UC San Diego

San Diego Linguistic Papers, Issue 8

Title

Intrinsic fundamental frequency of Amharic vowels

Permalink

Author

Van Doren, Maxine

Publication Date

2021-01-15

Intrinsic fundamental frequency of Amharic vowels

Maxine Van Doren

Department of Linguistics, University of California San Diego

9500 Gilman Dr.

La Jolla, CA 92093, USA

mvandore@ucsd.edu

Intrinsic fundamental frequency (IF0) is the cross-linguistic tendency for high vowels to have a higher fundamental frequency (F0) than low vowels. IF0 is often thought to result from mechanical coupling between the tongue and larynx, though the exact causal relationship remains unknown. The purpose of the present study is to test the predictions of four articulatory hypotheses using audio and electroglottographic data from Amharic vowels. The current study analyzed data from 6 adult native speakers of Amharic. F0 and contact quotient (CQ) were measured using audio and electroglottography (EGG) recordings. The results suggest that IF0 may have multiple sources. The acoustic analysis reveals that F0 is positively correlated with vowel height: as F1 lowers, F0 increases. This may be due to a combination of jaw lowering and tongue compression. However, contrary to some articulatory hypotheses, CQ results suggest that changes in tongue or jaw height do not induce changes on vocal fold tension. In addition, a phonological grouping of vowel F0 emerged, such that vowels within a phonological height category did not differ from each other in terms of F0, even if they differed in F1. This finding suggests that additional factors, such as auditory enhancement, may also contribute to the IF0 effect.

I. INTRODUCTION

Cross-linguistically, the fundamental frequency (F0) of vowels has been found to correlate with vowel height, with high vowels having a higher F0 than low vowels (Whalen & Levitt, 1995). This phenomenon is referred to as the intrinsic fundamental frequency of vowels (IF0) or intrinsic pitch of vowels (IPV). In a study of 35 languages representing 11 different language families, Whalen and Levitt (1995) found evidence of intrinsic F0 of vowels in every reported language, with effect size varying from 5 Hz to over 20 Hz (Whalen & Levitt, 1995). The authors suggest that the phenomenon is universal, but the cause of the phenomenon remains unclear. Commonly cited hypotheses attribute IF0 to mechanical coupling of the oral vocal tract and laryngeal structures, where the F0 perturbation is due to tongue or jaw movement inducing changes on laryngeal structures (Ladefoged, 1964; Ohala, 1978; Ewan, 1979; Chen, Whalen & Tiede, 2019). Yet, precisely how tongue or jaw movement might lead to changes in vocal fold vibration remains unclear. Other hypotheses posit that IF0 is due to volitional control for vowel production, either as auditory enhancement to cue vowel category (Diehl, 1991) or as part of the articulatory gesture for the vowel (Sapir, 1989). The goal of the present study is to test predictions of articulatory accounts of IF0 and examine the relationship between oral vocal tract movement (i.e., tongue and jaw) and laryngeal movement proposed by each hypothesis.

a. Background

Perhaps the most frequently cited account of IF0 is the TONGUE-PULL HYPOTHESIS, originally proposed by Ladefoged (1964). The tongue-pull hypothesis states that IF0 is driven by the tongue body's high position pulling on the hyoid bone. The upward movement of the hyoid bone in turn pulls the thyroid cartilage forward, leading to increased laryngeal height and vocal fold tension.

Together, the increased laryngeal height and vocal fold tension are believed to raise F0 (Ladefoged, 1964). This is schematized in Figure 1.

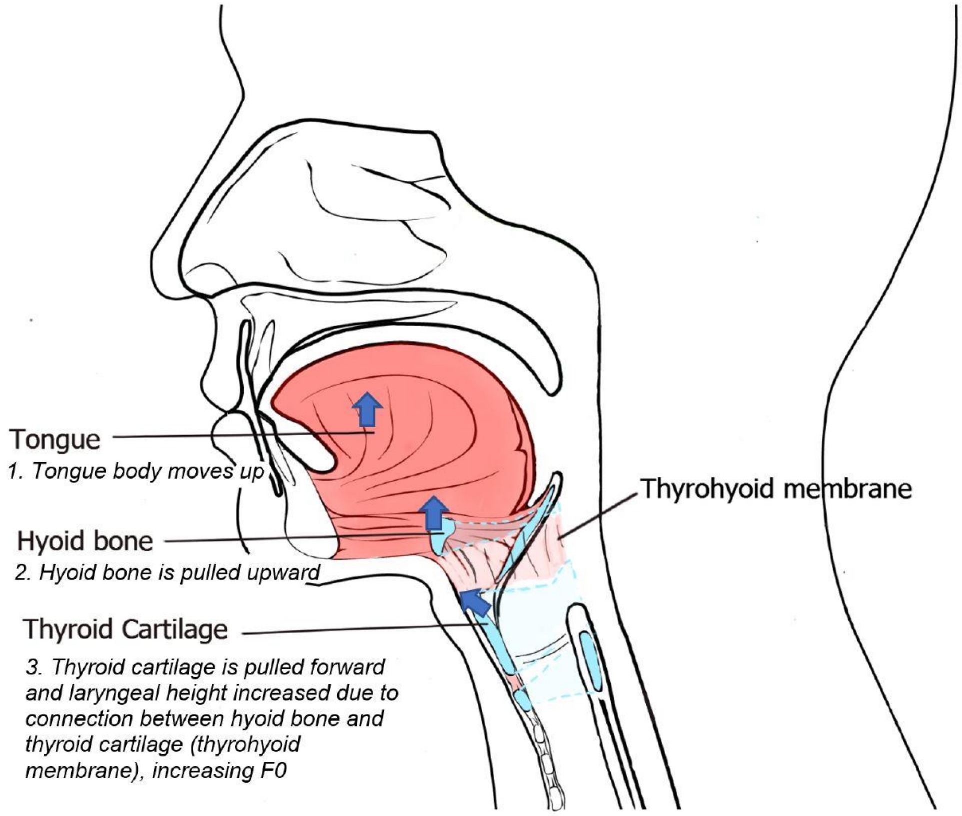


Figure 1. Schematic of the tongue-pull hypothesis with arrows indicating direction of articulator movement. As tongue body moves upward (1), the tongue pulls on the hyoid bone (2), causing the thyroid cartilage to move forward (increased laryngeal height) (3), which leads to increased F0 (Color online). Image source: OpenStax AnatPhys fig.22.6 - Divisions of the Pharynx - English labels" at AnatomyTOOL.org by OpenStax, license: Creative Commons Attribution. Source: book 'Anatomy and Physiology', https://openstax.org/details/books/anatomy-and-physiology. Adapted by B.

Evidence for the tongue-pull hypothesis has been inconsistent. Since the tongue-pull hypothesis assumes that the high tongue position causes laryngeal raising, it predicts that a high laryngeal position correlates with high F0. Generally, laryngeal height has been found to correlate with fundamental frequency (Hong, Hong, Jun & Hwang, 2015); however, this is not always the case. Sapir's (1978) study of laryngeal height during vowel production found that English speakers tended to produce high-vowel /u/ with a lower laryngeal height compared to low-vowel /a/, even though /u/had a higher F0 than /a/ (Sapir, 1989). Based on Sapir's (1989) findings regarding laryngeal height during vowel production, we would expect /a/ to have a higher F0 than /u/; however, data from Whalen and Levitt (1995) and Sapir (1989) suggest the opposite pattern holds: /u/ has a higher F0 than /a/ despite having a lower laryngeal position. It is also important to note that an increase in F0 is not caused exclusively by larynx raising. Volitional F0 raising is primarily caused by activation of the cricothyroid, the muscle located between thyroid and cricoid cartilage (refer to Figure 2 for identification of thyroid and cricoid cartilage). Contraction of this muscle causes the thyroid cartilage to tilt forward, lengthening the vocal folds and increasing F0. F0 can also raise due to an increase in vocal fold stiffness and increased airflow (Zhang, 2016). Therefore, it is not surprising that laryngeal height alone does not explain IF0 despite finding correlation between laryngeal height and fundamental frequency in other tasks (e.g., pitch glides, singing).

Ohala (1978) provides another version of the tongue-pull hypothesis. Rather than relating tongue position to laryngeal height, he proposes that the high tongue position of high vowels leads to a pulling of the aryepiglottic folds, which then pull the ventricular ("false") vocal folds away from the true vocal folds; this in turn increases vertical tension of the true vocal folds. The vertical movement of the false vocal folds reduces the damping effect of the false vocal folds on the true vocal folds; together with the increase of vertical tension of the true vocal folds, this leads to the increased F0 in high vowels (Ohala, 1978; Ohala & Eukel, 1987). This is schematized in Figure 2.

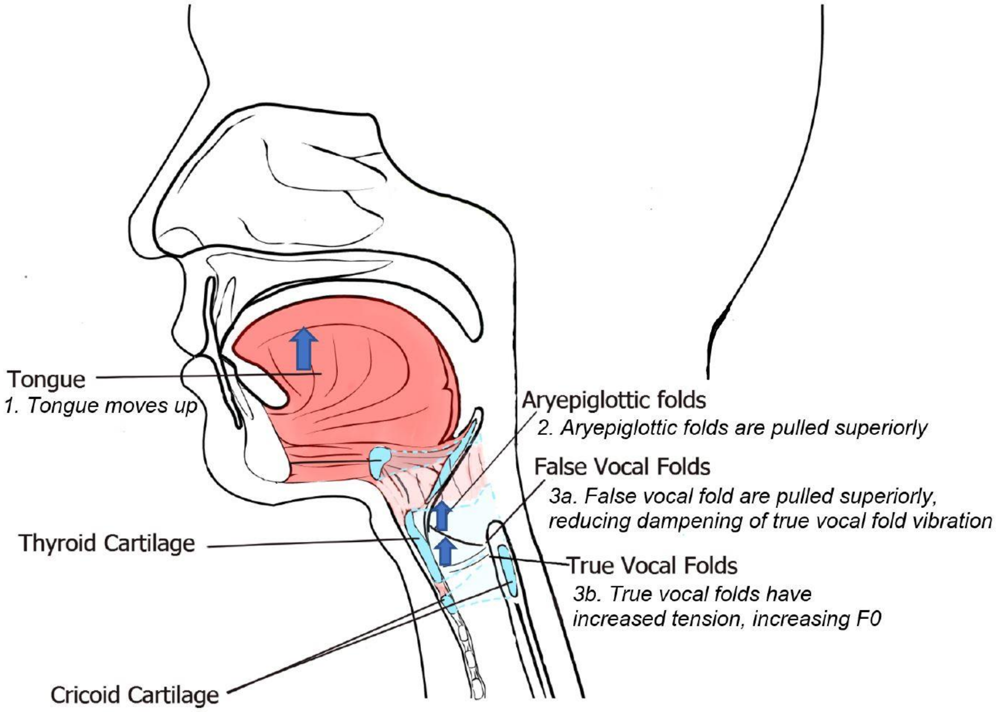


Figure 2. Schematic of revised tongue-pull hypothesis with arrows indicating direction of articulator movement. Tongue moves upward (1), pulling on the aryepiglottic folds (2), which causes the false vocal folds to move away from the true vocal folds (3a), reducing dampening of true vocal fold vibration and (3b) increasing vertical tension of the true vocal folds (Color online). Image source: OpenStax AnatPhys fig.22.6 - Divisions of the Pharynx - English labels" at AnatomyTOOL.org by OpenStax, license: Creative Commons Attribution. Source: book 'Anatomy and Physiology', https://openstax.org/details/books/anatomy-and-physiology. Adapted by B. Ramos for this paper.

