The acoustic correlates of stress and tone in Chácobo (Pano): A production study

Article *in* The Journal of the Acoustical Society of America · April 2020

DOI: 10.1121/10.0001014

2 authors:

Adam Tallman
Friedrich Schiller University Jena
17 PUBLICATIONS **52** CITATIONS
SEE PROFILE

Jose Elias-Ulloa
University of Oxford
29 PUBLICATIONS **115** CITATIONS
SEE PROFILE

The acoustic correlates of stress and tone in Chácobo (Pano): A production study

Adam J. R. Tallman, and José Elías-Ulloa

Citation: The Journal of the Acoustical Society of America **147**, 3028 (2020); doi: 10.1121/10.0001014
View online: https://doi.org/10.1121/10.0001014
View Table of Contents: https://asa.scitation.org/toc/jas/147/4
Published by the Acoustical Society of America

JASA ARTICLE

The acoustic correlates of stress and tone in Chácobo (Pano): A production study

Adam J. R. Tallman[1] and José Elías-Ulloa[2,a)]
[1]*Université de Lyon II, Lyon, France*
[2]*Stony Brook, New York 11794, USA*

ABSTRACT:
This paper investigates the acoustic correlates of word initial prominence and tone in Chácobo, a southern Pano language of the northern Bolivian Amazon. This paper reports the results of a production study with five speakers producing trisyllabic words with a word final high tone. It tests a claim found in the literature that there is an additional word-initial prominence in such forms and determines its acoustic correlates compared with high tone. This study used trisyllabic forms in morphophonological contexts where these forms would appear with a final high tone. In such forms, high tone and word-initial prominence do not overlap (e.g., "panaβ̰í" "asaí"). The paper takes into account five acoustic correlates across the three syllables of these words: *F*1, *F*2, *F*0, duration, intensity. The paper finds that the initial syllable in these words shows a statistically significant increase in intensity. There is significant speaker variation with respect to whether duration is a correlation of initial stress and the results do not provide clear evidence that initial stress is marked with duration. The final high-tone marked syllables are distinguished based on an increase in *F*0 and secondarily with vowel duration. Whether pitch is also a correlate of stress requires future research. The paper interprets these results as suggesting that Chácobo is a language in which tone and stress co-exist. The study thus provides instrumental evidence for a phenomenon suggested to occur in some Pano languages. The limitations of this study, including the fact the results have more than one interpretation in light of current discussions concerning stress and tone, are also discussed. © *2020 Acoustical Society of America*.
https://doi.org/10.1121/10.0001014
(Received 30 April 2019; revised 2 March 2020; accepted 9 March 2020; published online 30 April 2020)
[Editor: Richard A. Wright] Pages: 3028–3042

I. INTRODUCTION

Chácobo is an underdescribed southern Pano language of the northern Bolivian Amazon. High tones can occur on one of the first two syllables of a word, although they are often reduced or deleted through tone sandhi rules. Suffixes and enclitics also contain high tones in Chácobo (see Tallman, 2018). Although Chácobo is a tonal language, there have been reports based on impressionistic observations that the language has non-tonal prominence in certain syllables. In some cases, such prominence has been interpreted as stress or as the presence of a mid tone (e.g., Prost 1960, 1967; Shell, 1975, p. 56; Iggesen and Pilar, 2007; Iggesen 2006, 2007; Tallman 2017, 2018). In this article, we are interested in investigating one of those locations reported to have non-tonal prominence; that is, prominence on the initial syllable of a Chácobo word when the first and second syllables are both forced to surface toneless. This particular case occurs, for instance, when the ergative enclitic combines with a trisyllabic noun. The Chácobo ergative morpheme is a floating high tone that must dock onto the final syllable of the noun phrase. This floating high tone demands to be the only high tone on the noun: all other high tones are removed or blocked from surfacing. This results in trisyllabic nouns where the first two syllables are toneless. This study investigates whether the first two syllables in such nouns still show some type of prominence despite the fact that they do not host a high tone.

Our results show that the first syllable of Chácobo trisyllabic nouns in the ergative case tends to show a spike in intensity compared to the second syllable. Duration might be a correlate, but the speaker variation present in the results, together with the fact that we only have five participants, make it hard to draw a general conclusion. These results are important because they show, from an acoustic perspective, that tonal languages can use phonetic properties implemented, for instance, through intensity, to indicate the prominence of certain syllables when tones are not involved (see Gordon and Roettger, 2017, p. 6).

This article is structured as follows. Section II provides general information about the Chácobo language and its speakers. Section III presents the basic properties of Chácobo phonology, particularly about its segmental inventory and the distribution of high tones. The possibility that the phonological facts should be understood primarily in terms of a primary and secondary stress distinction is also taken up. Section IV presents the stimulus design; Sec. V, presents information about the language consultants that participated in the research project. Section VI describes the different types of measurements carried out. Section VII

a)Electronic mail: jose.elias-ulloa@stonybrook.edu, ORCID: 0000-0003-4448-8522.

https://doi.org/10.1121/10.0001014

discusses the statistical analyses used and their relationship to the hypotheses tested and explored in this paper. Sections VIII and IX present the results of the statistical analyses. Section X discusses the results in terms of the main hypotheses of this paper. It also discusses the typological and theoretical relevance of the results. Section XI discusses the limitations of this study. We offer concluding remarks in Sec. XII.

II. CHÁCOBO LANGUAGE: SPEAKERS AND FIELDWORK CONTEXT

Chácobo is a Southern Pano language spoken by approximately 1200 people in about 20 communities on the Geneshuaya, Ivon, Benicito, and Yata rivers in the Northern Bolivian Amazon. The communities are located in the Chácobo-Pacahuara Original Communitarian Land (Tierra Comunitaria de Origen, TCO), a region where the Chácobo and Pacahuara exert control over resources and land-use to varying degrees. In the communities within the TCO, the language is spoken in all aspects of daily life. Children still learn Chácobo as their first language, acquiring Spanish in school. Of the 1200 Chácobo speakers, there are approximately 300 that live in the neighboring Bolivian town of Riberalta, to the north of the TCO. The approximate location of the Chácobo TCO is provided in Fig. 1 (see Tallman, 2018, for more information).

The Chácobo learn a Spanish-based orthography that was developed by Summer Institute of Linguistics (SIL) missionaries in the 1960s. Despite the importance of tone in Chácobo grammar, it is not consistently marked in the orthography, although they are able to consistently identify tonally marked segments if asked. Nevertheless, the fact that most speakers can read segmental representations of their language means that production studies can be conducted by having speakers read constructed word lists from a computer screen. This paper focuses on the results of a production study conducted with five Chácobo speakers. Three of the production tasks were conducted in Riberalta and two were conducted in Alto Ivon, the largest Chácobo community. This study was conducted in the context of a documentation project of the Chácobo language undertaken by one of the coauthors.

III. CHÁCOBO PHONOLOGY

Chácobo has a relatively small segmental inventory with four vowels /i, o, a, ɨ/ and 15 consonants /p, k, t, ts, tʃ, m, n, y, w, β̞, ɾ, s, ʂ, ʃ, h/. Syllable structure is (C)V(C), where onsets and codas are optional. While any consonant can occupy the onset position, only sibilants can occur in the coda position. Most roots are disyllabic, although trisyllabic roots are not rare at all. The data in (1) and (2) illustrate the different syllable structures found in Chácobo both in disyllabic and trisyllabic words, respectively. The language has only one tonal unit: a high tone that is usually implemented as rising fundamental frequency ($F0$) between a low tonal target and a high one. In terms of Hyman (2006, 2009) classifications, Chácobo has a privative tonal system in which a high tone opposes to zero. We mark the high tone with an acute accent.

(1) a. β̞á.ʂɨʂ “whisper”
b. bɨ.mɨ́s “good hunter”
c. tsí.pis “fart”
d. nɨʂ.tí “catfish”
e. páʂ.na “hunger”
f. β̞a.kɨ́ “boy, child”
g. ká.tɨ “back”

(2) a. waʂ.mɨ́.nɨ “cotton”
b. a.wí.ni “woman, wife”
c. po.ʔí.ko “vulture”

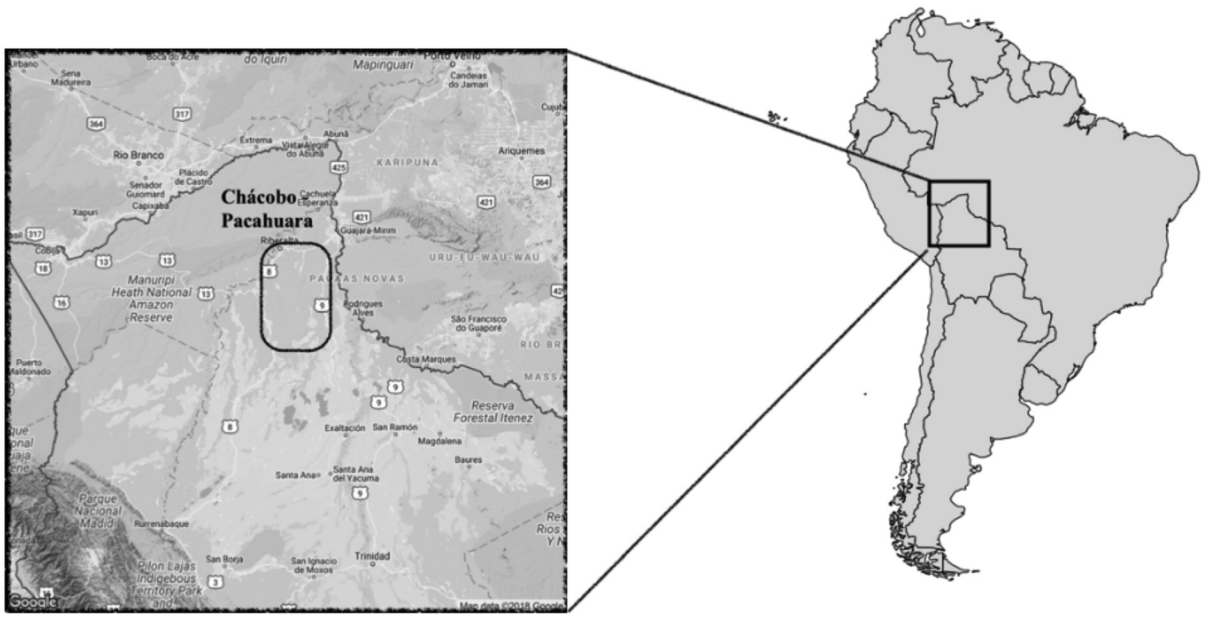

FIG. 1. Approximate location of Chácobo speakers in the northern Bolivian Amazon.

d. tɬ́.mis.ko "island"
e. ʂó.to.ko "pigeon"
f. ká.to.ko "kidney"
g. pa.ʔo.kí "ear"
h. nɬ.ʂo.mí "duck"
i. ʂoʂ.to.ɾó "spring"

According to the available sources (Tallman 2017, 2018), a low pitch is mapped to all syllables that do not contain a high tone such that a bisyllabic form which is underlyingly ØH is realized as LH. The results of our study call the latter aspect of the phonological description into question, implying that a distinction between low tone marked syllables and toneless syllables is important for a correct description of the phonetics even if a distinction between low and high tone is not phonologically contrastive.

