be accounted for by the declination principle. The proximity principle operates differently regarding the duration correlate. When it is operative, it causes <her> to undergo shortening. This is the reason why <her$^{1}$> is shorter than <her$^{2}$> and <her$^{3}$>. An exception such as <her$^{4}$> is easily explained by the presence of a demarcative pause that brings about terminal lengthening. This is the reason why <her$^{4}$> is longer than <her$^{1}$>, <her$^{2}$>, and <her$^{3}$>. More generally, exceptions to the proximity principle call for greater interspeaker variabilities.

## ABOUT THE AUTHOR

**Ettien Koffi**, Ph.D. linguistics (Indiana University, Bloomington, IN) teaches at Saint Cloud State University, MN. He is the author of five books and author/co-author of several dozen articles on acoustic phonetics, phonology, language planning and policy, emergent orthographies, syntax, and translation. His acoustic phonetic research is synergetic, encompassing L2 acoustic phonetics of English (Speech Intelligibility from the perspective of the Critical Band Theory), sociophonetics of Central Minnesota English, general acoustic phonetics of Anyi (a West African language), acoustic phonetic feature extraction for application in Automatic Speech Recognition (ASR), Text-to-Speech (TTS), and voice biometrics for speaker verification. Since 2012, his high impact acoustic phonetic publications have been downloaded **28,515** times (analytics provided by Digital Commons) and read **13,680** times (analytics provided by Researchgate.net) as of February 2022. He can be reached at enkoffi@stcloudstate.edu.

**Matt Beckstrand** is a 2021 graduate from St. Cloud State University in Minnesota, USA. Unfortunately, upon graduation, his school email address expired. So, he could not be reached for his personal email address.

**Works Cited**

Boersma, Paul, and David Weenink. 2019. *Praat*. Praat: doing Phonetics by Computer. https://www.fon.hum.uva.nl/praat/.

Anne Curzan and Michael Adams. 2006. *How English Works: A Linguistic Introduction.* 1st Edition New York: Pearson.

Fry, Dennis. B. 1958. Experiments in the Perception of Stress. *Language and Speech* 1 (2): 126-152.

Hyman, Larry. 1978. Word Demarcation. In *Universals of Human Language*, Volume 2, Phonology, ed. by Joseph H. Greenberg, pp. 443-470. Stanford, CA: Stanford University Press.

Jongman, A., Wayland, R., Wong, S. (2000). Acoustic Characteristics of English Fricatives. *Journal of the Acoustical Society of America*. 108: 3(1) 1252-1263

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Cambridge, MA: Blackwell Publishers.

Koffi, Ettien. 2021a. *Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility*. Boca Raton, Florida: CRC Press.

Koffi, Ettien. 2021b. A Comprehensive Review of Duration and Its Linguistic Implications. *Linguistic Portfolios* 10: 2-27.

Koffi, Ettien. 2020. A Comprehensive Review of Intensity and Its Linguistic Implications. *Linguistic Portfolios* 10: 2-28.

Koffi, Ettien. 2015. The Pronunciation of Voiceless TH. *Linguistic Portfolios* 10: 2-23.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: Acoustic and Perceptual Evidence. *Journal of the Acoustical Society of America* 59 (5): 1206–1221.

Liu, Chang. 2013. Just Noticeable Difference of Tone Pitch Contour Change for English and Chinese Native Listeners. *Journal of the Acoustical Society of America*, 134 (4): 3011-3020.

Stevens, Kenneth N. 2000. Acoustic Phonetics. Cambridge, MA: The MIT Press.

t' Hart, Johan. 1981. Differential Sensitivity to Pitch Distance, Particularly in Speech. *Journal of the Acoustical Society of America*, 69 (3): 811-821.

Weinberger, Steven. H. 2020. *The Speech Accent Archive*. Speech Accent Archive. http://accent.gmu.edu/index.php.