To test the tongue-pull hypothesis, Ohala and Eukel (1987) had participants produce 6 English vowels with bite blocks of varying widths; under these conditions, it was hypothesized that the size of the bite block would be correlated with the magnitude of the F0 effect, reflecting the increased tongue-pull for bite blocks of greater width. They found that found high vowels had a higher pitch than low and mid vowels in all conditions, with low and mid vowels pairing together. The magnitude of the effect increased only with the largest bite block (10 mm). These findings were interpreted as support for the revised tongue-pull hypothesis.

This hypothesis has several underlying assumptions, some of which remain untested. The first is that IF0 is due to passive laryngeal movement that is an automatic consequence of tongue movement for vowel articulation. Evidence from Ohala and Eukel's (1987) bite block study as well as findings in infant speech (Whalen, Levitt, Hsiao & Smorodinsky, 1995) and speech of Deaf adults (Bush, 1981) support this assumption. Additionally, Whalen et al. (1998) completed an EMG study of cricothyroid activity during vowel production. If IF0 were volitional (i.e., not passive), they would expect to see increased cricothyroid activity for high vowels (high F0) compared to low vowels (low F0). Instead, the authors found that cricothyroid activation did not follow the F0 pattern found in vowels. They interpreted their results as support of the hypothesis that IF0 is a result of passive laryngeal movement and an automatic consequence of vowel articulation since there was no evidence of volitional muscle contraction leading to changes in F0 (Whalen, Gick, Kumada & Honda,1998). Second, Ohala's tongue-pull hypothesis assumes that the false vocal folds dampen true vocal fold vibration in a neutral (i.e., mid) tongue position and that vocal fold dampening lowers F0; therefore, the tongue pulling the false vocal folds vertically allows for increased F0 by reducing the dampening effect. There is some research to bear on these assumptions: Bailly et al. (2014) found that medial compression of the false vocal folds (without false fold vibration) leads to reduced fundamental frequency (Bailly, Bernardoni, Muller, Rohlfs & Hess, 2014). This suggests that the false vocal folds can, indeed, dampen true vocal fold vibration, as predicted by Ohala. However, Bailly's (2014) results pertained to medial compression (adduction) of the false vocal folds, not vertical distance between the true vocal folds and false vocal folds. Still, it stands to reason that if medial compression of the false vocal folds leads to reduced F0, then reduction in ventricular space would be expected to impede true vocal fold vibration as well. Whether vertical tongue movement leads to increased ventricular space is an open question. Finally, the hypothesis also assumes that vertical tongue motion leads to increased vertical tension of the true vocal folds, though it is unclear what is meant by *vertical* vocal fold tension as opposed to longitudinal vocal fold tension (e.g., via

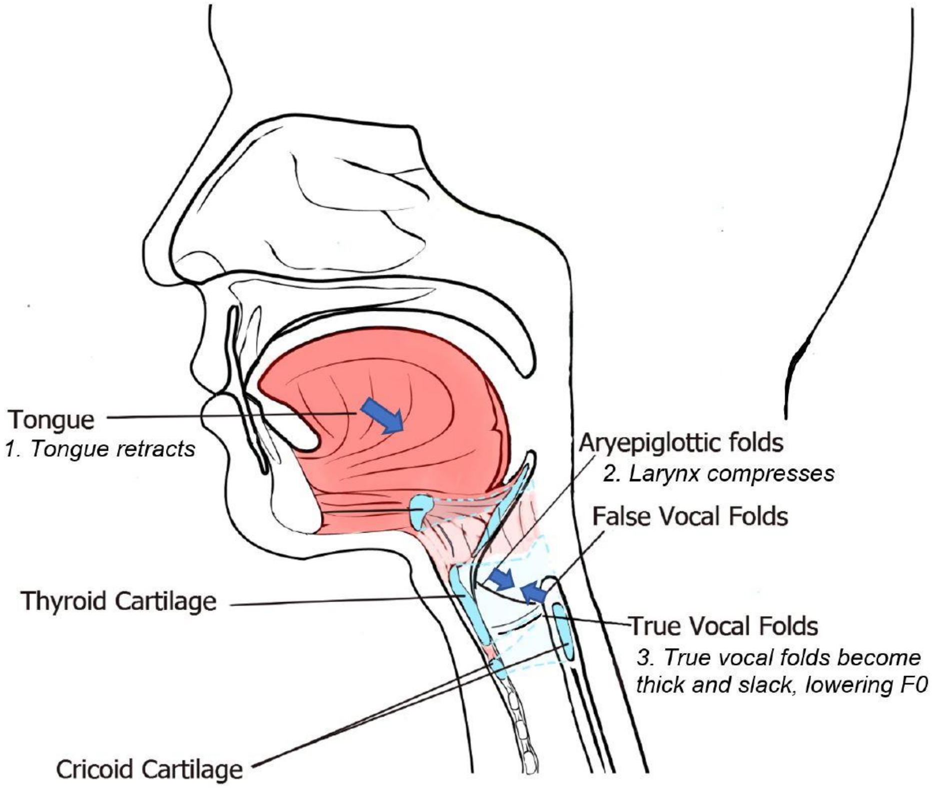


Figure 3. Schematic of tongue-compression hypothesis. Tongue retraction (1) causes laryngeal compression in anterior-posterior dimension (2). This causes vocal folds to become thick and slack (3), reducing F0 (Color online). Image source: OpenStax AnatPhys fig.22.6 - Divisions of the Pharynx - English labels" at AnatomyTOOL.org by OpenStax, license: Creative Commons Attribution. Source: book 'Anatomy and Physiology, https://openstax.org/details/books/anatomy-and-physiology. Adapted by B. Ramos for this paper.

thyroarytenoid contraction), and the assumption that vertical fold tension is increased with tongue raising remains to be validated.

In contrast to the tongue-pull hypothesis, the TONGUE-COMPRESSION HYPOTHESIS proposes that the tongue lowering and retraction during the production of low vowels leads to anterior-posterior pharyngeal compression, which causes the vocal folds to thicken and slacken, leading to the lowered F0 of low vowels (Ewan, 1979). This is schematized in Figure 3. Like the tongue-pull hypothesis, the tongue-compression hypothesis assumes that IF0 is an automatic consequence of vowel articulation. In his study, Ewan (1979) demonstrated that vowels following Arabic pharyngeal approximants had lowered F0 compared to other Arabic consonants; he proposes that this is due to the compression of the tongue for the pharyngeal consonants causing the F0 perturbation on the following vowel. However, studies of pharyngeal consonants suggest that the primary articulatory motion is epiglottal retraction, not tongue lowering (Laufer & Condax, 1979; Esling, 1996). Laufer & Condax (1979) also found that the vowel /a/ in Hebrew is produced with epiglottal retraction as the primary articulation, similar to pharyngeal consonants. Nevertheless, the predictions of the tongue-compression hypothesis remain unchanged whether the epiglottis retracts as the primary articulation for low vowels or as a consequence of tongue retraction, because both scenarios are hypothesized to have the same effect of laryngeal compression. Under the tongue-compression hypothesis, high and mid vowels would be expected to pair together given the lack of tongue compression during production for those vowels, and /a/ would have the lowest fundamental frequency.

The tongue-compression hypothesis remains largely untested in the literature, though results from some studies can be used to support the claim that tongue compression leads to lowering of F0. In an MRI study of laryngeal configuration during vowel production, Moisik et al. (2019) found that tongue retraction led to laryngeal compaction and that low vowels are more likely to have creaky voice quality as a result of this unless compensatory adjustment is made (Moisik, Esling, Crevier-Buchman & Halimi, 2019). Similarly, Lancia and Grawunder (2014) found that /a/ was more prone to laryngealization compared to a more fronted vowel, such as /i/, due to tongue retraction required for /a/ (Lancia & Grawunder, 2014). This is consistent with the model of the vowel space provided by Esling (2005), who claims that low-back vowels are produced by the *laryngeal articulator*; that is, vowels such as /a/ can be considered pharyngeal vowels given the relationship between tongue retraction and the larynx. In contrast, other vowels do not involve the laryngeal articulator as a primary articulator. Esling adds that vowels involving tongue retraction are more susceptible to increased laryngeal constriction. With this view, the findings presented in Esling 2005, Lancia and Grawunder 2014, and Moisik et al. 2019 partly support Ewan's tongue-compression hypothesis that a retracted tongue position leads to low F0, under the assumption that increased laryngeal constriction or creakiness for low back vowels is correlated with decreased F0. The support for the tongue-compression hypothesis is only partial, however. Recall that Ewan's tongue-compression hypothesis also stated that tongue retraction leads to vocal fold slackening. Yet findings from the studies cited here suggest that tongue retraction causes decreased F0 by *increasing* laryngeal constriction, not by vocal fold *slackening*, as Ewan predicts.

In addition to movement of the tongue, there is some evidence that jaw height also influences IF0, though the results can vary by speaker (Zawadzki & Gilbert, 1989; Pape & Mooshammer, 2006; Chen et al., 2019). Zawadzki and Gilbert (1989) found that jaw height was a better predictor of vowel F0 than tongue height for 3 of 5 participants in their study of 11 American English vowels. Similarly, Pape and Mooshammer (2006) found both jaw height and tongue height to be important articulatory movements in predicting F0, though neither articulator explained F0 differences between German tense and lax vowels in their study. Recently, Chen et al. (2019) completed a study of IF0 in 8 English vowels using x-ray microbeam data of 40 speakers from two databases. While tongue height and F1 show high correlations with F0, their findings suggest that jaw height contributes more to IF0 than tongue height. The authors state that the muscular chain between the mandible, hyoid, and larynx may be behind the connection between jaw height and changes in F0, though they do not offer a predicted sequence of actions responsible for IF0. Erickson et al. (2017) propose that jaw opening leads to backward translation of the hyoid bone. The hyoid bone is connected to the thyroid cartilage through thyrohyoid muscles; when the hyoid bone is moved back, they propose that the thyroid cartilage is also forced to rotate posteriorly