Despite the transcription practices adopted throughout this paper, which follow Tallman (2018), there is a possibility that underlying H tones are in fact bitonal L + H units and that other syllables are unspecified. This is different from the description in Tallman (2018), which assumes that all syllables that do not contain an underlying tone receive a default low tone and that there is no difference between different syllables that do not bear an underlying high tone with respect to tonal structure. The possibility of a bitonal L + H analysis emerges from careful consideration of the spectrograms and is not reflected in previous descriptions of the language. We note that the current study does not argue that the bitonal L + H analysis of underlying tone is clearly true, rather we point it out as a possibility that confounds the interpretation of some of the acoustic correlates to initial prominence (see Sec. VI).

As illustrated by the data in (3) and (4), the high tone is lexically contrastive

(3) a. ká.ʂa "be angry (v., intr.)"
b. ka.ʂá "play (v. intr.)
c. tɬ́.sa "cook (v., tr.)
d. tɬ.sá "loosen (v., tr.)"
(4) a. má.nɬ "knife (n)"
b. ma.nɬ́ "transform (v., intr.)"
c. á.ni "lake (n)"
d. a.ní "grow (v.)"

Certain affixes are lexically specified with their own high tone while others have no tonal specification. As a result, a single prosodic word can occur with more than one high tone. This can be observed in (5) (see the description from section 6.2.1 of Tallman, 2018).

(5) [pí.tɬ.kɬ́.kɬ]
/*pí-tɬk-ɬ́=kɬ*/
wing-break-itr=dec:past
"S/he broke her/his own arm."

In Chácobo, adjacent high tones undergo a tone reduction process such that HH changes to ØH. An example of the tone reduction rule applying at a root+affix boundary is provided in (6b) [compare with (6b) and (6c) where there is the right adjacent clitic does not contain a high tone].[1]

(6) a. [ha.ʂa.já.ma.kɬ]
haʂá=jáma=ki
run-neg=dec:nonpast
"S/he didn't run."
b. [ha.ʂá.kɬ]
haʂá=ki
run=dec:nonpast
"S/he ran."
c. [ha.ʂá.tsi.ki]
haʂá=tsi=ki
run=imm=dec:nonpast
"S/he is running now."

Examples of the tone reduction rule applying between words are provided in (7) and (8)

(7) [ʃi.no.ká.ʂa.kɬ]
ʃinó káʂa=kɬ
monkey be_angry=dec:past
"The monkey became angry"
(8) [ʃi.no.hó.ʂo]
ʃinó hóʂo
monkey white
"white monkey"

In Chácobo, roots vary in terms of whether they contain an underlying high tone. There are roots with no tonal specification, like the proper name /ɾaʂi/ "Rabi," while others do come with a high tone, like the proper name /ɾámi/ "Rami." Prior to the clause-type morpheme (=*kɬ* in the example below), these nouns both show a high tone but when they occur after the clause-type rank morpheme; as shown in (9) and (10), /ɾaʂi/ surfaces completely toneless, while /ɾámi/ does show its high tone (see Tallman, 2018, p. 95 for spectrograms of these examples).

(9) [ɨ.a. tsá.ja.kɬ. **ɾa.ʂi**]
/ɨ-*a* *tsaja=kɬ* ***rabi***/
1sg-acc see=dec:past **Rabi**
"Rabi has looked at me."
(10) [ɨ.a. tsá.ja.kɬ. **ɾá.mi**]
/ɨ-*a* *tsaja=kɬ* **ɾámi**/
1sg-acc see=dec:past **Rami**
"Rami has looked at me."

By default, as in (11), underlyingly toneless roots are assigned a high tone on the initial syllable when they occur prior to the clause-type rank morpheme. However, high tone insertion can be blocked if a high-tone bearing affix or clitic is attached to the root. This is the case of =*jó* "completive," which blocks the insertion of any high tone, so if the verb root does not have one underlyingly, it will appear toneless. This case is shown in (12). If the root has its own underlying high tone, then =*jó* cannot remove it, and, as shown in (13), the word surfaces with more than one high tone.

(11) Inserted high tone on the first syllable
[há.na.kɬ]
/hana=kɬ/
leave=dec:past
"S/he left it/them/him/her"

https://doi.org/10.1121/10.0001014

(12) Clitic with lexical tone blocks high tone insertion on root
[ha.na.jó.kɨ]
/hana=jó=kɨ/
leave=cmpl=dec:past
"S/he left everything."

(13) Word with more than one high tone
[ká.ʂa.jó.kɨ]
/káʂa=jó=kɨ/
become.angry=cmpl=dec:past
"They all became angry."

A detailed description of the distribution of underlying high tones, high-tone insertion, tone sandhi is provided in Tallman (2018, pp. 84–102, 415–480). Note that the only default tone insertion rule inserts the high tone on the first syllable, but this only occurs if no other high tones are present in the prosodic word as in (11). There is no evidence that any of the other high tones present in Chácobo grammar are inserted through rule. Apart from the possibility that the first syllable potentially has an increase in intensity, the fact that it is the only target of high tone insertion provides some additional evidence that this syllable should be regarded as stressed. This should not be meant to imply that there is evidence for metrical structure throughout Chácobo grammar as implied by Tallman (2017) (see below).

Chácobo also has floating high tones, some of which code grammatical categories (e.g., genitive, ergative, vocative). As with many cases of grammatical tone (see Rolle, 2018), grammatical tone in Chácobo conditions tone sandhi rules which are not general across the grammar. Ergative case is marked by a floating high tone that docks to the final syllable of a noun phrase in transitive subject (A) position. The grammatical floating high tone erases all other underlying high tones on the element it combines with (see Rolle, 2018, pp. 19–33, for such processes from a typological perspective). This is illustrated in (14): the noun *ínaka* "dog, pet" surfaces with just one high tone. It loses its underlying high tone on the first syllable when the ergative high tone docks to the final syllable: [i.na.ká], *[í.na.ká].[2] We capitalize on this aspect of Chácobo tonology in our construction of word lists since it allows us to construct example sentences with words that have a predictable tonal shape (see Sec. IV).

(14) [i.na.ká mi.tʃi tɬ́.ʂa.kɨ]
/ínaka=́ mitʃi tɬ́ʂ-a=kɨ/
dog/pet=erg cat bite-tr=dec:past
"The dog bit the cat."

We point out that two primary sources on Chácobo prosody (Tallman, 2017, 2018) adopt different terminology to describe the prosodic patterns of the language. Based on data gathered during a short pilot fieldtrip conducted in 2011, Tallman (2017) presents an analysis that accounts for the distribution of some of the pitch patterns on simplex bisyllabic and trisyllabic noun roots when elicitated in isolation. Later on, based on a longer and more comprehensive fieldwork conducted between 2012 and 2018, Tallman (2018) presents a more thorough analysis that accounts for the prosodic patterns across the whole grammar. Tallman (2017) describes the high tones as being mapped on to underlyingly lexically stressed syllables. On the other hand, Tallman (2018) describes the high tones as underlyingly present, following the description we have provided above.[3]

One predictive difference between Tallman (2017) and Tallman (2018) is that the Tallman (2017, p. 111) analysis entails that it is impossible to have more than one high tone in a prosodic word. However, that high tone is not culminative in the prosodic word in Chácobo as can be seen from (5) and (13). Tallman (2017, p. 108) also states that high tone is obligatory in the phonological phrase but provides no phonological phrasing rules. His position is based only on the stipulation the noun compounds are phrases in Chácobo, for which no evidence is provided (see Tallman, 2018, pp. 463, 1106–1117, for a description of noun compounds and noun-noun combinations in Chácobo). If we assume as a null hypothesis that the phonological phrase maps over noun phrases, then examples in (7) and (9) also provide counterevidence against the Tallman (2017) analysis, because *rabi* "Rabi (Chácobo name)" should receive a high tone according to Tallman's rules of high tone insertion. This is not an issue related to word versus phrasal prominence because *rámi* "Rami (personal name)," which has an underlying high tone, does surface with a high tone on the first syllable, as discussed by Tallman (2018, p. 93).

Another difference between Tallman (2017) and Tallman (2018) regards the position of default high tone assignment. For the latter analysis, this would have to be the second syllable of a word (Tallman, 2017, p. 111). However, there is no evidence that this is the default pattern. In fact, the evidence suggests that if a default high tone insertion is needed, it is word initial as shown by considering the examples in (9) and (10) (see Tallman, 2018, pp. 84–102, 416–438, for more details). The fact that the only position in Chácobo grammar where there is evidence for an inserted high tone is the first syllable is one of the reasons why this syllable is considered stressed by Tallman (2018).

Finally, Tallman (2017) proposes to account for the distribution of high tones metrically following de Lacy (2002). However, in order to provide evidence for a metrical system in the context of a tone language, there should be independent evidence for it (see Gordon, 2016, p. 218). Such evidence would have to account for phenomena independently of what is already accounted for by the tone sandhi rules and the presence of underlying high tones, both of which are independently necessary in Chácobo (see Tallman, 2018, pp. 415–480, for a detailed discussion of the tone facts[2]). The fact that one can imagine that iambic or trochaic feet are present on any given sequence of a low pitch followed by a high pitch or a high pitch followed by a low pitch does not in and of itself provide evidence for a metrical system.

Another reason for adopting the terminological framework of Tallman (2018) instead of Tallman (2017) is that the latter study only accounted for simplex (single morpheme) bisyllabic and trisyllabic noun roots elicited in

isolation, whereas Tallman (2018) accounts for tone patterns throughout the entire grammar. Regardless of these issues, even if the lexical/post-lexical distinction is posited and the analysis of Tallman (2017) is somehow redesigned to account for the Chácobo data, the issue is irrelevant to the paper, because, in the context of bisyllabic and trisyllabic nouns, which is the data examined here, both analyses make identical predictions regarding the surface phonetic patterns. Nevertheless, in Sec. XI, we discuss the possibility of interpreting the difference between stress and high tone as a difference between secondary and primary stress in light of the results of the production study, even though the latter analysis is not supported by the phonological facts.