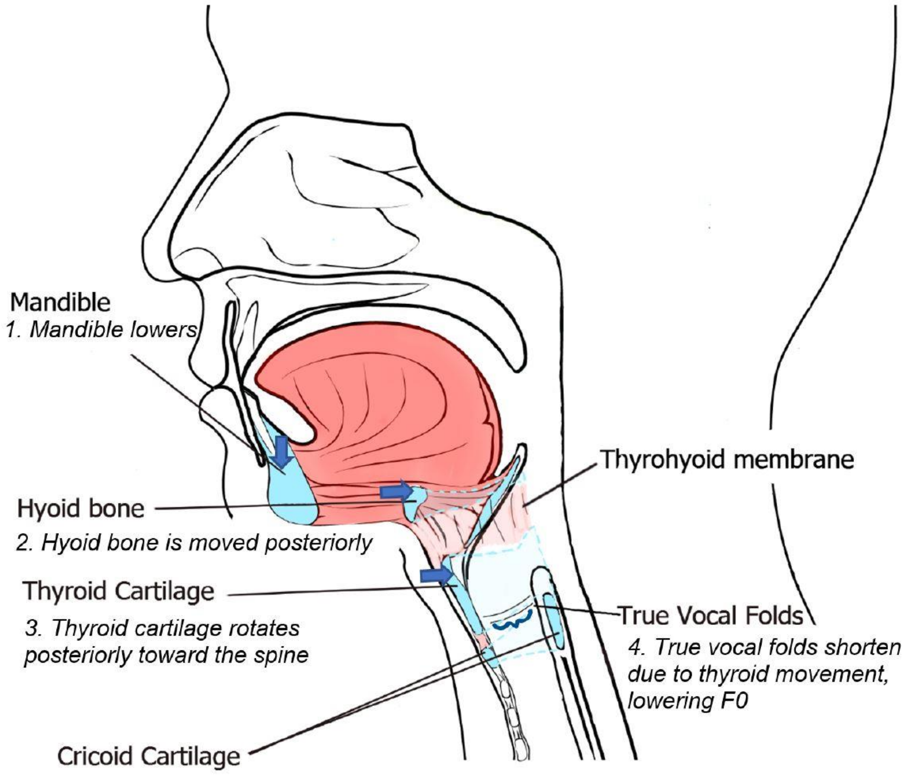


Figure 4. Schematic of jaw height hypothesis. As jaw lowers (1), hyoid is pushed posteriorly (2). The thyroid is then rotated posteriorly (3) due to connection with the hyoid bone. Thyroid movement causes the vocal folds to shorten (4), reducing F0 (Color online). Image source: OpenStax AnatPhys fig.22.6 - Divisions of the Pharynx - English labels" at AnatomyTOOL.org by OpenStax, license: Creative Commons Attribution. Source: book 'Anatomy and Physiology', https://openstax.org/details/books/anatomy-and-physiology. Adapted by B. Ramos for this paper.

toward the spine and away from the cricoid cartilage, inducing vocal fold shortening, which lowers F0 (Erickson, Honda, and Kawahara, 2017) (Figure 4).

The hypothesized link between jaw movement and IF0 remains untested in the literature. Interestingly, in both Zawadzki & Gilbert (1989) and Chen et al. (2019), /a/ and /ɔ/ seem to have lower F0 values than expected given their jaw height alone when compared to vowels of similar jaw height. Though neither study compared F0 of front and back low vowels, the results would be compatible with the tongue compression hypothesis and suggest that IF0 has multiple sources rather than a link between the vocal folds and a single articulator, such as the jaw.

Unlike the previous hypotheses, Sapir (1989) proposes instead that the laryngeal configuration during vowel production is volitional and forms part of the articulatory goal for the vowel to reach its acoustic target. His HORIZONTAL-VERTICAL PULL HYPOTHESIS proposes that IF0 is due to extrinsic laryngeal muscle activation leading to changes in laryngeal configuration during vowel production (Sapir, 1989). According to his account, extrinsic laryngeal muscle activation, not tongue movement, causes changes in laryngeal and pharyngeal configurations in order to reach an acoustic target for vowel quality (i.e., F1 and F2 targets). The laryngeal configurations also have consequences for fundamental frequency. As summarized in Sapir (1989), during production of /i/, contraction of the suprahyoid muscles cause the hyoid to move forward and superiorly, which leads to increased pharyngeal space in the anterior-posterior dimension and lowers F1 for /i/. This same movement causes increased true vocal fold tension, which increases F0. Since low vowels do not target a low F1, there is less activation in the suprahyoid muscles, and the vocal fold tension is not altered, leading to a lower F0 for low vowels compared to high vowels. This hypothesis remains largely untested outside of Sapir's own studies (Sapir 1989), though evidence from Hong et al. (2015) suggests that extrinsic laryngeal muscles are involved in pitch control. The results of Hong et al. (2015) are based on their participants' productions of /a/ at continuously increasing pitch from minimum to maximum. It is unclear if the same extrinsic laryngeal muscle activation found under these conditions apply to the patterns found in speech, where F0 varies within a smaller range. A summary of articulatory accounts of IF0 are presented in Table I.

TABLE I. Summary of articulatory explanations of IF0

Hypothesis	Mechanism
Tongue-pull (Ohala)	High tongue → False folds move away from true folds + increased vertical tension of true vocal folds → Increased F0
Tongue-compression (Ewan)	Retracted tongue → Laryngeal compression → Thick and slack vocal folds → Reduced F0
Jaw height (Erickson et al.)	Jaw lowers → Hyoid moves posteriorly → Thyroid rotates posteriorly → Vocal folds shorten $→ Reduced F0
Horizontal-vertical pull (Sapir)	Phoneme-specific extrinsic laryngeal muscle activation leads to changes in F0

Non-articulatory explanations of IF0 have also been proposed, including the acoustic coupling hypothesis (Flanagan & Landgraf, 1968) and the auditory enhancement hypothesis (Diehl, 1991). The auditory enhancement hypothesis proposes that phonetic features of consonants and vowels covary to enhance the auditory feature of that segment. With respect to IF0, the hypothesis suggests that changes in IF0 are due to purposeful enhancement of vowels to make contrasts perceptually distinct (Diehl, 1991). Whalen and Levitt (1995) point out that the near universality of IF0 has been used as an argument against the hypothesis (Whalen & Levitt, 1995). If the difference in F0 is driven by linguistic factors and is not merely an automatic consequence of articulation, we would predict that some languages do not show the IF0 effect. Moreover, the effect has been demonstrated in infant babbling prior to the onset of meaningful speech production (Whalen et al.,

1995) and in the speech of Deaf adults (Bush, 1981), suggesting that the F0 of vowels is not due to auditory enhancement of the vowel contrasts. However, it is important to note that evidence against the hypothesis as the *cause* of the cross-linguistic phenomenon does not rule out the possibility that IF0 can be exploited for language-specific enhancement. In fact, Van Hoof et al. (2011) found that Dutch, with a 12-vowel system, had a greater IF0 effect than Moroccan Arabic, with only 3 vowels. The authors argue that languages with a more crowded vowel space might exploit IF0 to enhance the differences between high and low vowel contrasts (Van Hoof et al., 2011). Additionally, studies have found that the effect size of IF0 varies by sociolinguistic factors such as regional variety (Jacewicz & Fox, 2015) or ethnolect (Holt, Jacewicz, & Fox, 2019). Finally, IF0 has also been found in esophageal speech, where no laryngeal structures are present (Gandour & Weinberg, 1980), which could be taken as evidence in support of auditory enhancement. In sum, evidence suggests that auditory enhancement may be a factor in the differences in magnitude of the IF0 effect between languages and language varieties. If IF0 were entirely dependent on tongue position or jaw height, then we would expect similar effect sizes cross-linguistically. The fact that we see such variations suggests that volitional control may play a role in IF0 either as auditory enhancement or as part of the articulatory goal as proposed by the horizontal-vertical pull hypothesis.

Finally, the acoustic coupling hypothesis proposes that the increased F0 of high vowels is an effect of the first formant on the fundamental frequency (Flanagan & Landgraf, 1968). The hypothesis states that when the vocal folds vibrate at a rate that is near the resonant frequency of F1, acoustic coupling between the vocal tract and vocal folds occurs, such that F0 increases due to coupling with F1. Since high vowels have a low F1 that can be near to F0, acoustic coupling occurs for high vowels, causing an increase in F0 for high vowels. Ewan (1979) tested the acoustic coupling hypothesis, comparing nasal /m/ in two phonetic environments, /ama/ and /umu/. Despite the first nasal formant being low in both environments, the F0 of /m/ differed depending on the vowel context. The nasal in /ama/ had lower F0 than the nasal in /umu/. In both cases, the F0 of the intervocalic /m/ was consistent with the vowel context and did not appear to be influenced by the first nasal formant (Ewan, 1979). Guérin and Boë (1980) also found evidence that acoustic coupling cannot account for IF0. Using a two-mass voice source model to simulate the effect of vocal tract changes on F0, the authors found that F0 was positively correlated with F1. This is the opposite pattern found in natural speech as it predicts increased F0 with low vowels, not high vowels (Guérin & Boë, 1980). Due to the evidence suggesting acoustic coupling does not account for IF0, this hypothesis has generally been abandoned in the literature as a possible explanation.

b. Current Study

The cause of IF0 remains unknown, though the summary of previous research presented above suggests IF0 may have multiple sources. The purpose of the present study is to investigate the predictions of articulatory hypotheses by comparing high, mid, and low vowels in Amharic using acoustic data and electroglottography (EGG) to examine the relationship between tongue position, F0, and Contact Quotient (CQ), a measure of vocal fold contact area during vibration that is thought to relate to vocal fold tension (Herbst, 2020). Amharic is a Semitic language spoken in Ethiopia and by a large diaspora community. Its seven-vowel inventory allows for a detailed comparison of IF0 along the dimensions of both height and frontness as it contrasts front, central,

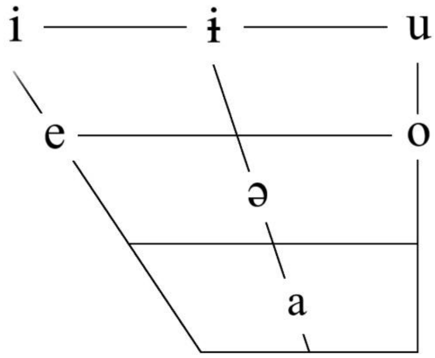


Figure 5. Amharic vowel inventory

and back vowels among high and mid vowels in addition to a single low vowel, /a/ (Figure 5) (Hayward & Hayward, 1992). Although previous research by Ado (2011) on Amharic vowels did not find a statistically significant difference between F0 of high vowel /i/ and low vowel /a/, the difference in F0 values for /i/ and /a/ was consistent with effect sizes found in Whalen and Levitt (1995). Previous studies examining the cause of IF0 have relied heavily on data from European languages (e.g., English or German). This study contributes acoustic and articulatory data from an under-represented language to address a long-standing question in phonetics.

The predictions of each hypothesis are as follows (Table II): under the tongue-pull hypothesis, high vowels in Amharic should have the highest F0, followed by mid and low vowels, which should pair together. Conversely, under the tongue-compression hypothesis, high and mid vowels would be expected to pair together given the lack of tongue retraction required for those vowels. However, since we know from the recent body of literature on IF0 in vowels that IF0 is gradient, it is likely that both tongue pull and tongue compression would be responsible for IF0. In this case, high, mid, and low vowels should all differ in F0. Additionally, since both hypotheses ultimately relate the cause of IF0 to differences in vocal fold tension, differences in contact quotient would be expected to follow the predicted F0 patterns. Since testing the horizontal-vertical pull hypothesis requires measuring muscle activation (e.g., with electromyography), this hypothesis is not tested directly in this study. Similarly, the current study does not directly assess the contribution of jaw height to F0, but predictions based on these hypotheses are provided in Table II.