The fact that the Tallman (2017) preliminary analysis makes incorrect predictions does not necessarily rule out some type of metrical analysis accounting for the phenomena if perhaps the facts reviewed above can be accounted for in a different manner. However, the distribution of high tones in forms that are longer than trisyllabic calls such an approach into question.

The tonal analysis states that the language has morphemes that are either specified or not specified as containing underlying tones. A tone reduction rule which also applies between word/stem boundaries as in (7) and (8) prevents adjacent high tones from co-occuring. The analysis predicts that Chácobo prosodic words will allow any combination of H and toneless syllables as long as the H tones are not adjacent. This prediction is born out.

(15) ØØØH
[ha.na.jo.wɨ́]
hana=jó=wɨ́
leave=cmpl=imper
"Leave everything!"

(16) ØØHØ
[ha.na.jó.kɨ]
leave=cmpl=dec:past
"S/he left everything."

(17) ØHØØ
[ha.ná.tsi.ki]
haná=tsi=ki
vomit=now=dec:nonpast
"S/he is vomiting now."

(18) HØØØ
[há.na.tsi.ki]
hana=tsi=ki
"S/he is becoming angry"

(19) HØHØ
[pí.tɨ.kɨ́.kɨ]
pí-tɨk-ɨ́=kɨ
wing-break-itr=dec:past
"S/he broke its wing."

(20) HØØH
[hó.tɨ.kɨ.wɨ́]
hó=tɨkɨ(n)=wɨ́
"Come again!"

(21) ØHØH
[ha.βá.kɨ.rɨ́]
haβá=kɨ=rɨ́
run=dec:past=regret
"It is regrettable that s/he ran."

The only caveat is that ØØØØ sequences are difficult to elicit in isolation. Even if we find a form that is completely toneless, a high tone insertion rule operates on prosodic words that occur before the clause-type morpheme or in isolation resulting in a HØØØ pattern. However, if a final high tone from a verb form is deleted by tone sandhi, such forms can appear.

(22) ØØØØ
[ha.na.ʔa.ka .ɾá.mi .kɨ]
hana-ʔaká ɾámi =kɨ
leave-pass Rami =dec:past
"Rami was left (somewhere)."

Consider how a metrical analysis might capture such patterns. The lexical high tones are independently necessary based on (9) and (10). It seems unavoidable to posit the tone reduction rule HH −> ØH (*or* HLH −> LH if the bitonal L + H analysis is adopted) based on (7) and (8). It is clear that nothing is captured by stipulating that feet are mapped on to the forms above apart from what is already necessary from the tonal analysis. Given the evidence, maintenance of a metrical analysis in Chácobo is equivalent to arguing that metrical structure exists but has no empirical signal in the grammar. Some linguists might prefer to adopt such an assumption. However, given the scope and concerns of this paper which is ostensibly about describing what structures are empirically motivated in Chácobo, we choose not to adopt to describe the results in terms that imply another layer of latent metrical structure.

In this study, we focus on cases like the noun *ínaka* "dog, pet" whose two first syllables are forced to remain toneless when the noun occurs in ergative case: *inaká*. Except for a few brief comments on spectrograms in Tallman (2018), there is no previous literature that deals with the phonetics of those two initial toneless syllables using instrumental studies. However, impressionistic observations suggest that the initial syllable in those cases presents some sort of prominence, sometimes described as a tonal difference (the presence of a mid tone) or as stress prominence (see Shell, 1975; Iggesen, 2006, 2007; Tallman, 2017, 2018). Since we are interested in how the initial prominence is distinguished from tone in general, we will also look at the phonetic correlates of the final high tone itself to see whether there are phonetic differences between the initial prominence and the high tone in Chácobo.

Before moving into the experimental design and results, we must mention that this study contains two types of data analysis: hypothesis confirmation and hypothesis generation (Wickham and Grolemund, 2017; Haig, 2018, *inter alia*). In the realm of hypothesis confirmation, we are interested in testing to see whether there is any evidence for stress-based prominence. The experiment below was designed explicitly to address this issue by investigating words sentence-medially with prosodic patterns where the distinction

https://doi.org/10.1121/10.0001014

between stress and tone is thought to be manifest. In the realm of hypothesis generation, we are interested in describing what the acoustic correlates of these categories might be. We also discuss previously unstated hypotheses that we have developed for Chácobo in light of the phonetic data we gathered. This concerns whether a relatively higher pitch on the "stressed" syllable can really be considered a correlate of such syllables in the context of our study given that such stressed syllables straddle the boundary between a very high clause initial high tone and what could be analyzed as a bitonal L + H* contour mapped to the last two syllables of the words we investigated. The possibility that this analysis could be correct partially confounds the original hypothesis we set out to (dis)confirm concerning the existence of initial stress in Chácobo, at least to the extent that pitch is considered to be a correlate of stress. This is because, in light of the extra high H tone clause-initially and the possibility of a L + H* hypothesis, it is no longer possible to interpret a higher pitch on the first syllable as opposed to the second as indicative of this syllable having a specific pitch target unequivocally. Stated in terms of the impressionistic descriptions we have reviewed above, we cannot yet determine whether a surface mid pitch/tone surfaces on the "stressed" syllables because it is possible that a higher pitch occurs on this syllable compared to the unstressed syllables because of pitch movements from the surrounding tones. It is possible that in other contexts, an increase in pitch on the stressed syllable could be observed, but this seems to be ruled out as a matter of principle in the sentences we gathered in our data because of the high pitch on the first syllable of the sentence (see Sec. X for more discussion). We thus confirm that there are important acoustic correlates for the stress-tone distinction in Chácobo but point out that the data lend themselves to conflicting hypotheses regarding the interpretation of pitch on the first syllable. Since this is the first instrumental study of Chácobo stress and tone, it is to be expected that much of it will be descriptive. Ideally, a phonetic study reporting the results of a production experiment would be more concerned with hypothesis confirmation rather than hypothesis generation. However, this is the first study of the phonetics of any language of the Pano language. Furthermore, taking a more exploratory stance on the data allows us to detect flaws in the research design which will help provide a sounder basis for future research.

IV. LINGUISTIC MATERIALS

We sought to elicit words with three properties. First, the words that we collected were all trisyllabic. We avoided words with fewer syllables in order to examine whether the initial syllable presented any sort of prominence and isolate any interference from the final syllable which bears a high tone. Second, we chose words where the vowels of the first two syllables were the same, either [a] or [o]. Thus, we control the intrinsic prominence vowels could show due to their degree of sonority. By guaranteeing that the two first vowels have the same quality, we can also study whether there is any asymmetric behavior in vowel quality through the examination of their first ($F1$) and second ($F2$) formants. Third, we chose words where the second and third consonants were sonorants (either nasals /n/, /m/, the approximant /β̞/, or the flap /ɾ/). This was done so that we could observe the uninterrupted behavior of $F0$ during the production of the target word. The words shown in (23) were found to meet the criteria above.

(23) Nouns used in this study presented with and without the ergative case marker:

	Citation form	Ergative form	Gloss
a.	kamáno	kamanó	"jaguar"
b.	kanaɾá	kanaɾá	"jabiru (type of bird)"
c.	kaɾáma	kaɾamá	"rubber"
d.	koɾóno	koɾonó	"cloud (usually dark)"
e.	nonomí	nonomí	"muscovy duck"
f.	panána ~ panáɨ̨	pananá ~ panaɨ́	"palm leaf"
g.	panaβ̞í	panaβ̞í	"asai plant"
h.	poɾóma	poɾomá	"axe"
i.	maβ̞aɾá	maβ̞aɾá	"bald person"

We recorded the forms in the second column in (23) in sentential contexts. Thus, we were able to elicit those nouns with the same suprasegmental patterns in similar syntactic contexts where they are in ergative or genitive case. The ergative high tone docks to their final syllable and deletes all of the underlying tones of the citation form. For example, *kamáno* "jaguar" surfaces as *kamanó* because of the presence of the ergative or genitive case markers. Sentences were constructed for semantic coherence so that the relevant words would occur sentence medial; (24) provides an example of one of the sentences recorded. The target word *kamanó* "jaguar (erg.)" appears in bold.

(24) [tʃá.ʂo **ka.ma.nó** kɨ.jó.ɾo.ʔá.kɨ]
tʃáʂo kamáno=´ kɨjó=ɾoʔá= kɨ
deer jaguar=erg finish=limit=dec:pst
"The jaguar only finished the deer"

All target words were elicited with a phonological pattern such that the first and the second syllables (s1 and s2) are toneless and the final third syllable (s3) has a high tone. Target words did not appear as the initial or final word in the phrase to avoid any phonological phenomena that affect words that appear at phrase edges. Speakers were shown a screen on a laptop computer with the relevant sentences in the Chácobo's Spanish-based orthography. They were asked to repeat the sentence three times. Across the five speakers (Sec. V) of this study, we expected to measure a total of 1170 vowels (=3 vowels per phrase × 26 phrases × 3 repetitions × 5 speakers). The full list of sentences recorded is provided in Appendix.

In the process of gathering data, there were occasional malfunctions of our recording equipment, and at times background noise interfered with the recording to such an extent that some utterances were unusable. This is understandable given that the recordings were made in Amazonian communities (Alto Ivon and Riberalta), where the climate

TABLE I. Language consultant information.

Speaker	Age	Gender	Location
NDT	Aprox. 50	F	Community of Cachuelita, close to the Geneshuaya river
GOS	Aprox. 60	M	Currently in Riberalta, originally from the community of Alto Ivon.
PCS	Aprox. 60	M	Community of Alto Ivon.
JAM	Aprox. 60	F	Community of Alto Ivon.
CMO	Aprox. 65	M	Currently in Riberalta, originally from the community of Alto Ivon.

sometimes interferes with the proper functioning of recording equipment and background noise (produced chickens, birds, and children) often disrupts linguistic work. Some recordings were, therefore, lost. Sometimes participants in the study offered additional tokens of utterances without prompting from the researchers. These utterances were invariably instances of sentences from the stimuli. We added these additional tokens to make-up for occasionally lost tokens due to malfunctions in the recording equipment or background noise. The total number of actual vowels measured in our study was 1197 or 399 utterances.

V. SUBJECTS

The speakers were all approximately between the ages of 50 and 70 at the moment of the data collection. They all speak the "central" dialect of Chácobo spoken in or near Alto Ivon (see Tallman, 2018, Chap. 1). Some of them, like GOS, PCS, and CMO are teachers with a higher education and a higher level of bilingualism in Spanish. However, all the speakers interviewed use Chácobo with family and friends on a daily basis. Their information was gathered through a number of conversations Tallman has had with them over the years. See Table I.

VI. MEASUREMENTS

An H4N Zoom recorder was used with mono recording setting and a SM10 head-worn microphone. The audio files were recorded in PCM WAV format in 16 bits with a sampling frequency of 44.1 KHz. We used Praat TextGrids to transcribe the data collected (Boersma and Weenik, 2019). We manually transcribed the three vowels in the target word of each of the 26 phrases recorded per speaker. The beginning and end of each vowel in the target word were determined manually based on changes in intensity and the salience of the vowel formants. We used the Praat script functionality to extract *F*0, *F*1, *F*2, and duration measurements from each of the three vowels in the target word. In order to filter out micro-prosodic phenomena that could affect the *F*0 contour, making the identification of the pitch movements difficult, we used the Praat functions *kill octave jump* and *pitch smooth* (with 10 Hz bandwidth).

In Fig. 2, we show the three syllables, s1, s2, and s3, in the target word *camanó* "jaguar (ergative)" that occurs in the phrase *játsi camanó rësomáquë* ("So the jaguar made him die"). For illustrative purposes, we indicate the different points where measurements were taken in the vowel [o] of the word *camanó* "jaguar (ergative)." Those same measurements were carried out in the vowels of each target word. The capital letters A and B show the beginning and the end of the vowel, respectively. Those two points were used to calculate duration and the mean intensity of each vowel. For extracting *F*0, *F*1, and *F*2 values, we excluded 5% of the vowel duration at each edge of the vowel (the limits of those exclusion areas are represented by the capital letters C and D). This was done to make sure as much as possible that any micro-prosodic effect left between the transition of consonants and vowels would not interfere with the formant measurements.

Table II presents a summary of the measurements extracted from each vowel in the target words with the help of a Praat script.

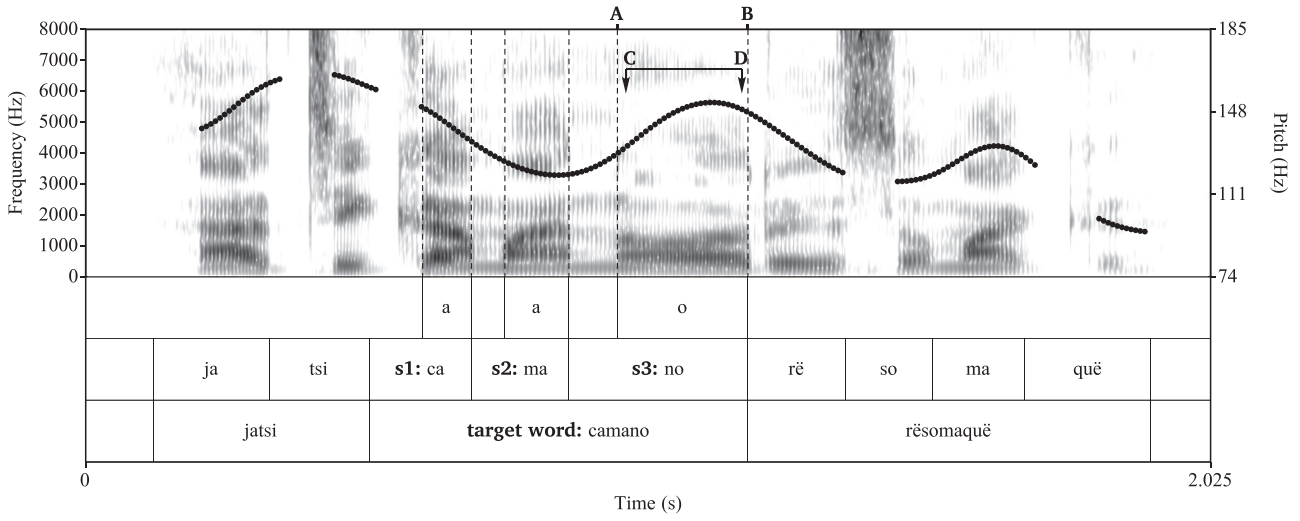

FIG. 2. Segmentation, transcription, and points of measurements.

https://doi.org/10.1121/10.0001014

TABLE II. List of measurements taken in each vowel.

	Measurement Location	Type of measurement
i.	Between points A and B	Vowel duration (s)
ii.	Between points A and B	Intensity mean across the vowel (dB)
iii.	Between points C and D	*F*1 and *F*2 mean across the vowel (Hz)
iv.	Between points C and D	*F*0 mean across the vowel (Hz)

The decision to code duration, intensity, values of *F*0, *F*1, and *F*2 was made because they are frequently noted as being correlates of stress prominence in the literature (e.g., Elias-Ulloa, 2011; Gordon and Roettger, 2017; Himmelmann, 2006; Himmelmann and Ladd, 2008; Ladefoged, 2003; Lieberman, 1960; Sluijter, 1995; Llisterri *et al.*, 2003; Ortega-Llebaria and Prieto, 2010; Roettger and Gordon, 2017). We expected that pitch would be either the only or the most important correlate of tone and that any other type of prominence would be marked by a combination of the variables that were measured.

The transcription of tones was done automatically by the script and then double-checked manually. The pitch movements associated with each vowel in the target words were calculated by the script inspecting the *F*0 contour. If the highest value of *F*0 followed the lowest value during the duration of a vowel, the pitch movement of that vowel was recorded as rising in the database. If the lowest value followed the highest one, then the pitch movement was recorded as falling. Rising pitch movements were coded with an H symbol in the TextGrids. An illustrative spectrogram is provided in Fig. 3. Target words appear in bold. High tones are marked by the symbol H, and L% stands for a final low boundary tone.[4]

The researchers noticed, in contrast to our expectations, that the pitch contour of the initial syllable in the target word typically started high at the beginning of s1 and reached its lowest point at the end of the second vowel in s2. The literature reviewed in Sec. II on Chácobo prosody suggests that there might be an uptick in pitch associated with stress on the first syllable. However, an alternative analysis emerges from the observation of the pitch contours in our data. The apparent uptick in pitch of the first syllable

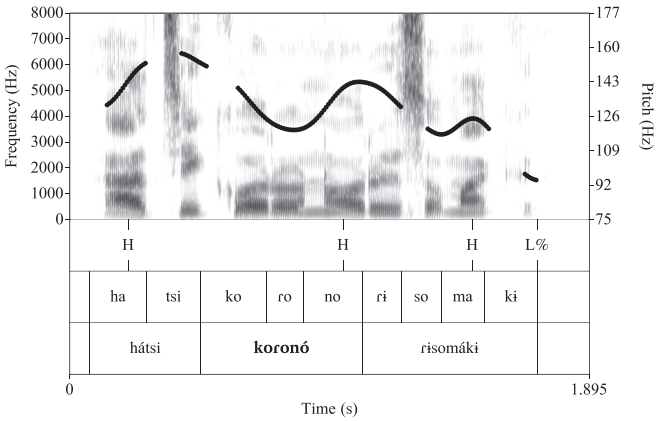

FIG. 3. Pitch contour for [ha.tsi **ko.ɾo.nó** ɾɨ.so.má.kɨ] "The rain cloud made (them) die."

of the target word compared to the second could be the result of *F*0 interpolation between an extra prominent sentence initial high tone that occurs on the first word of the sentence and the high tone that occurs in the second word. This interpretation is possible because in Chácobo, high tones are realized as rising *F*0. In fact, the attested pitch contours suggest that Chácobo high tones could be better described phonologically as L + H* bitonal units and that the apparently higher pitch of the first syllable compared to the second related to an extra high pitch associated with the first high tone of the sentence followed by a gradual declination until the H segment of the L + H* bitonal unit is reached. On this approach, one could argue that s1 is unspecified with respect to tone. Thus, the interpolation between the peak of the high tone in the first word to the valley of the high tone on the second word creates the phonetic effect of a falling pitch during the first syllable of the target word. This is distinct from the Tallman (2018) analysis, which assumes that a low pitch is assigned to all syllables which do not contain an underlying high tone. A surface distinction between syllables that are unspecified with respect to tone and those that contain an underlying L is not made in Tallman (2018). However, from a phonetic perspective, the data we examined in this study are also consistent with an analysis that associates the higher *F*0 on the first syllable of targets words with tonal interpolation rather than with one that posits a tonal unit (surface M tone) as a correlate of stress in s1. However, we emphasize that we did not design our study to investigate this question and thus a definitive conclusion cannot be arrived at concerning the increase in pitch of the second syllable. For complete transparency, we present descriptive and inferential statistics regarding *F*0 on the first syllable. We emphasize that the results cannot be straightforwardly interpreted in light of the plausibility of the bitonal L + H sketch above. We hope that future research will be able to determine whether the increase in pitch of the first syllable vis-à-vis the second is better associated with the surrounding phonological context or with the first syllable having pitch/tonal target.

VII. STATISTICAL DESCRIPTION AND ANALYSIS ON STRESS AND TONE CATEGORIES

Box plots displaying the differences between *F*1, *F*2, *F*0, vowel duration and intensity are provided.[4] The third syllable is removed for *F*1 and *F*2 values because we were not able to control for vowel quality of the third syllable in our stimuli. The data below are presented with the normalized (*z*-score) values that are used for statistical analysis. Figure 4 provides *F*1 values for the vowel /a/. Figure 5 provides *F*1 values for the vowel /o/. Figure 6 provides *F*2 values for the vowel /a/. Figure 7 provides *F*2 values for the vowel /o/. Figure 8 provides values for vowel duration. Figure 9 provides values for intensity (dB). Figure 10 provides values for pitch.

Statistical analyses were conducted to compare the first and second syllables of the target words. Both of syllables, s1 and s2, are toneless phonologically. We were interested in

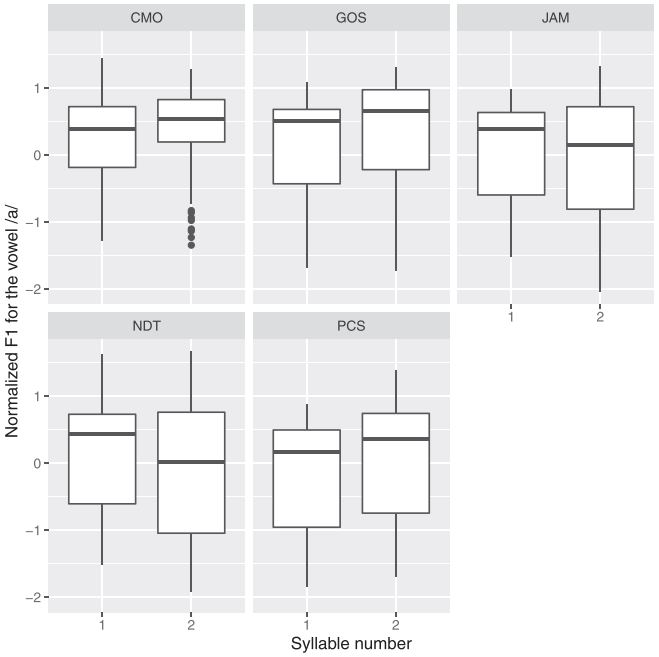

FIG. 4. *F*1 across the five Chácobo speakers by syllable number on trisyllabic nouns with a final high tone for the vowel /a/.