Three of the articulatory hypotheses (tongue-pull, tongue-compress, and jaw height) assume that IF0 is an automatic consequence of articulator movement. That is, as the tongue or jaw moves, it induces changes on the larynx leading to F0 perturbations. It follows, then, that these changes should occur regardless of vowel category. For example, under the tongue-pull hypothesis, an /i/ that is produced with a higher tongue position should have a higher F0 and more tense vocal folds than an /i/ produced with a lower tongue position. Therefore, predictions are also provided for F0 and CQ based on F1 and F2, independent of vowel category in Table III.

TABLE II. Summary of F0 and CQ predictions for vowel categories based on articulatory explanations of IF0.

Hypothesis	Vowel category predictions for F0 and CQ
Tongue-pull (Ohala)	/i, ɨ, u/ > /e, ə, o, a/
Tongue-compression (Ewan)	/i, ɨ, u, e, ə, o/ > /a/
Jaw height (Erickson et al.)	/i, ɨ, u/ > /e, ə, o/ > /a/; no CQ predictions
Horizontal-vertical pull (Sapir)	Phoneme-specific extrinsic laryngeal muscle activation; F0 and CQ may vary by vowel category

TABLE III. Summary of F0 and CQ predictions as a factor of F1/F2 based on articulatory explanations of IF0.

Hypothesis	F1/F2 category predictions for F0 and CQ
Tongue-pull (Ohala)	F0: As F1 increases, F0 decreases; no F2 effect CQ: Increased for low F1 values (increased tension)
Tongue-compression (Ewan)	F0: Decreases with high F1 and low F2 CQ: Decreases under same condition (reduced tension)
Jaw height (Erickson et al.)	F0: As F1 increases, F0 decreases; no F2 effect CQ: No predictions
Horizontal-vertical pull (Sapir)	N/A; F0 is dependent on vowel category

II. METHODS

a. Participants

Participants included 8 native speakers of Amharic (5 men and 3 women) living in the San Diego area but who were originally from Addis Ababa, Ethiopia (Table IV). All participants were L2 English speakers and speak Amharic as the primary language in their home. Recordings from Speaker 1 and Speaker 7 were excluded due to noisy EGG signals that resulted in inability to calculate contact quotient. Individuals with a history of speech, language, hearing, or neurological disorders were excluded from participation, as were those who were not literate in Amharic, because the study protocol required reading the Amharic orthography.

TABLE IV. Participant demographics. Years in the US for Speaker 1 is unknown.

Speaker	Sex	Age	Years in US	Other Languages
1	F	45	NA	English, Arabic
2	M	44	17	English
3	M	36	7	English
4	M	36	13	English
5	F	19	5	English
6	M	19	5	English
7	F	19	13	English
8	M	19	10	English

c. Stimuli

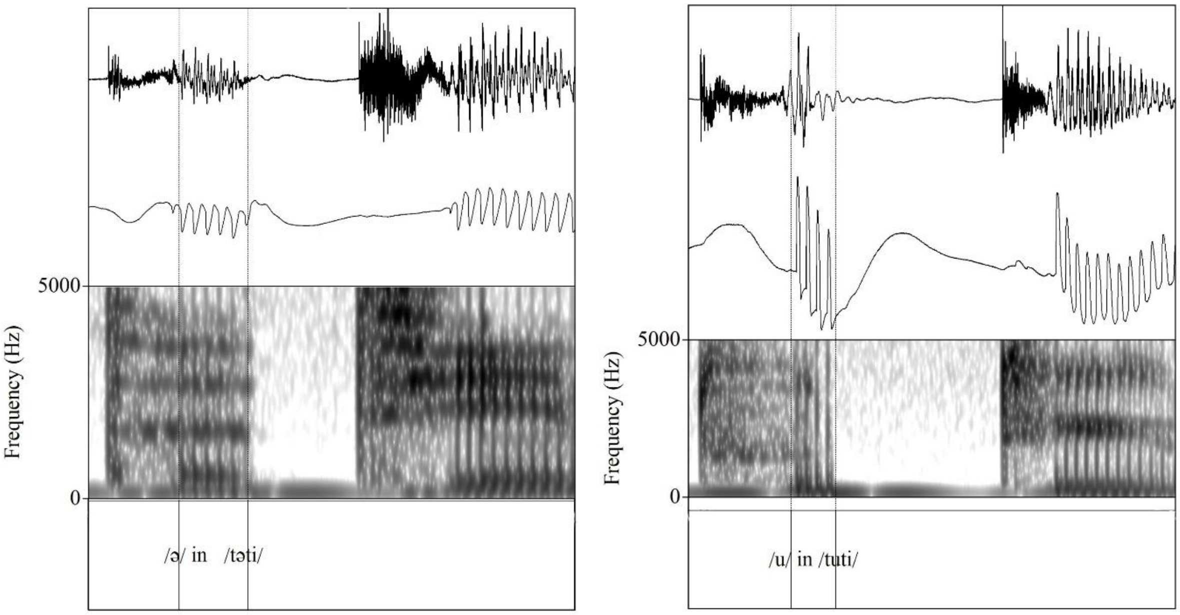
Stimuli consisted of /tVtV/ frames, where the first vowel was the target Amharic phoneme /i, ɨ, u, e, ə, o, a/ and the second vowel alternated between six Amharic vowels /i, u, e, ə, o, a/. The sixth order vowel, /ɨ/ was not included as a second vowel in the /tVtV/ frame as it cannot occur word-finally. All permutations resulted in nonce words, which were reviewed by a native Amharic speaker prior to use in the study to ensure that none of the words were in fact real words. Nonce words were chosen for maximal control of the phonetic environment of the target vowel because a minimal set of real words with all vowels in the same CVCV context could not be found. This was important because stress and consonant voicing are known to cause F0 perturbations on the adjacent vowel (Kirby & Ladd, 2016, 2015; Gordon & Roettger, 2017). Since the effect size of IF0 can be as small as 5 Hz (Whalen & Levitt, 1995), F0 perturbations due to differences in the phonetic environment could result in amplification or masking of an effect. Stimuli were embedded in a carrier phrase, /jɨh k'al _____ nəw / (*This word is _____)* to control for phrasal effects on F0. The complete list of stimulus words can be found in the Appendix.

d. Task

Stimulus sentences were presented one at a time on a screen using Amharic orthography, which is an alphasyllabary, where one character represents a CV syllable. Participants were instructed to read each sentence aloud at a comfortable rate and loudness and to self-advance to the next sentence at their own pace. Target words within the carrier sentences were repeated five times for each participant, with sentences presented in the same order each time. A sixth recording was collected for Speaker 6 due to frequent reading disfluencies during the first recording. There were 30 repetitions of each target vowel, yielding 210 stimulus items per participant. The author was present in the recording booth at the time of the recording to ensure participants read all stimulus items. If a word was read disfluently or in error, the speaker was asked to repeat it, and the second production was used for analysis. Reading was judged to be disfluent if the speaker produced repetition, prolongation, or blocking of segments. In the case of Speaker 6, who had a higher rate of disfluencies in reading compared to other participants, 88% of his recordings were segmented for analysis (compared to 100-98% for other participants), though the overall number of tokens analyzed was equal to other participants given the additional recording.

e. Instrumentation and segmentation

Recordings were done in a sound-attenuated booth. Audio data were recorded at a 44.1 kHz sampling rate and 16-bit quantization rate using a Shure SM10A head-mounted microphone. The electroglottography (EGG) signal was recorded using an EG2-PCX electroglottograph from Glottal



(a) (b)

Figure 6. Example of vowel segmentation. The top waveform in both images corresponds to the audio signal and the second waveform corresponds to EGG signal. 6a on the left shows an example of segmentation of vowel /ə/ in /təti/ with boundaries at the beginning and end of clear formants. 6b on the right shows segmentation of vowel /u/ in /tuti/ with boundaries beyond the beginning and end of clear formants to capture all glottal pulses in EGG signal.

Enterprises. EGG and audio signals were collected simultaneously using a Focusrite Scarlett 6i6 2nd generation pre-amplifier and digitizer. EGG measures vocal fold contact area during vocal fold vibration (Herbst, 2020). Contact quotient (CQ) is a measure of the duration vocal fold contact in relation to the total duration of the vibratory cycle. An increase in contact quotient should be interpreted as increased vocal fold adduction (e.g., through general laryngeal and pharyngeal constriction) or increased contraction of the thyroarytenoid--the muscle responsible for increased true vocal fold tension (Herbst, 2020). Therefore, CQ can be used to examine differences in tension proposed by the tongue-pull and tongue-compression hypotheses.

Audio and EGG data were segmented in Praat (Boersma & Weenink, 2020). Target vowels were segmented from the onset of clear formants to the end of clear formants (Figure 6(a)). In cases of short vowels consisting of three or fewer glottal pulses in the EGG waveform, the boundary for vowel segments were extended just beyond the beginning and end of glottal pulses regardless of formant onset and offset in the spectrogram. This was done in order to capture all EGG pulses for analysis (Figure 6(b)).

Independent variables include phonemic vowel category, F1, and F2. Dependent variables include F0 and contact quotient. Mean F0, F1, and F2 were collected automatically every millisecond over the duration of the segmented vowel using the STRAIGHT algorithm for F0 (Kawahara, de Cheveigné & Patterson, 1998) and the Snack algorithm (Sjölander, 2004) for F1/F2 in VoiceSauce (Shue, Keating, Vicenik & Yu, 2011). F1 and F2 were corrected for vowels that were judged to be mistracked based on values that fell outside of the normal range for the vowel category (e.g., an of F2 of 2000+ Hz for /u/). To correct mistracked values, F1 and F2 were first re-collected automatically in VoiceSauce using the Praat settings. If those values were also judged to be mistracked by the same criteria, values were corrected manually in Praat. EGG waveforms were analyzed using EGGworks. Mean CQ was collected over the duration of the segmented vowel using the hybrid method (Howard, 1995). The hybrid method uses two different methods for defining vocal fold contacting and decontacting. Onset of vocal fold closure is taken to be the peak in the derivative of the EGG signal. The end of contact was determined using a 25% threshold. Contact quotient values below 0.2 and above 0.8 were excluded, as these values are believed to be outside of the range of normal voicing. Mean values were calculated for all measures from the middle one third of each vowel for analysis. This was done to further control for co-articulatory effects of the adjacent consonants on F0, F1, or F2 of the vowel. F1, F2, F0, and CQ values were z-scored for each speaker, and values greater than 2.5 standard deviations away from each speaker's mean were excluded as these represented outliers.

f. Analysis

The data were analyzed in R (version 3.6.2, R Core Team, 2019) using raw values from the middle one-third of each vowel. Raw data for each token were z-scored for each speaker and values falling greater than 2.5 standard deviations away from the mean were excluded as these were believed to represent outliers. Linear mixed effects analyses were performed using *lme4* (Bates, Machler, Bolker, & Walker, 2015) to compare the effect of vowel category on F0 and CQ. The contribution of duration was also assessed as a fixed effect to account for microprosody, as shorter vowels might be expected to exhibit F0 perturbations in the middle one third of the vowel, while longer vowels (such as /a/) would not. Nested model comparison was used to assess the contribution of each fixed effect to the model. Post-hoc pairwise comparison with Tukey alpha correction was performed to compare F0 between vowel categories using the *lsmeans* package (Length, 2016). Additionally, since tongue position is predicted to have an effect on F0 and CQ as a result of mechanical coupling, independent of vowel category, linear mixed effects analyses were done to investigate the relationship between F0, F1 and F2 (acoustic correlates of tongue height and backness, respectively) as well as between CQ, F1 and F2. F1 and F2 were treated fixed effects; their interaction was also included. Participant was entered as random intercepts in all models to account for speaker variability. Nested model comparison was used to examine the factors that predicted F0 and CQ.