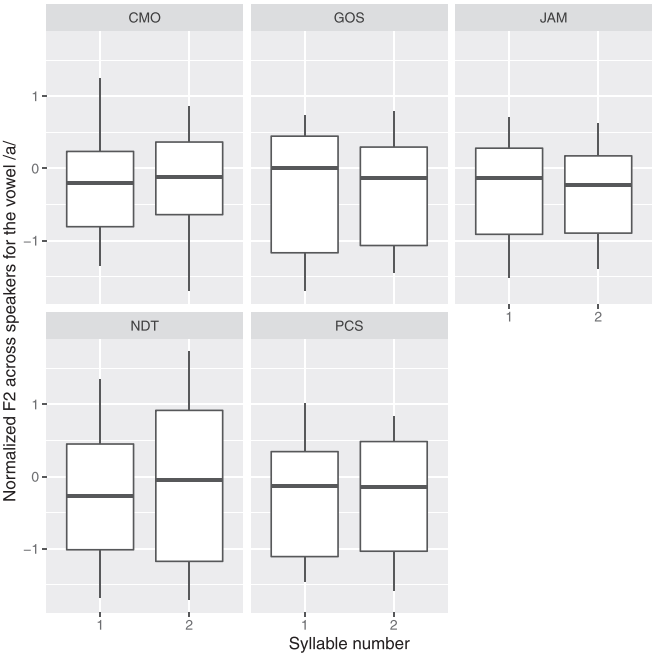

FIG. 6. *F*2 across the five Chácobo speakers by syllable number on trisyllabic nouns with a final high tone for the vowel /a/.

seeing whether s1 is more prominent according to some correlates despite the fact that neither of these syllables contains a high tone. We used multilevel analysis of variance (ANOVA) models using the lmerTest package in R (Kuznetsova *et al.*, 2019) to test for the significance of each of the relevant correlates discussed in Sec. VI. Throughout, speaker and word are used are random effects. We do not present confidence intervals for the effect of speaker because it is not relevant to our research questions at the present time.

In the following discussion, s1 corresponds to the first syllable which is prominent despite being toneless; syllable s2 stands for the second syllable, which is non-prominent (unstressed) and toneless; and s3 stands for the third syllable, which contains a high tone. Where we find a difference between s1 and s2 we refer to this as "initial prominence." We also compare s3 to s1 and s2. This is done to determine the acoustic correlates of syllables with a high tone versus those that are toneless. We used *z*-scores for the statistical

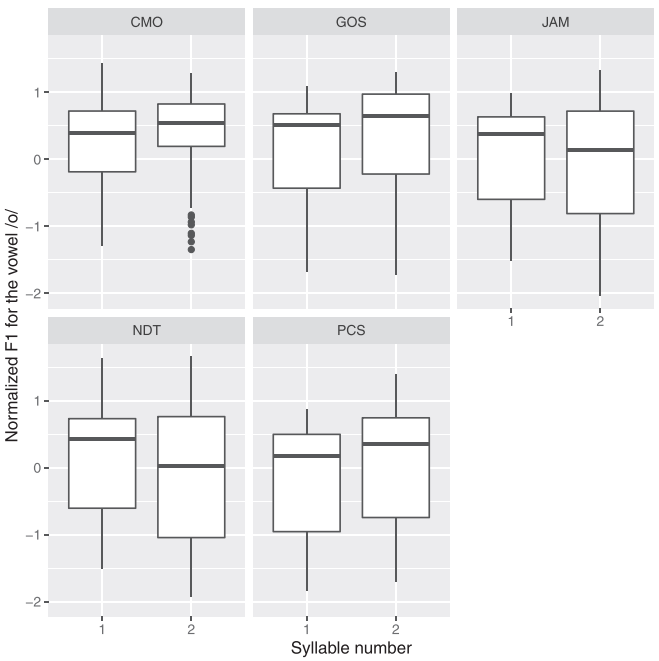

FIG. 5. *F*1 across the five Chácobo speakers by syllable number on trisyllabic nouns with a final high tone for the vowel /o/.

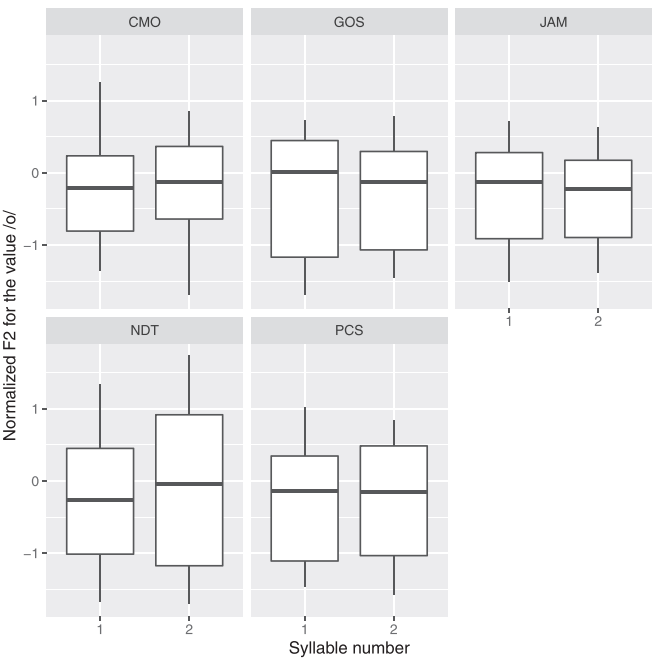

FIG. 7. *F*2 across the five Chácobo speakers by syllable number on trisyllabic nouns with a final high tone for the vowel /a/.

https://doi.org/10.1121/10.0001014

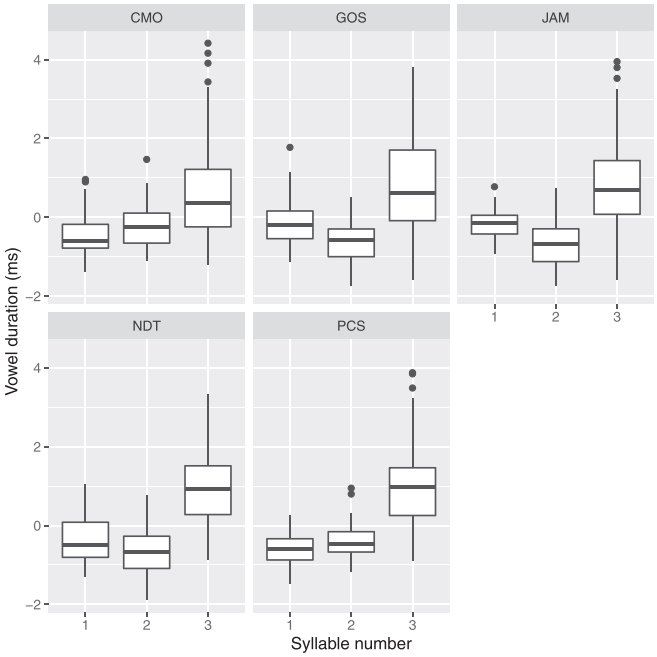

FIG. 8. Vowel duration (s) across the five Chácobo speakers by syllable number on trisyllabic nouns with a final high tone.

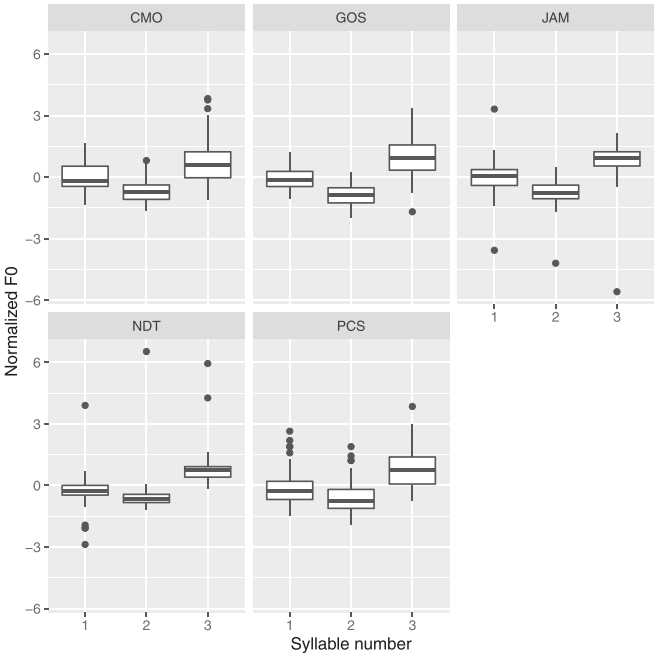

FIG. 10. *F*0 across the five Chácobo speakers by syllable number on trisyllabic nouns with a final high tone.

analyses: all data were normalized within speakers by subtracting the mean values and dividing by the standard deviation for a given speaker for each data point. Finally, Fig. 8 suggests that intensity on the first syllable is not just higher than that of the second, but also higher than the third syllable that contains a high tone, except for one speaker (NDT). We are interested in knowing whether this difference is statistically significant. Based on our understanding of prominence of the first syllable being stress-based, we expect that it will have a higher intensity than the high tone marked syllable as well.

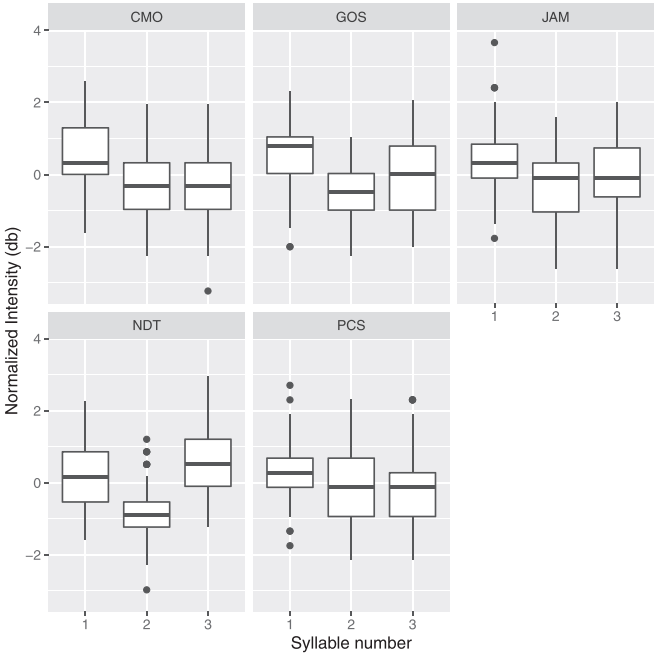

FIG. 9. Intensity across the five Chácobo speakers by syllable number on trisyllabic nouns with a final high tone.