III. RESULTS

a. Fundamental Frequency

TABLE V. Mean and standard deviation of F0, CQ, F1, and F2 for each vowel category.

Vowel	F0 Hz (sd)	CQ Hz (sd)	F1 Hz (sd)	F2 Hz (sd)	Duration ms (sd)
i	155.15 (43.66)	0.55 (0.08)	341.68 (59.08)	2246.27 (235.38)	90.04 (26.96)
ɨ	152.99 (40.56)	0.55 (0.09)	432.12 (44.86)	1770.81 (192.99)	96.31 (39.05)
u	152.15 (45.17)	0.54 (0.09)	397.26 (115.23)	1322.86 (197.7)	88.7 (30.11)
e	146.48 (40.39)	0.54 (0.07)	437.5 (41.64)	2145.02 (252.16)	117.4 (31.46)
ə	144.49 (40.99)	0.54 (0.07)	598.18 (83.23)	1664.41 (187.7)	106.27 (34.6)
o	146.97 (40.97)	0.55 (0.07)	492.43 (95.94)	1110.4 (352.45)	115.08 (28.56)
a	141.03 (38.98)	0.55 (0.07)	764.33 (110.7)	1371.89 (147.56)	126.3 (26.85)

The means and standard deviations of fundamental frequency (F0), contact quotient (CQ), F1, F2, and duration for each vowel category are provided in Table V. F0 is highest for /i/ and lowest for /a/, consistent with findings of previous studies in IF0 (Whalen & Levitt, 1995; Ado, 2011). Overall, high vowels have the highest mean F0, followed by mid vowels, and finally /a/, though there are within-height differences between high and mid vowels. CQ results will be discussed in section 3.2.

A linear mixed effects analysis was performed to assess the effect of each factor (vowel, F1, F2) on F0. Significance of a targeted factor was assessed with nested model comparison between a model that included the targeted factor to a baseline one where that factor was removed from the model, but whose random error structure was identical. This was done using the *anova()* function in R (R Core Team, 2019) There was a significant main effect of *vowel quality* on F0 (X^2 (1) = 271.33, p < .001). See Table VI for coefficients and standard errors for the model of F0 as a function of *vowel category*. Post-hoc pairwise comparisons between vowels in the model with Tukey alpha correction found no difference between any two pairs of high vowels /i, ɨ, u/, but F0 values for all high vowels

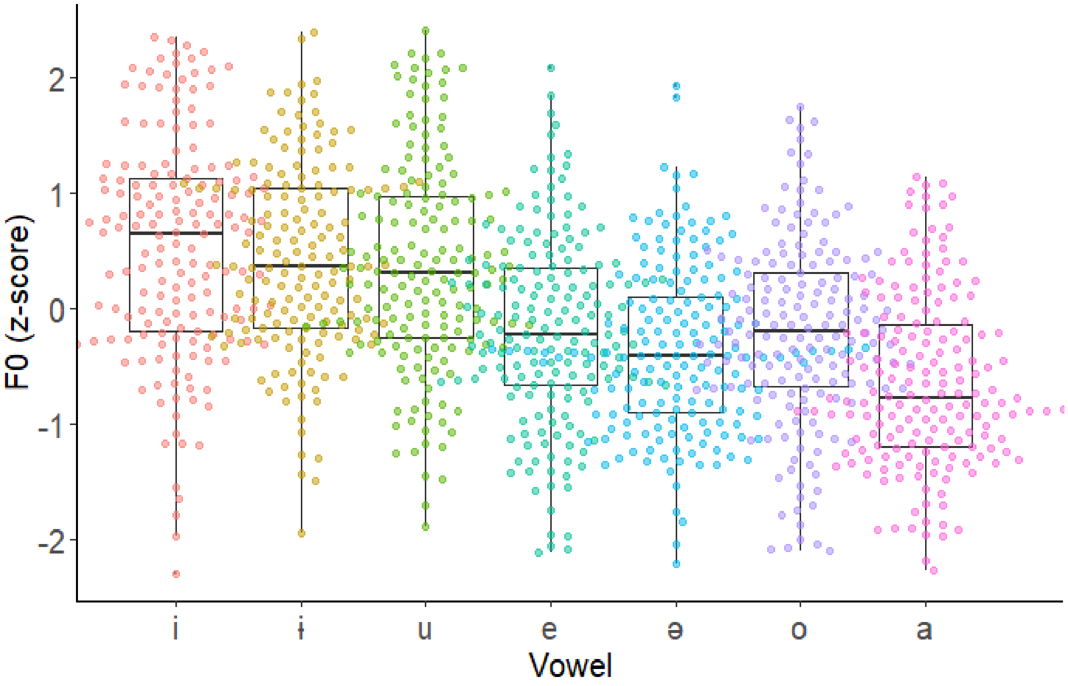


Figure 7. Beeswarm plots and overlaid boxplots of F0 (z-scored) as a function of vowel category (Color online).

were significantly higher than mid and low vowels. Mid vowels /e, ə, o/ were not significantly different from each other with respect to F0. All non-low vowels differed significantly from F0 of /a/, with /a/ having the lowest F0, as expected. These findings are illustrated in Figure 7, where high vowels have the highest F0s, followed by mid vowels, and finally low vowel /a/. Notably, there are no significant differences in F0 between front and back vowels at the same phonological height category.

TABLE VI. Coefficients, standard errors, t-values, and p-values of F0 as a function of vowel category with participant as a random intercept. P-values for this model were obtained using the *lmerTest* package (Kuznetsova et al., 2017).

Fixed Effect	Estimate	Standard Error	T-Value	P-Value
i (Intercept)	158.50	17.59	9.26	< .001
ɨ	-1.018	0.84	-1.21	0.23
u	-1.19	0.83	-1.43	0.15
e	-6.48	0.83	-7.82	< .001
ə	-8.07	0.83	-9.77	< .001
o	-6.22	0.83	-7.49	< .001
a	-10.99	0.82	-13.35	< .001

To ensure that differences across vowel categories were not due to inherent differences in duration of Amharic vowels (Fantaye, Yu, & Hailu, 2019), a model with fixed effects of duration and vowel was compared to the model with vowel as the only fixed effect. The main effect of *duration* was not found to be significant (X^2 (1) = 0.7568, p = 0.3843), which suggests that duration does not affect F0 once vowel quality is accounted for.

Next, we turn to the effect of F1 and F2 on F0. Recall, that the tongue-pull, tongue-compression, and jaw height hypotheses state that intrinsic F0 of vowels is caused by mechanical coupling of the tongue and larynx. Based on this, we would predict that F0 perturbations would differ even within vowel categories if tongue or jaw height differs between two vowels of the same category. Therefore, the effects of F1 and F2 on F0 were also examined independent of vowel category to investigate changes in F0 as a result of tongue position. The relationship between F1, F2, and F0 was examined with a linear mixed effects model including participant as a random intercept.

Table VII shows the summary of the full model with coefficients and standard errors of fixed effects. The interaction between F1 and F2 was not found to be significant (X^2 (1) = 0.7121, p = 0.3988). That is, changes in F1 do not contribute to the effect of F2 on F0. This is expected as we do not see an effect of vowel backness on F0 that differs depending on vowel height in Figure 8.

Results revealed a significant main effect of F1 (X^2 (1) = 192.31, $p < 0.001$), indicating that F0 decreases as F1 increases. This is seen in Figure 8 and Figure 9(a), which show that F0 decreases as speakers produce vowels lower in the vowel space. The main effect of F2 was not significant (X^2 (1) = 0.2034, p = 0.652), indicating that vowel frontness does not have an effect on F0. This is seen in Figure 9(b); as F2 increases (increased frontness), F0 remains largely unchanged.

TABLE VII. Coefficients, standard errors, and t-values, and of F0 as a function of F1 and F2 with participant as a random intercept. Nested model comparison revealed a significant main effect of F1; F2 and F1*F2 did not significantly improve the model.

Fixed Effect	Estimate	Standard Error	T-Value
Intercept	158.5	17.59	9.01
F1	-177	8755	-2.02
F2	1793	2492	0.72
F1*F2	<0.01	<0.01	-0.84

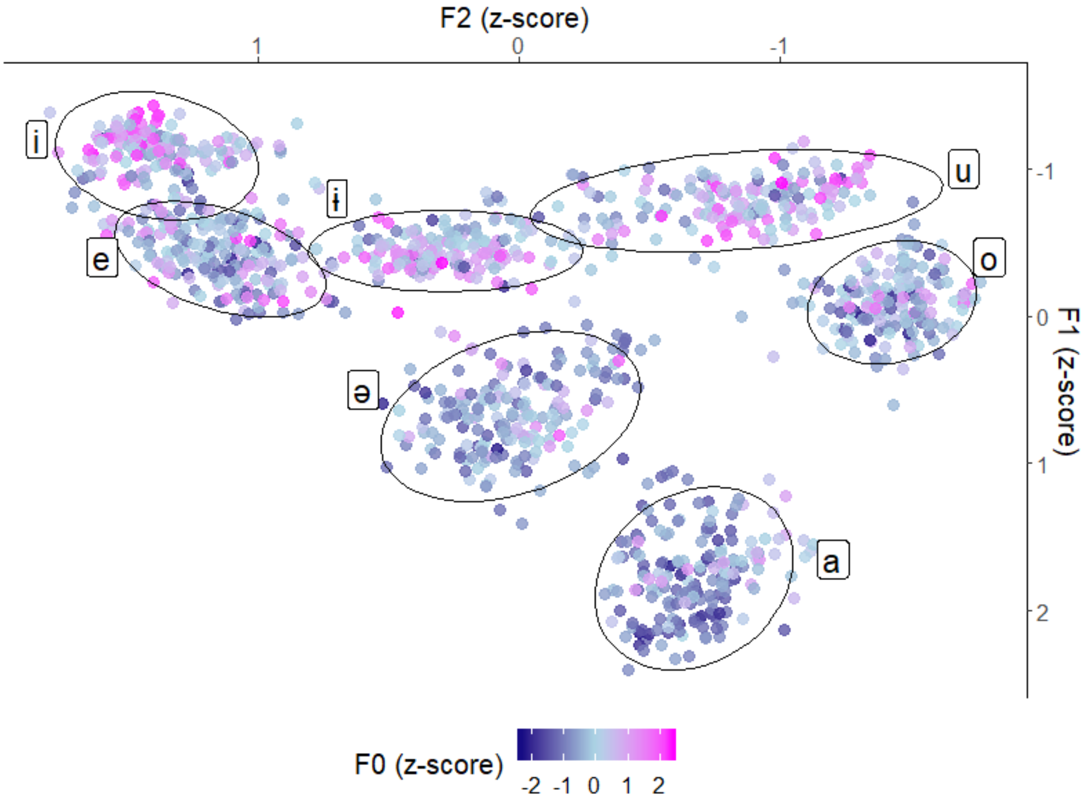
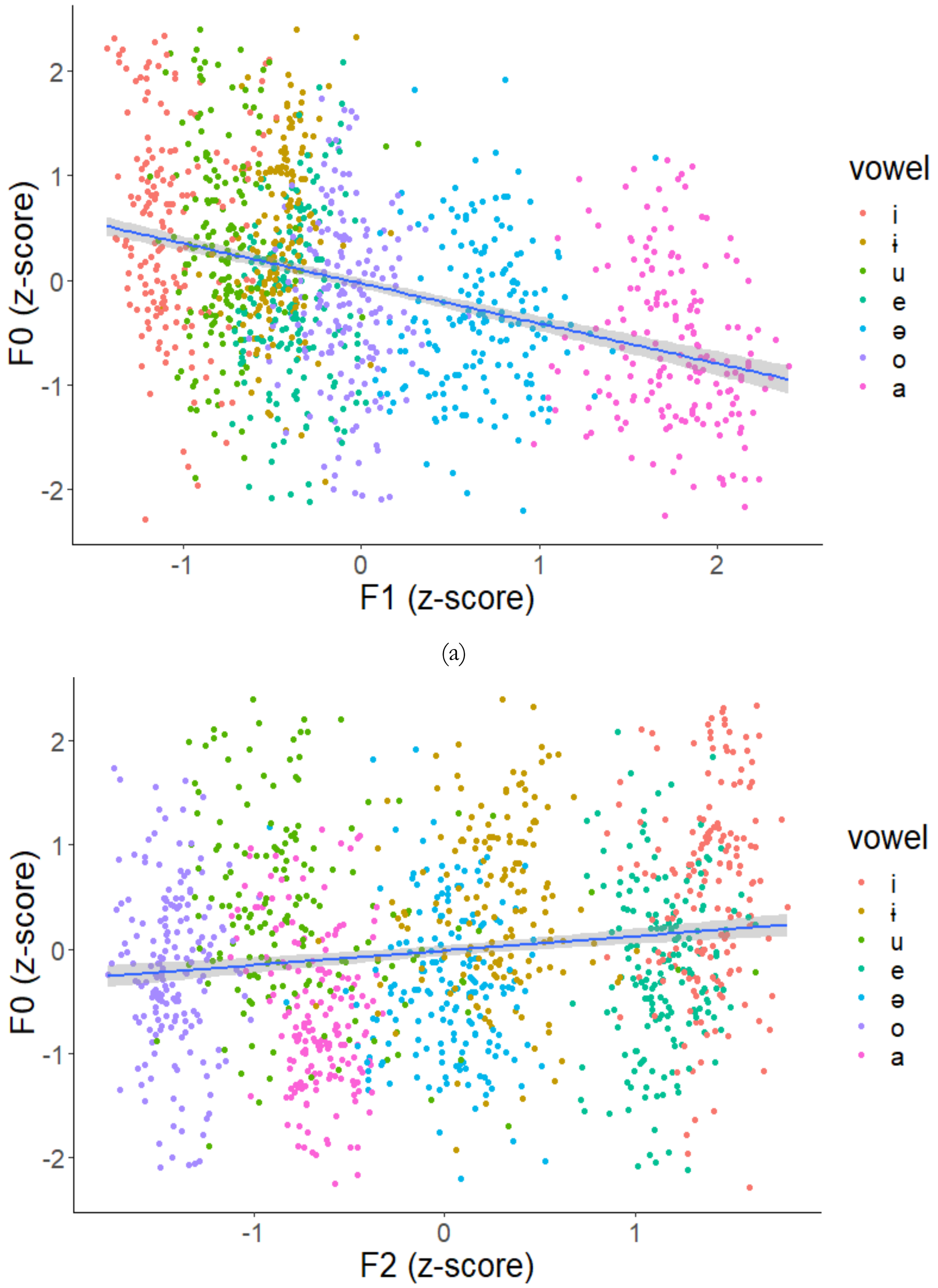


Figure 8. Scatterplot of z-scored F1~F2 with z-scored F0 represented as color gradient; magenta indicates higher F0 and blue indicates lower F0 (Color online).

b. Contact Quotient

CQ was used as measure of vocal fold tension. Recall that CQ is a measure of the relative duration vocal fold contact during a vibratory cycle, and an increase in contact quotient is typically interpreted as increased vocal fold adduction or contraction of the thyroarytenoid (Herbst, 2020) Mean CQ values for each vowel category can be seen in Table V. Based on the mean values from the middle one-third of each vowel, CQ does not appear to vary by vowel height or backness. The main effect of *vowel* was not significant, indicating that CQ does not vary as a function of vowel (X^2 (1) = 5.5879, p = 0.4709). These results are seen in Figure 10, which shows similar CQ values across vowel categories.



(a)

(b)

Figure 9. Scatterplots of F0 (z-scored) as a function of F1 (z-scored) in Figure 9a and F0~F2 (z-scored) in Figure 9b with vowel represented in color (Color online).

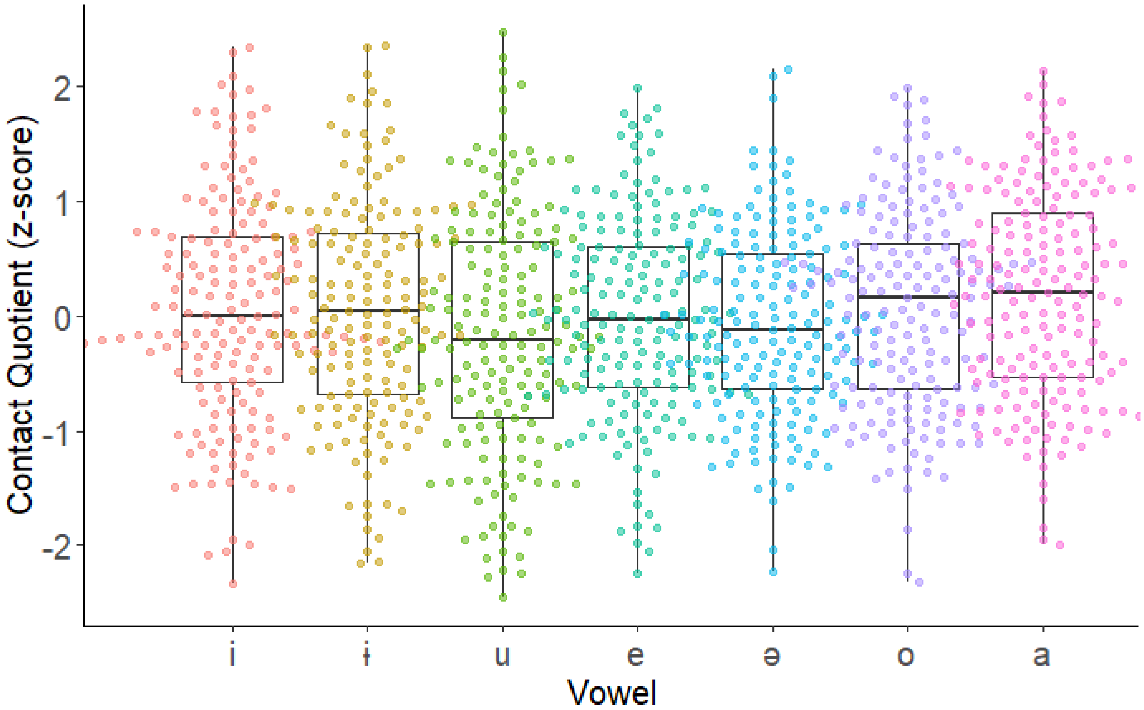
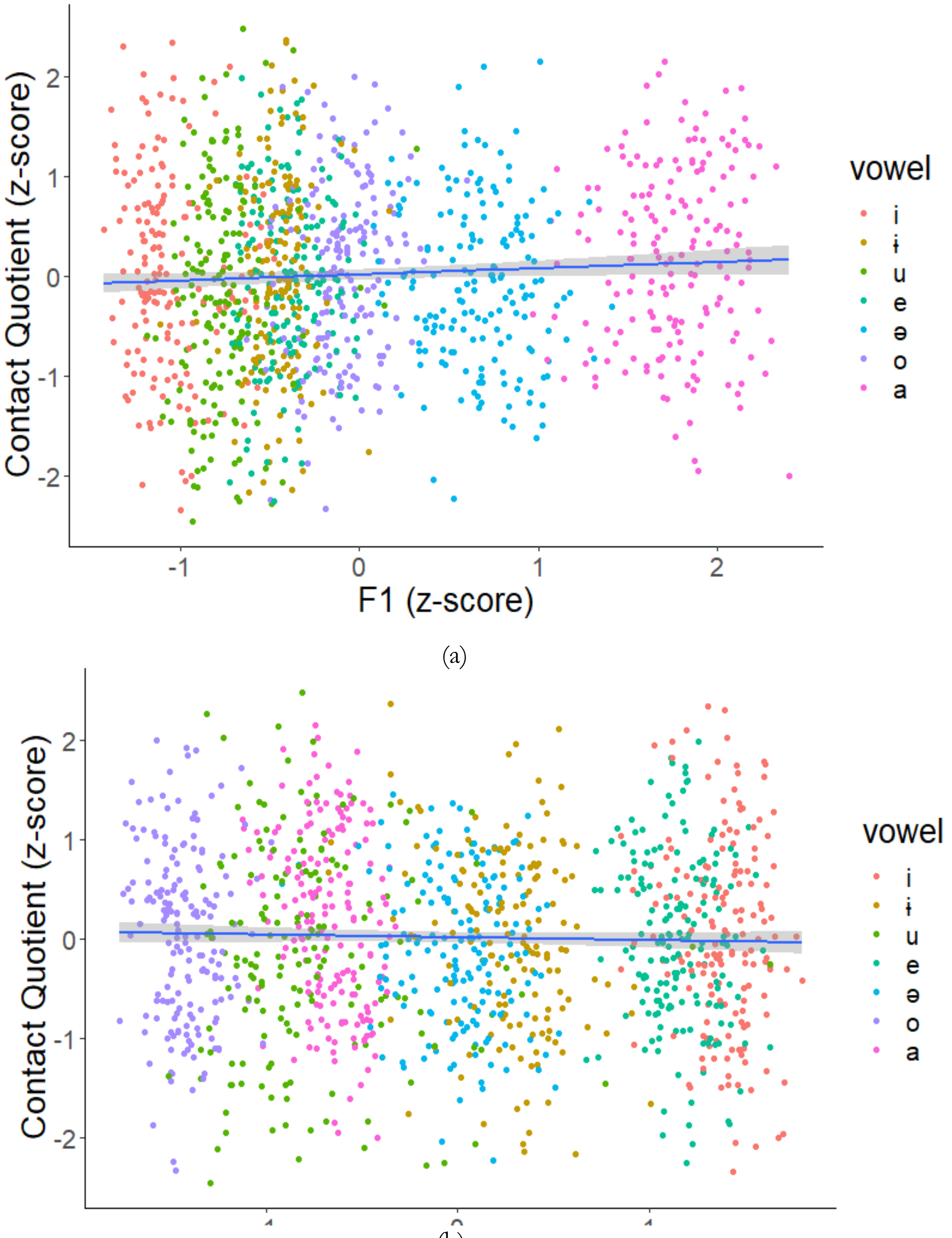


Figure 10. Beeswarm plots and overlaid boxplots of CQ (z-scored) as a function of vowel category (Color online).