VIII. RESULTS—COMPARISON OF THE FIRST VERSUS THE SECOND SYLLABLE IN TARGET WORDS

A. s1 vs s2: *F*1

The overall mean difference in Hz for *F*1 values between first and second syllables with the vowel /a/ is 7.1 Hz [s1 (694.8826)–s2 (702.0076)]. The overall mean difference in Hz for *F*1 values between first and second syllables with vowel /o/ is 8.9 Hz [s1(535.5547)–s2(526.625)]. A multilevel ANOVA model shows no significant main effect of initial prominence on normalized *F*1 as can be seen from Table III. The data gathered provide no evidence to support *F*1 as a correlate of initial prominence. The full model is reported in Table III.

B. s1 vs s2: *F*2

The overall mean difference in Hz for *F*2 between first and second syllables with the vowel /a/ is –1.9 Hz [s1 (1518.152)–s2(1520.045)]. The overall mean difference in Hz for *F*2 between first and second syllables with the vowel

TABLE III. Type III ANOVA table with Satterhaite's method with *F*1 as the dependent variable.

			95% Confidence intervals	
	F value	Pr (>*F*)	Lower bound	Upper bound
Syllable	0.7989	0.37172	0.02089647	0.153090971
Speaker	14.0439	4.566×10^{-11}		

TABLE IV. Type III ANOVA table with Satterthwaite's method with $F2$ as the dependent variable.

			95% Confidence Interval	
	F value	Pr ($>F$)	Lower bound	Upper bound
Syllable	0.2369	0.6266	−0.03689062	0.06132831
Speaker	6.1348	7.327×10^{-5}		

/o/ is –8.8 Hz [s1(1109.602)–s2(1118.445)]. A multilevel ANOVA model shows no significant main effect of initial prominence on normalized $F2$ as can be seen from Table IV, which displays the results of the model. We conclude that the data gathered provide no evidence to support $F2$ as a correlate of initial prominence.

C. s1 vs s2: *F*0 mean

The overall mean difference in $F0$ between the first and second syllables in the vowel /a/ is 9.5 Hz [s1(153.6932)–s2(144.2159)]. The overall mean difference in $F0$ between first and second syllables in the vowel /o/ is 8.9 Hz [s1(157.2344)–s2(148.3516)]. A multilevel ANOVA shows that a significant mean effect of initial prominence on normalized mean $F0$. The results are displayed in Table V.

However, as mentioned in Sec. VI, it is not clear whether this result can be interpreted as providing evidence that the first syllable is marked with pitch given the possibility that the results are due to declination from an extra-high high tone that appears sentence initially to an L + H bitonal contour on the last two syllables. Thus, while the results from the ANOVA model are suggestive, they do not provide strong evidence that pitch is a correlate of initial prominence in Chácobo because the interpretation is confounded by the possibility that the first syllable has a higher pitch due to the surrounding tonal context. We discuss this further in Sec. XI.

D. s1 vs s2: intensity

The mean difference in intensity between first and second syllables with the vowel /a/ is 2.791667 dB [s1(52.1553)–s2(49.36364)]. The overall mean difference in intensity between first and second syllables with the vowel /o/ is 1.117188 dB [s1(50.24219)–s1(49.125)]. A multilevel ANOVA model shows a significant main effect of initial prominence on normalized intensity (dB) as can be seen from the Table VI. The data we gathered suggests that intensity is a correlate of initial prominence.

TABLE V. Type III ANOVA table with Satterthwaite's method with $F0$ as the dependent variable and the first and second syllables as independent variables. (***: highly statistically significant.)

			95% Confidence Interval	
	F value	Pr ($>F$)	Lower bound	Upper bound
Syllable	135.7217	<0.01***	−0.67422209	−0.48036317
Speaker	0.6631	0.6178		

TABLE VI. Type III ANOVA table with Satterthwaite's method with intensity (dB) as a dependent variable. (***: highly statistically significant.)

			95% Confidence Interval	
	F value	Pr ($>F$)	Lower bound	Upper bound
Syllable	142.1564	<0.01***	−0.8653843	−0.62145912
Speaker	4.9748	0.0005793		

E. s1 vs s2: duration

The overall mean difference in duration between first and second syllables is 2.7 milliseconds with the vowel /a/ [s1(88.30303)–s2(85.64015)]. The overall mean difference in duration between first and second syllables with the vowel /o/ is 14.7 milliseconds [s1(90.73438)–s2(76.0625)]. A multilevel ANOVA model shows a significant main effect on normalized duration (seconds) as can be seen from Table VII, which contains the results of the model.

The boxplots in Fig. 6 suggest that speakers vary in terms of whether they use duration to mark the stress distinction. The mean differences in terms of durations presented in the previous paragraphs are thus somewhat misleading in that they aggregate data from speakers regardless of whether they seem to mark the contrast or not. Table VIII makes the speaker variation clear.

We currently do not have an explanation for the speaker variation observed above and the fact that for two speakers the correlation is in the wrong direction makes this analysis suspect (although see Fig. 8 that contains the normalized values). It is possible that the correlation is spurious, but it is also possible that some speakers mark the contrast with duration and others do not. While multilevel ANOVA model provides weak evidence for the association between vowel duration and prominence, the data from our study are inconclusive with respect to vowel duration.

IX. RESULTS—COMPARISON OF THE FIRST AND SECOND SYLLABLES VERSUS THE THIRD SYLLABLE IN TARGET WORDS

A. s1/s2 vs s3: intensity

The difference between the first two syllables and the last syllables is marginal in terms of intensity. The first two syllables have a mean intensity of 1 dB lower than the third syllable [s1/s2(50.75947)–s3(51.7644)]. A multilevel ANOVA model shows that normalized intensity has no significant main effect on the low/high tone distinction. Results of the model are presented in Table IX.

TABLE VII. Type III ANVOA table with Satterthwaite's method with duration (s) as the dependent variable. (***: highly statistically significant.)

			95% Confidence Intervals	
	F value	Pr ($>F$)	Lower bound	Upper bound
Syllable	28.3333	<0.01***	−0.2515917	−0.1163747368
Speaker	1.7424	0.1387		

https://doi.org/10.1121/10.0001014

TABLE VIII. Mean values of vowel duration across the speakers of this study.

	Mean syllable 1 (ms)	Mean syllable 2 (ms)	Mean difference (ms)
JAM	92.21429	71.53571	20.67858
GOS	95.12987	81.7013	13.42857
NDT	101.4865	89.35135	12.13515
CMO	87.81818	95.25974	−7.44156
PCS	69.7875	76.225	−6.4375

B. s1/s2 vs s3: duration

The overall mean duration of the high tone marked syllable (syllable 3) is 44 ms higher than the mean duration of the syllables which do not contain a high tone (syllables 1 and 2) [s1/s2 (85.80485)–s3(129.7806)]. A multilevel ANOVA model shows that normalized duration has a significant main effect on the high tone - toneless distinction. The results are displayed in Table X. High tone marked syllables are longer than toneless syllables in our production study.

C. s1/s2 vs s3: *F*0 mean

The overall mean *F*0 of the high tone marked syllable (syllable 3) is 21.2 Hz higher than the mean *F*0 of the low tone marked syllables (syllables 1 and 2) [s1/s2(150.2079)–s3 (171.3597)]. A multilevel ANOVA model shows that normalized *F*0 displays a significant main effect on the high tone—toneless distinction as displayed in Table XI.

X. RESULTS: S1 VS S3 ON INTENSITY

Results from Secs. VIII and IX above show that intensity is a correlate of stress, but not of tone. The results of Sec. VIII contrasted the first syllable with the second. In this section, we are interested in testing whether there is a significant difference between s1 and s3. Figure 8 suggests that intensity is higher in the first syllable compared to the third for all speakers except NDT. Cross-linguistic literature on stress suggests that in tonal languages, intensity is a correlate of stress (Gordon and Roettger, 2017, p. 6). While Sec. VIII shows that intensity marks the distinction between the first and second syllable, it is not clear whether the strength of this variable holds when comparing the stressed syllable with the high tone marked syllable.

To apply the multilevel model, words with /i/ and /ɨ/ on the final syllable were removed so that the vowel values across the two syllables were comparable. This resulted in removal of 112 data points containing words where the final syllable was /i/ or /ɨ/ such as panaβí "asai" and pananɨ́ - palm." The first syllable was coded as a having the dummy variable 1 for stress and the third syllable was coded for the dummy variable 0. A multilevel ANOVA with normalized intensity (dB) as the dependent variable, stress as the predictor (where s1 is stressed and s3 is not) with word and speaker as random effects shows that the stress of the first syllable as a predictor is significantly higher than the high tone on the third syllable. The results of the analysis are displayed in Table XII.

The results provide evidence that intensity is significantly higher in the first (stressed) syllable compared to the third (high tone marked) syllable.

TABLE IX. Type III ANOVA table with Satterthwaite's method with intensity as the dependent variable.

			95% Confidence Interval	
	F value	Pr (>*F*)	Lower bound	Upper bound
Syllable	4.1683	0.8629	−0.1053841	0.1257874
Speaker	0.0362	0.9999		

TABLE X. Type III ANOVA table with Satterthwaite's method with duration as the dependent variable. (***: highly statistically significant.)

			95% Confidence Interval	
	F value	Pr (>*F*)	Lower bound	Upper bound
Syllable	762.7634	<0.01***	−0.1053841	0.1257874
Speaker	0.0422	0.9966		

XI. DISCUSSION AND INTERPRETATION OF THE RESULTS

In this section, we summarize and discuss the results. First, we interpret the results as supporting the idea that there is a phonetic distinction between stress and tone in Chácobo. Then we consider the possibility that the results could be understood as contrasting primary and secondary stress rather than a tone-stress distinction.

The results above report descriptive and inferential statistics concerning the phonetic correlates of prominence in the syllables of trisyllabic words that are marked with a final high tone. Based on previous literature, we expected that the first syllable would be prominent, but we also tested for the phonetic correlates of the high tone marked syllable, since we are interested in whether there are differences in the way stress and tone, are marked in Chácobo on the assumption that word initial prominence can be regarded as stress. The results presented above provide evidence that the first syllable is prominent compared to the second and third syllables with respect to intensity. There is a statistically significant difference between syllable 1 and syllable 2 with respect to intensity. There is a weaker but still statistically significant difference between syllable 1 and syllable 3 in intensity across the five speakers. Note that this model included the speaker where the difference in intensity could not be

TABLE XI. Type III ANOVA table with Satterthwaite's method with *F*0 as the dependent variable and the first and second syllables versus the third syllable as independent variables. (***: highly statistically significant.)

			95% Confidence Interval	
	F value	Pr (>*F*)	Lower bound	Upper bound
Syllable	588.2310	<0.01***	−1.318323798	−1.1214154
Speaker	0.0138	0.9996		

TABLE XII. Type III ANOVA table with Satterthwaite"s method with normalized intensity. (*: statistically significant.)

			95% Confidence Interval	
	F value	Pr ($>F$)	Lower bound	Upper bound
Stress (s1 vs s3)	4.0451	0.0447*	0.003795009	0.2911929
Speaker	1.6372	0.1631		

observed. The first syllable appears to be longer than the second syllable overall, but the results do not provide convincing evidence for duration being a marker of initial stress because only three of the five speakers have mean values of duration that are higher than the mean vowels for duration of the second syllable. Pitch is also higher on the first syllable compared to the second. However, the higher pitch on the first syllable compared to the second could also be explained by intonational context since the initial syllable plausibly falls between a clause initial high tone and bitonal L + H segment. Unfortunately, our study was not able to find any evidence for or against the proposal that stress in Chácobo is marked by a level of pitch intermediate between high tone and low pitch could be tested. On the other hand, $F1$ and $F2$ do not show any unexpected change on the first or second syllable of those Chácobo words. This is perhaps surprising given that the language only contains four phonemically distinct vowels and distinctions in quality could be profitably used in marking prominence contrasts.

The final high tone marked syllable is distinguished from the other two syllables by pitch and secondarily by duration. Intensity is not a correlate of high tone marked syllables. Not all languages use vowel duration to mark lexical pitch (e.g., Pierrehumbert and Beckman, 1988, for Japanese; Hualde *et al.*, 2000 and Hualde *et al.*, 2002 for Basque). However, in Uspantenko (Mayan) high tone is marked through a combination of higher pitch and vowel duration (Bennet *et al.*, 2019). If the descriptive labels applied to the different types of prominence found on the first syllable and the third syllable from Tallman (2018) are adopted, then this study has found evidence for another language where tone and vowel duration are correlates of high tone marked syllables.

The results of this study are not surprising from a typological perspective. In their typological overview of the phonetics of tone and stress, Gordon and Roettger (2017) note that there is a correlation between stress being encoded by intensity and the presence of a tone system for a given language. They state, "It is likely no coincidence that six of the seven tone languages in the database for which an overall measure of intensity was taken (all except Thai) employed intensity as a marker of stress" (Gordon and Roettger, 2017, p. 6). Chácobo, therefore, fits well into the typological pattern in the sense that its stressed syllable is marked with intensity. The interpretation whereby the first syllable marks stress and the third marks high tone is also in line with what several other researchers of Pano languages have been impressionistically suggesting; that is, that in some Panoan languages tone and stress coexist (Zariquiey, 2011; Tallman, 2018; Neely, 2019). This article provides the first acoustic evidence supporting such an analysis.

A reader who is skeptical of the labels applied in the description of some Pano languages might wonder why the distinction between the first syllable and the third syllable cannot be understood as a distinction between secondary and primary stress rather than a distinction between stress and high tone. It is not clear that the phonetic results support a secondary versus primary stress analysis for the data of our study. Note that higher vowel duration marks off the third syllable as distinct from the others. But vowel duration does not clearly distinguish between the first and the second syllable. Presumably, this should mean that the third syllable bears the primary stress. However, under this interpretation, there would be no way of explaining the fact that the intensity on the first syllable is higher than that on the third. The phonetic evidence thus suggests that the categories are of different types not manifestations of different levels of the same category.[5] However, we do not discard completely the possibility that the final syllable of the word could be regarded as containing a primary stress and a high tone, the phonological facts reviewed in Sec. III notwithstanding. It is possible that the initial intensity is phrase-level intensity (see Sec. X) or a correlate of some other prosodic boundary. Perhaps the duration on the final syllable is not a correlate of high tone as described in Bennet *et al.* (2019) but rather indicates primary stress. Ideally adopting such an analysis would also entail finding independent evidence for an underlying stress that is not already accounted for our initial prominence and the fact that high tones are present underlyingly.[6] Currently, no such evidence is available.

XII. LIMITATIONS OF THIS STUDY

This study has a number of limitations that we mention here. The first limitation of this study is that the low number of speakers makes it hard to come to general conclusions about the population of the whole (Ladefoged, 1997, p. 140; Roettger and Gordon, 2017, p. 6). This is particularly troublesome in the context of vowel duration where certain speakers seem to be marking the first syllable as longer than the second while other speakers do not. Future research, with a larger sample of speakers, may reveal that the statistical model which indicates that their duration is a statistically significant correlate of stress, may turn out to be spurious. Another limitation to the study which we have highlighted throughout is that we were unable to come to conclusions regarding whether pitch is a correlate of initial prominence because of a confounding intonational contour from the beginning of the utterance to a low pitch target on the second syllable. Future research will seek to investigate words in intonational contexts where an increase in pitch of the first syllable relative to the second could be observed if it were present.

An obvious limitation to this study is that it does not contain a complementary perception experiment to test whether and the extent to which the results from production are perceptually available. We do not know, for instance, if

https://doi.org/10.1121/10.0001014

the difference in intensity marked on the first syllable is perceptually salient to the speakers of the language, even if the results from our study suggest it would be.

Another potential limitation is that our experiment did not adequately distinguish between word and phrase level intensity (see Roettger and Gordon, 2017, p. 4). While we controlled for the possibility that intensity could be sentence initial by creating stimuli where our target words were sentence-medial, future research should test whether the intensity effect is present for words that are phrase internal [e.g. *toa kamanó* "that jaguar (erg.)" rather than *kamanó* "jaguar (erg.)" as done in this study].

A final limitation of this study is that we only investigated words with a single prosodic shape. Ideally, the generalizations that can be inferred from the results of the experiment extend to stress and high tone across all contexts, but future research may show that this is not the case.

XIII. CONCLUSION

This paper provides the first instrumental study of the phonetics of suprasegmental categories in Chácobo, a southern Pano language of the northern Bolivian Amazon. We provided an analysis of production data gathered from five Chácobo speakers. We found that in words whose two initial syllables are toneless, the initial syllable shows a sudden increase in intensity, which points out the co-existence of stress and tone categories in Chácobo. High tone is marked by vowel higher pitch whereas stress is marked by intensity. This is in keeping with typological findings of the phonetic correlates of stress in other tonal languages (Gordon and Roettger, 2017, p. 6). The finding that high tone is marked by high pitch and vowel duration is typologically attested as well (Bennet *et al.*, 2019), but it is unclear whether this pattern can be thought to be common or rare cross-linguistically. The co-existence of stress and tone in Pano languages has been suggested by Panoanists based on impressionistic observations but never supported by instrumental nor statistical analysis. Future research will investigate whether the finding presented in this paper is also reported in other environments and whether the analysis is confirmed through perception experiments. Finally, this study highlights the need for carrying out more instrumental work documenting different phonetic aspects of indigenous languages that can help gain a better understanding of their sound systems beyond what can be provided by impressionistic observations.

APPENDIX: LIST OF STIMULI USED IN THE STUDY

TABLE XIII. List of stimuli used in the study.

	Chácobo orthography	Phonemic representation	Local Spanish translation	English translation
1	cháxo camanó quëyóroháquë	tʃáʂo kamanó kɨjóɾoʔákɨ	El tigre lo mató al huaso no más.	The jaguar only finished (eating) the deer.
2	máquë canará quëyóroháquë	mákɨ kanaɾá kɨjóɾoʔákɨ	El jabiru terminó la palometa no más.	The jabiru only finished (eating) the piranha.
3	jẽ́në caramá yoshíhuayóquë	hɨ́nɨ kaɾamá joʃíwajókɨ	La goma de siringa fregó todo el agua.	The rubber completely corrupted (poluted) the water.
4	huási coronó mapároháquë	wási koɾonó mapároʔákɨ	La neblina tapó la pampa no más.	The clouds covered just the savannah / grass.
5	xẽ́na nonomí quëyóroháquë	ʂɨ́na nonomí kɨjóroʔákɨ	El pato simurani terminó los gusanos no más.	The muscovy duck just finished (eating) the worms.
6	xóbo pananẽ́ mapároháquë	ʂóβo pananɨ́ mapáɾoʔákɨ	El palmito cubrió la casa no más	The palm tree just covered the house.
7	xóbo panabí mapároháquë	ʂóβo panaβɨ́ mapáɾoʔákɨ	La planta de asaí cubrió la casa no más.	The asaí plants covered only the house.
8	jíhui poromá tecároháquë	híwi poɾomá tɨkáɾoʔákɨ	El hacha quebró el palo no más.	The axe only broke the stick.
9	játsi camanó rësomáquë	hatsi kamanó ɾɨsomákɨ	Entonces el tigre le hizo morir.	So the jaguar made him die.
10	játsi canará rësomáquë	hátsi kanaɾá ɾɨsomákɨ	Entonces el jabiru le hizo morir.	So the jabiru made him die.
11	játsi caramá rësomáquë	hátsi kaɾamá ɾɨsomákɨ	Entonces la goma le hizo morir.	So the rubber made him die.
12	játsi coronó rësomáquë	hátsi koɾonó ɾɨsomákɨ	Entonces la neblina le hizo morir.	So the cloud made him die.
13	játsi nonomí rësomáquë	hátsi nonomí ɾɨsomákɨ	Entonces el pato simunari le hizo morir.	So the muscovy duck made him die.
14	játsi pananá rësomáquë	hátsi pananá ɾɨsomákɨ	Entonces el palmito le hizo morir.	So the palm tree made him die.
15	játsi panabí rësomáquë	hátsi panaβɨ́ ɾɨsomákɨ	Entonces la panabi le hizo morir.	So the asai made him die.
16	játsi poromá rësomáquë	hátsi poɾomá ɾɨsomákɨ	Entonces la hacha le hizo morir.	So the axe made him die.
17	játsi joní mabará rësomáquë	hátsi honí maβaɾá ɾɨsomákɨ	Entonces el hombre calvo le hizo morir.	So the bald man made him die.
18	játsi camanó baquẽ́ rësoníquë	hátsi kamanó βakɨ́ ɾɨsoníkɨ	Entonces el hijo del tigre le hizo morir.	So the child of the jaguar made him die.
19	játsi canará baquẽ́ rësoníquë	hátsi kanaɾá βakɨ́ ɾɨsoníkɨ	Entonces el hijo de jabiru le hizo morir.	So the jabiru made him die.
20	játsi caramá ibó rësomáquë	hátsi kaɾamá iβó ɾɨsomákɨ	Entonces el dueño de la goma le hizo morir.	So the owner of the rubber made him die.
21	játsi coronó ibó rësomáquë	hatsi koɾonó iβó ɾɨsomákɨ	Entonces el dueño de la neblina le hizo morir.	So the owner of the clouds made him die.
22	játsi nonomí baquẽ́ rësoníquë	hátsi nonomí βakɨ́ ɾɨsoníkɨ	Entonces el hijo del pato simunari le hizo morir.	So the child of the muscovy duck made him die.
23	játsi pananẽ́ ibó rësomáquë	hátsi pananɨ́ iβó ɾɨsomákɨ	Entonces el dueño del palmito le hizo morir.	So the owner of the palm tree made him die.
24	játsi panabí ibó rësomáquë	hátsi panabí iβó ɾɨsomákɨ	Entonces el dueño de la panabi le hizo morir.	So the owner of the asai made him die.
25	játsi poromá ibó rësomáquë	hátsi poɾomá ɾɨsomákɨ	Entonces el hacha le hizo morir.	So the axe made him die.
26	játsi jóni mabará baquẽ́ rësoníquë	hátsi hóni maβaɾá βakɨ́ ɾɨsoníkɨ	Entonces el hijo del hombre calvo le hizo morir.	So the child of the bald man made him die.