As previously stated, the tongue-pull, tongue-compression, and jaw height hypotheses state that intrinsic F0 of vowels is caused by mechanical coupling of the tongue and larynx. Under the tongue-pull and tongue-compression hypotheses, tongue movement induces changes in vocal fold tension, which in turn lead to F0 perturbations. Based on this, we would predict that vocal fold tension would differ even within vowel categories for the tongue-pull or tongue-compression hypothesis. The effects of F1 and F2 on CQ were examined independent of vowel category to investigate changes in CQ as a result of tongue movement. The relationship between F1, F2, and CQ was examined with a linear mixed effects model including participant as a random intercept and F1 and F2 as fixed effects. Table VIII shows the summary of the full model with coefficients and standard errors of fixed effects. Nested model comparison revealed the interaction between F1 and F2 to be significant (X^2 (1) = 9.9346, $p < 0.05$); however, the significance was due to improved model fit



(a)

(b)

Figure 11. Scatterplots of CQ (z-scored) as a function of F1 (z-scored) in Figure 11a and as a function of F2 (z-scored) in 11b with vowel represented in color (Color online).

without the interaction term compared to with. The effect of F2 was found to be not significant (X^2 (1) = 0.8707, p = 0.3508). Similarly, was the effect of F1 found to be not significant (X^2 (1) =

2.5164, p = 0.1127). That is, tongue height and backness, as measured by acoustic correlates F1 and F2, do not independently contribute to changes in vocal fold tension, as measured by CQ. This can be seen in Figure 11, where CQ remains generally unchanged as a function of F1 (Figure 11a) or F2 (Figure 11b).

TABLE VIII. Coefficients, standard errors, and t-values of CQ as a function of F1 and F2 with participant as a random intercept. Nested model comparison revealed that none of the factors improved model fit (F1, F2, or the interaction).

Fixed Effect	Estimate	Standard Error	T-Value
Intercept	<0.01	<0.01	13.60
F1	<0.01	<0.01	2.91
F2	<0.01	<0.01	2.72
F1*F2	<0.01	<0.01	-3.16

IV. DISCUSSION

The purpose of the study was to examine IF0 of Amharic vowels and compare predictions of different articulatory hypotheses that have been proposed in the literature. Predictions of each hypothesis are outlined in Table II. Under the tongue-pull hypothesis, high vowels in Amharic were expected to have the highest F0, followed by mid and low vowels, which were expected to pair together. According to the tongue-compression hypothesis, high and mid vowels were expected to pair together, and low vowel /a/ was predicted to have the lowest F0. If both tongue-pull and tongue-compression both play a role in IF0, high, mid, and low vowels were expected to differ in F0. Additionally, since both hypotheses relate the cause of IF0 to differences in vocal fold tension, differences in contact quotient were expected to follow the predicted F0 patterns. Since testing the horizontal-vertical pull hypothesis requires measuring muscle activation (e.g., with electromyography), predictions for this hypothesis were not described. Similarly, the current study did not directly assess the contribution of jaw height to F0, though F0 was expected to vary generally by vowel height based on previous studies.

The study found that F0 varied as a factor of vowel category due to differences in vowel height. High vowels had the highest F0, followed by mid vowels, and finally, low vowels. No significant differences were found within height groups (e.g., /i/ was not different from /u/). Similarly, when examining the effect of F1 and F2 on F0, the study found that F1 is predictive of F0 differences, while F2 is not. F0 varied by vowel height but not by backness. With respect to CQ, the study revealed no effect of vowel category on CQ, nor was there an effect of F1 or F2 on CQ independent of vowel category.

TABLE IX. Coefficients, standard errors, t-values, and p-values of F1 as a function of vowel with participant as a random intercept to compare /ɨ/ to /i/ and mid vowels. P-values for this model were obtained using the *lmerTest* package (Kuznetsova et al., 2017).

Vowel	Estimate	Standard Error	T-Value	P-Value
ɨ (Intercept)	429.11	22.61	19.32	<.001
i	-88.36	4.33	-20.40	<.001
e	10.02	4.31	2.33	0.02
u	-47.24	4.33	-10.91	<.001
ə	170.36	4.30	39.66	<.001
o	52.62	4.31	12.21	<.001
a	331.9	4.28	77.58	<.001

Another interesting finding emerged from the data. Vowel F0 seemed to pattern together based on phonological height, despite differences in F1 within a phonological height group. For example, /ɨ/ and /i/, both high vowels, have different F1 values despite having similar F0 values in the pairwise comparison. A post-hoc linear mixed effects model was fitted to the data with *vowel* entered as a fixed effect and *participant* as a random intercept to explore differences in F1 as a function of vowel category. The model revealed a significant difference between /ɨ/ and /i/. The coefficients and intercepts of the model are presented in Table IX.

TABLE X. Coefficients, standard errors, t-values, and p-values of F1 as a function of vowel with participant as a random intercept to compare /e/ to /ə/ and /o/. P-values for this model were obtained using the *lmerTest* package (Kuznetsova et al., 2017).

Vowel	Estimate	Standard Error	T-Value	P-Value
e (Intercept)	439.12	22.19	19.79	<.001
ɨ	-10.02	4.31	-2.33	0.02
i	-98.37	4.26	-23.09	<.001
u	-57.26	4.26	-13.44	<.001
ə	160.35	4.22	37.98	<.001
o	42.61	4.24	10.06	<.001
a	321.88	4.20	76.56	<.001

In a similar vein, all mid vowels, /e, ə, o/ were found to have similar F0 values as they did not differ significantly from each other in the pairwise comparison; however, mid vowels also had different F1 values from each other. An additional post-hoc linear mixed effects model was performed with factor re-leveling to allow for comparison of mid vowels. The results revealed significantly different F1 values between /e/ and /o/ as well as between /e/ and /ə/ (Table X). While /ə/ and /o/ were not compared to each other, given the difference in coefficients and t-values, it is likely these two vowels have significantly different F1 values as well.

The findings from these exploratory analyses are similar to those by Turner & Verhoeven (2011), who found that F0 varied by vowel height in Belgian Standard Dutch, though close-mid and open-mid vowels did not differ from each other despite having differences in F1. This suggests that there may be multiple factors driving IF0. On the one hand, the universality of IF0 and evidence from articulatory studies suggest IF0 is an automatic consequence of mechanical coupling between articulators. Indeed, the general trend of F0 varying by height in this study seems to support the hypothesis that IF0 is a result of mechanical coupling. On the other hand, the phonological grouping of F0 despite differences in F1 suggests that other factors contribute to F0, such as auditory enhancement or forming part of the articulatory goal for vowel production (Sapir's horizontal-vertical pull hypothesis). The implications of these findings as they relate to the proposed causes of IF0 are discussed below.

The tongue-pull hypothesis cannot account entirely for the results of the study for three reasons. First, F0 of mid and low vowels did not pair together. Instead, F0 was found to be gradient with different values for high, mid, and low vowels; this finding was expected based on results of previous studies. Second, the study found that high vowels, /i/ and /ɨ/, and mid vowels /e, ə, o/ differed significantly with respect to F1 yet did not differ in F0. In other words, differences in F1 did not correlate with differences in F0 for these groups of vowels, which does not support the prediction that higher tongue position necessarily leads to increased F0. Instead, a phonological grouping of F0 was found between high and mid vowels. Finally, there were no differences found in CQ, suggesting no changes in vocal fold tension. Therefore, the tongue-pull hypothesis does not appear to be the best account of IF0. It is possible that tongue-pull contributes to the effect for high vowels, but tongue-pull alone does not account for the data.

Likewise, the tongue compression hypothesis cannot fully account for this study's findings for the reasons cited above: namely, the gradient effect and lack of difference in CQ for /a/ compared to other vowels. Tongue retraction may contribute to laryngeal changes that cause reduced F0, but the laryngeal changes due to tongue retraction do not seem to be those predicted in the hypothesis: thick, slack vocal folds. Instead, we see that vocal fold tension, as measured via CQ, during low vowel /a/ is similar to other vowels. Interestingly, the CQ results did not reveal *increased* contact/tension for /a/, predicted by Esling 2005 and Moisik et al. 2019.

There are possible explanations for the lack of difference in CQ between vowels. Recall that EGG measures vocal fold contact area and increased contact is typically a result of increased laryngeal constriction or thyroarytenoid contraction. Previous studies that use CQ as a measure of vocal fold tension compared very different phonation patterns, such as normal and disordered voicing (Childers, 1990), hypophonic and hyperfunctional voicing (Szkiekowska, Krasnodebska, Miaskiewicz & Skarżyński, 2018), or modal and non-modal phonation (Scherer, 1987; Kochetov, 2020; Herbst, 2020). Participants in the present study were judged to have normal voice quality, reported no history of voice or speech disorders, and produced speech with modal phonation during the study task. Therefore, it might be the case that there are small changes in vocal fold *tension*, but they were not captured in this study because there were not significant changes in vocal fold *contact area* that are typically seen with different types of phonation. In this case, CQ might not be appropriate to measure small differences in tension with modal voicing, or the sample size needs to be much larger to capture the very small effect size with modal voicing. Of course, it is also possible that there were no changes in vocal fold tension that caused the F0 changes seen in the IF0 effect. If the mechanism behind IF0 causes a passive stretching of the vocal folds without activation of the cricothyroid or thyroarytenoid, there would be no increase in thyroarytenoid activation and potentially no laryngeal constriction causing changes in vocal fold contact area. In this case, the CQ results correctly captured the null effect.

Finally, the F0 findings may be consistent with the jaw height hypothesis. Recall that though F1 is thought to reflect tongue height, Chen et al. (2019) found that jaw height was a better predictor of F0 than tongue height. Therefore, differences in F0 between high, mid, and low vowels, may reflect different jaw heights. For the mid vowels, tongue height might differ while jaw height is similar, explaining the finding of different F1 values for mid vowels while they had similar F0 values. In addition, as previously stated, tongue compression may play a role in further lowering F0 beyond what is expected by jaw height alone. Therefore, IF0 may be an effect of multiple mechanisms including lingual and mandibular movement. Additionally, auditory enhancement cannot be ruled out as a contributing factor in the F0 results of this study. The F0s values may be a result of auditory enhancement of the phonological grouping of the vowels to contrast high, mid, and low vowels in Amharic. Sapir's vertical-horizontal pull hypothesis also cannot be ruled out as a contributing factor as the phonological grouping of F0 might be due to differences in the articulatory goal of the vowels. While no independent evidence from Amharic explains why these phonological groupings emerge, findings for Dutch in Turner & Verhoeven 2011 indicate that the phonological grouping of F0 is not unique to Amharic and warrants further cross-linguistic investigation.

V. CONCLUSION

The cause of intrinsic fundamental frequency of vowels has been long studied yet still remains unknown. Proposed hypotheses that account for the universal phenomenon primarily relate IF0 to mechanical coupling between oral vocal tract structures and the larynx, where movement of the oral structures causes changes in vocal fold tension, which lead to perturbations in F0. The purpose of the present study was to investigate IF0 in Amharic vowels in an effort to compare hypotheses. The results of the current study suggest that IF0 may be due to multiple mechanisms including the effect of jaw displacement creating gradient perturbations in F0 and tongue compression further decreasing pitch for low back vowels, though the exact link between tongue or jaw movement and the larynx remains an open question.