[1]Note that if the underlying high tones are in fact bitonal L+H units, then the tone reduction is slightly different (H.LH > L.H as opposed to H.H > L.H).

[2]An anonymous reviewer implies that they think that the tone erasure rule provides evidence for metrical structure in the language and that possibly there is no evidence for underlying high tones at all. On this account, a Chácobo form such as *ínaka* would have an underlying stress as in /ˈinaka/. A high tone maps onto the first syllable in the default form. The ergative form maps stress on to the final syllable resulting in the form /ˈinaˈka/. The underlying stress marked syllables are then somehow realized as secondary stress on the first syllable and primary stress on the second syllable. After this, a bitonal L+H unit maps onto an iambic foot parsed on the right edge of the word. However, if high tones are post-lexical and then insert on stressed syllables, it is unclear how to derive the form in (5) which contains two high tones. Chácobo forms with multiple high tones are not hard to come by in the grammar (see Tallman, 2018) as long as more than one high tone is present underlyingly. It is also unclear how to account for the difference between (9) and (10), where high tone surfaces on *rámi* but not on *raβi*. On the tone based account *rámi* contains an underlying high tone where *raβi* does not. However, the stress-based account predicts that either both of these forms should contain a high tone or neither of them should. Furthermore, adopting a metrical rule that deletes a prior high tone just in case the final syllable indexes an ergative, but not otherwise, is just as powerful (if not more) than simply adopting a grammatical tone sandhi rule. More discussion concerning the possibility of accounting for all the phenomena with a metrical analysis is provided in Secs. III and XI.

[3]Both sources describe high tone as an underlying high tone, but Tallman (2017) provides no description of how it interacts with lexical high tones.

[4]See supplemental material at http://dx.doi.org/10.1121/10.0001014 for the recordings, textgrids, and script used in this study, and database and R code used for visualizations and statistical analyses.

[5]A reviewer suggests a hybrid analysis where the trisyllabic forms have secondary stress on the first syllable and primary stress on the third, and that a high tone is mapped to the final syllable or else a bitonal L+H* unit is mapped to a bisyllabic foot on the right edge. Trisyllabic forms in our study are parsed as (ˈσ)(σˈσ) This analysis still fails to predict that intensity is higher on the first syllable. It is also unclear why under such an analysis vowel length is not a stronger correlate of secondary stress given that it is parsed as a single foot, whereas vowel duration is a correlate of the final syllable. Apart from the phonetic issues with such an analysis, from a phonological perspective, the analysis simply adds a derivation that accounts for no additional facts about the phonology apart from what is already described in Tallman (2018). As we stated in Sec. III, the fact that one can imagine feet parsed over LH or hearing level (HL) sequences does not automatically provide support for a metrical analysis of the language.

[6]Maintaining that high tone marked syllables *also* contain an underlying stress would mean adopting an extra category in the description of the language (a difference between secondary and primary stress) for which there are no phonological or phonetic consequences.

Bennet, R., Henderson, R., and Harvey, M. (**2019**). "Tono y entonación en Uspanteko" ("Tone and intonation in Uspanteko"), in *Proceedings of the Conference of Indigenous Languages of Latin America IX*, October 10–12, Austin, TX.

Boersma, P., and Weenink, D. (**2019**). "Praat: Doing phonetics by computer, version 6.1.05 [computer program]," retrieved 4 November 2019 from http://www.praat.org/.

de Lacy, P. (**2002**). "The interaction of tone and stress in Optimality Theory," Phonology **19**(1), 1–32.

Elías-Ulloa, J. (**2011**). *Una Documentación Acústica de la Lengua Shipibo-Conibo (Pano): Con un Bosquejo Fonológico* (An acoustic documentation of the Shipibo-Conibo language (Pano): with a phonological outline) (Fondo Editorial de la Pontificia Universidad Católica de Perú, Lima, Peru).

Gordon, M. K. (**2016**). *Phonological Typology* (Oxford University Press, Oxford, UK).

Gordon, M., and Roettger, T. (**2017**). "Acoustic correlates of word stress: A cross-linguistic survey," Ling. Vanguard **3**(1), 1–11.

Haig, B. D. (**2018**). *The Philosophy of Quantitative Methods: Understanding Statistics* (Oxford University Press, Oxford, UK).

Himmelmann, N. P. (**2006**). "Prosody in language documentation," in *Essentials of Language Documentation*, edited by J. Gippert, N. P. Himmelmann, and U. Mosel (De Gruyter Mouton, Berlin, Germany), pp. 163–182.

Himmelmann, N. P., and Ladd, D. R. (**2008**). "Prosodic description: An introduction for fieldworkers," Lang. Doc. Conserv. **2**(2), 244–274.

Hualde, J. I., Gaminde, G. E., and Smiljanić, R. (**2002**). "From pitch-accent to stress-accent in Basque," in *Laboratory Phonology 7*, edited by C. Gussenhoven and N. Warner (De Gruyter Mouton, Berlin, Germany), pp. 548–584.

Hualde, J. I., Smiljanić, R., and Cole, J. (**2000**). "One the accented/unaccented distinction in Western Basque and the typology of accentual systems," in *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Aspect*, February 18–21, Berkeley, CA, pp. 133–144.

Hyman, L. M. (**2006**). "Word-prosodic typology," Phonology **23**, 225–257.

Hyman, L. M. (**2009**). "How (not) to do phonological typology: The case of pitch-accent," Lang. Sciences **31**, 213–238.

Iggesen, O. A. (**2006**). "Case morphosyntax in Chacobo: Multiple alignment splits and suprasegmental marking," RCLT Seminar, Victoria, BC.

Iggesen, O. A. (**2007**). "The great killjoy of Chacobo case marking: How tone sandhi undermines suprasegmental morphology," RCLT Internal Seminar, Victoria, BC.

Iggesen, O., and Pilar, M. V. (**2007**). "El desarrollo de un marcador suprasegmental," in *Lenguas indígenas de América del Sur. Estudios descriptivo-tipológicos y sus contribuciones para la linguística teórica* [Indigenous languages of South America. Descriptive-typological studies and theoretical contributions], edited by A. Romero-Figuero, A. Fernández Garay, and A. Corberi Mori (Universidad Catolica Andres Bello, Caracas), pp. 187–199.

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., and Jensen, S. P. (**2019**). "Tests in Linear Mixed Effects Models, R package [computer program]," https://github.com/runehaubo/lmerTestR (Last viewed 2020-01-21).

Ladefoged, P. (**1997**). "Instrumental techniques for linguistic phonetic fieldwork," in *The Handbook of Phonetic Sciences*, edited by W. Hardcastle and J. Laver (Blackwell, Oxford), pp. 137–166.

Ladefoged, P. (**2003**). *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques* (Blackwell, Malden, MA).

Lieberman, P. (**1960**). "Some acoustic correlates of word stress in American English," J. Acoust. Soc. Am. **33**, 451–454.

Llisterri, J., Machuca, M., de la Mota, C., Riera, M., and Rios, A. (**2003**). "The perception of lexical stress in Spanish," in *Proceedings of the 15th International Congress of Phonetic Sciences*, August 3–9, Barcelona, Spain.

Neely, K. (**2019**). "The linguistic expression of affective stance in Yaminawa (Pano, Peru)," Ph.D. thesis, University of Berkeley, California, Berkeley, CA.

Ortega-Llebaria, M., and Prieto, P. (**2010**). "Acoustic correlates of stress in central Catalan and Castilian Spanish," Lang. Speech **54**(1), 73–97.

Pierrehumbert, J., and Beckman, M. (**1988**). *Japanese Tone Structure* (MIT Press, Cambridge, MA).

Prost, G. (**1960**). "Fonemas de la lengua Chacobo" ("Phonemes of the Chácobo language"), in *Notas lingüísticas de Bolivia (Phonemes of the Chacobo language); No. 2* (Publications of the Summer Institute of Linguistics in collaboration with the Department of Farming Affairs, La Paz, Bolivia, La Paz, Bolivia).

Prost, G. (**1967**). "Chacobo," in *Bolivian Indian Grammars 1* (Summer Institute of Linguistics, Dallas, TX).

Roettger, T., and Gordon, M. (**2017**). "Methodological issues in the study of word stress correlates," Ling. Vanguard **2017**, 1–11.

Rolle, N. R. (**2018**). "Grammatical tone: Typology and theory," Ph.D. thesis, University of Berkeley, California, Berkeley, CA.

Shell, O. A. (**1975**). *Estudios Panos III: Las Lenguas Pano y su Reconstrucción* (Panoan languages and their reconstruction) (Instituto Lingüístico de Verano, Lima, Peru).

Sluijter, A. (**1995**). *Phonetic Correlats of Stress and Accent* (Holland Institute of Generative Linguistics, Dordrecht, the Netherlands).

Tallman, A. J. R. (**2017**). "A metrical analysis of nouns in Chácobo (Pano)," Amerindia: Estudios Sincrónicos y Diacrónicos Sobre Lenguas Pano y Takana **39**(1), 105–128.

Tallman, A. J. R. (**2018**). "A grammar of Chácobo, a southern Pano language of the northern Bolivian Amazon," Ph.D. thesis, University of Texas at Austin, Austin, TX.

Wickham, H., and Grolemund, G. (**2017**). *R for Data Science* (O'Reilly, Sebastopol, CA).

Zariquiey, R. (**2011**). "A grammar of Kashibo-Kakataibo," Ph.D. dissertation, La Trobe University.