Hypotheses appealing to categorical distinctions arising from the articulatory gesture for the vowel (horizontal-vertical pull hypothesis) or auditory enhancement may also play a role in the effect as they could not be ruled out based on the current study. In this study, F0 differed by categorical distinctions (i.e., high vs. mid vs. low vowels). However, within height categories, vowels were found to have significantly different F1 values, despite having similar F0 values. This finding can be explained with two possible hypotheses: first, jaw height between the vowels did not differ but F1 did; second, the F0s values of high, mid, and low vowels is a result of auditory enhancement of the phonological grouping of the vowels to contrast them. Both causes may contribute to results seen in this study.

Results of the current study raise questions and directions for future work. First, the role of tongue compression in lowering F0 merits further investigation. Currently, the connection between tongue retraction and laryngeal movement remains understudied in the literature. Esling (2005) and Moisik et al. (2019) suggest that tongue retraction leads to increased laryngeal constriction. However, no evidence of laryngeal constriction with /a/ was found in this study.   Further cross-linguistic studies comparing low front vowels to low back vowels may address the question of whether tongue retraction affects F0. Additionally, articulatory studies such as ultrasound or MRI may shed light on the mechanical link between oral vocal tract structures and the larynx. In particular, the effect of jaw movement vs. tongue-height on F0 merits further systematic, cross-linguistic investigation as such studies are limited in the literature. Finally, future studies of IF0 should include all vowels in the vowel space to examine potential phonological constraints on IF0.

ACKNOWLEGEMENTS

I would like to thank Marc Garellek for his advising on this project as well as Sharon Rose and Will Styler for their comments. I would also like to thank Hira Rizvi for her help with annotations and B. Ramos for designing Figures 1-4 for use in this paper.

APPENDIX

List of stimulus words in carrier phrase ይሀ ቃል _____ ነው, /jɨh k’al _____ nəw / (*This word is* _____) . Each Amharic character represents a CV syllable in the stimulus word.

Amharic	IPA	Amharic	IPA
1. ተቲ	/təti/	22. ተተ	/tətə/
2. ቱቲ	/tuti/	23. ቱተ	/tutə/
3. ቲቲ	/titi/	24. ቲተ	/titə/
4. ታቲ	/tati/	25. ታተ	/tatə/
5. ቴቲ	/teti/	26. ቴተ	/tetə/
6. ትቲ	/tɨti/	27. ትተ	/tɨtə/
7. ቶቲ	/toti/	28. ቶተ	/totə/
8. ተቱ	/tətu/	29. ተቶ	/təto/
9. ቱቱ	/tutu/	30. ቱቶ	/tuto/
10. ቲቱ	/titu/	31. ቲቶ	/tito/
11. ታቱ	/tatu/	32. ታቶ	/tato/
12. ቴቱ	/tetu/	33. ቴቶ	/teto/
13. ትቱ	/tɨtu/	34. ትቶ	/tɨto/
14. ቶቱ	/totu/	35. ቶቶ	/toto/
15. ተታ	/təta/	36. ተቴ	/təte/
16. ቱታ	/tuta/	37. ቱቴ	/tute/
17. ቲታ	/tita/	38. ቲቴ	/tite/
18. ታታ	/tata/	39. ታቴ	/tate/
19. ቴታ	/teta/	40. ቴቴ	/tete/
20. ትታ	/tɨta/	41. ትቴ	/tɨte/
21. ቶታ	/tota/	42. ቶቴ	/tote/

REFERENCES (BIBLIOGRAPHIC)

Ado, Derib. (2011). An Acoustic Analysis of Amharic Vowels , Plosives and Ejectives. Ph.D. thesis Addis Ababa University Addis Ababa.

Bailly, L., Bernardoni, N., Muller, F., Rohlfs, A.-K., & Hess, M. (2014). Ventricular-fold dynamics in human phonation. *Journal of Speech, Language, and Hearing Research, 57*, 979-989.

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48.

Boersma, P., & Weenink, D. (2020). Praat: doing phonetics by computer.

Bush, M. (1981). Vowel Articulation and Laryngeal Control in the Speech of the Deaf. Ph.D. thesis MIT Boston.

Chen, W., Whalen, D. H., & Tiede, M. K. (2019). Mandibular contribution to vowel-intrinsic f0. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 152-156). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Childers, D. G. (1990). Electroglottography and Vocal Fold Physiology. *Journal of Speech Language and Hearing Research, 33*, 245-254.

Diehl, R. L. (1991). The Role of Phonetics within the Study of Language. *Phonetica, 48,* 120-134.

Erickson, D., Honda, K., & Kawahara, S. (2017). Interaction of jaw displacement and F0 peak in syllables produced with contrastive emphasis. *Acoustical Science and Technology, 38*, 137-146.

Esling, J, H. (1996). Pharyngeal consonants and the aryepiglottic sphincter. *Journal of International Phonetic Association, 26,* 65-88.

Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics/Revue Canadienne de linguistique, 50*, 13-44.

Ewan, W. G. (1979). Can intrinsic vowel F0 be explained by source / tract coupling? *The Journal of the Acoustical Society of America, 66*, 358-362.

Fantaye, T. G., Yu, J., & Hailu, T. T. (2019). Acoustic characterization of Amharic vowel sound units. *ACM International Conference Proceeding Series*, (pp. 36-40).

Flanagan, J., & Landgraf, L. L. (1968). Self-oscillating source for vocal tract synthesizers. *IEEE Transactions on Audio and Electroacoustics, 16,* 57-64.

Gandour, J., & Weinberg, B. (1981). On the relationship between vowel height and fundamental frequency : Evidence from esophageal speech. *Phonetica, 37,* 344-354.

Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard, 3*, 1-21.

Guérin, B., & Boë, L. (1980). Study of the influence of acoustic coupling of the voice source on F0 of oral vowels: Consequences for the study of intrinsic characteristics. *Phonetica, 37*, 169-192.

Hayward, K., & Hayward, R. J. (1992). Amharic. *Journal of the International Phonetic Association, 22*, 48-52.

Herbst, C. T. (2020). Electroglottography: An update. *Journal of Voice*, *4*, 503-526.

Holt, Y. F., Jacewicz, E., & Fox, R. A. (2019). Social cultural influences of fundamental frequency in vowels: The case of African American English. In P. W. Sasha Calhoun, Paola Escudero, Marija Tabain (Ed.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 874-878).

Hong, Y. T., Hong, K. H., Jun, J.-p., & Hwang, P. H. (2015). The effects of dynamic laryngeal movements on pitch control. *American Journal of Otolaryngology-Head and Neck Medicine and Surgery, 36*, 660-665.

Howard, D. M. (1995). Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice, 9*, 163-172.

Jacewicz, E., & Fox, R. A. (2015). Intrinsic fundamental frequency of vowels is moderated by regional dialect. *The Journal of the Acoustical Society of America*, *138*, EL400-EL410.

Kawahara, H., de Cheveign, A., & Patterson, R. (1998). An instantaneous-frequency based pitch extraction method for high quality speech transformation: revised TEMPO in the STRAIGHTsuite. In *Proceedings ICSLP'98, Sydney, Australia, December 1998*.

Kirby, J., & Ladd, D. (2016). Effects of obstruent voicing on vowel F0: Evidence from "true voicing" languages. *The Journal of the Acoustical Society of America,* 140.

Kirby, J. P., & Ladd, D. R. (2015). Stop Voicing and F0 Perturbations: Evidence From French and Italian. In *Proceedings of the 18th International Congress of Phonetic Sciences*.

Kochetov, A. (2020). Research methods in articulatory phonetics II: Studying other gestures and recent trends. *Lang Linguist Compass.,* (p. e12371).

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest Package: Tests in linear mixed effects models. J*ournal of Statistical Software, 82*, 1-26.

Ladefoged, P. (1964). Some possibilities in speech synthesis. *Language and Speech, 7*, 205-214.

Lancia, L., & Grawunder, S. (2014). Tongue-larynx interactions in the production of word initial laryngealization over different prosodic contexts: a repeated speech experiment. In S. Fuchs, M. Grice, A. Hermes, L. L, & D. Mnucke (Eds.), *Proceedings of 10th (ISSP)* April 2014 (pp. 245-248). Cologne.

Laufer, A. & Condax, I.D. (1979). The epiglottis as an articulator. *Journal of the International Phonetic Association, 9,* 50-56.

Lenth, R. V. (2016). Least-square means: The R package lsmeans. *Journal of Statistical Software, 69*, 1-33.

Moisik, S. R., Esling, J. H., Crevier-Buchman, L., & Halimi, P. (2019). Putting the larynx in the vowel space: Studying larynx state across vowel quality using MRI. In *The 19th International Congress of Phonetic Sciences* (pp. 1-5).

Ohala, J. (1978). Production of tone. In V. A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 5-39). New York: Academic Press.

Ohala, J., & Eukel, B. (1987). Explaining the intrinsic pitch of vowels. In R. Channon, & L. Shockey (Eds.), *In honor of Ilsa Lehiste*. Dordrecht: Foris.

Pape, D., & Mooshammer, C. (2006). Intrinsic F0 differences for German tense and lax vowels. In *Proceedings of the International Seminar on Speech Production ISSPAt: Ubatuba, Brasil.*

R Core Team (2019). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing,* Vienna, Austria.

Sapir, S. (1989). The intrinsic pitch of vowels: Theoretical, physiological, and clinical considerations. *Journal of Voice, 3*, 44-51.

Scherer, R. C., & Titze, I. R. (1987). The abduction quotient related to vocal quality. *Journal of Voice, 1*, 246-251.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. In *Proceedings of the ICPhS XVII* (pp. 1846-1849).

Sjolander, K. (2004). Snack sound toolkit.

Szkiekowska, A., Krasnodebska, P., Miaskiewicz, B., & Skarz̀yński, H. (2018). Electroglottography in the diagnosis of functional dysphonia. *Eur Arch Otorhinolaryngol*, *275*, 2523-2528.

Turner, P., & Verhoeven, J. (2011). Intrinsic vowel pitch: a gradient feature of vowel systems? In *The 17th International Congress of Phonetic Sciences*, 17-21 August 2011, Hong Kong.

Whalen, D. H., Gick, B., Kumada, M., & Honda, K. (1998). Cricothyroid activity in high and low vowels: exploring the automaticity of intrinsic F0. *Journal of Phonetics*, 27, 125-142.

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics,* 23, 349-366.

Whalen, D. H., Levitt, A. G., Hsiao, P.-l., & Smorodinsky, I. (1995). English-learning infants Intrinsic F0 of vowels in the babbling of 6- , 9- , and 12-month-old French and English-learning infants. *The Journal of the Acoustical Society of America*, 97, 2533-2539.

Zawadzki, P. A., & Gilbert, H. R. (1989). Vowel fundamental frequency and articulator position. *Journal of Phonetics*, 17, 159-166.

Zhang, Z. (2016). Mechanics of human voice production and control. *Journal of the Acoustical Society of America,* 140, 2614-2635.