CHAPTER 3

Prosodic and morphological focus marking in Ixcatec (Otomanguean)

Evangelia Adamou\textsuperscript{i}, Matthew Gordon\textsuperscript{ii} and Stefan Th. Gries\textsuperscript{ii}
\textsuperscript{i}CNRS / \textsuperscript{ii}University of California at Santa Barbara

This paper presents the first description of the expression of focus in Ixcatec, a nearly extinct language of Mexico. The study is based on experimental tasks carried out with the last three fluent speakers of Ixcatec. Prosodic analysis shows that in Ixcatec, a language with three lexical tones, contrastive focus is associated with raised F0, lack of focus is marked through lowered F0 and decreased duration, and corrective focus is signaled through various speaker-specific means. Finally, this study shows that morphological and phonetic properties display a complex interaction that contradicts the view that focus may be conveyed through either morphological or phonetic exponents but not both.

1. Introduction

The aim of this paper is to describe how focus is expressed in Ixcatec, a nearly extinct Otomanguean language spoken in Mexico (State of Oaxaca). Results of the analysis, which represents to the best of our knowledge the only phonetic study of focus in any Otomanguean language, show that Ixcatec, a language with three contrastive lexical tones (high, mid, and low), makes use of prosody to express focus, as observed in other tone languages such as Chinese (Xu 1999; Chen & Gussenhoven 2008; Chen 2010), Yongning Na and Vietnamese (Michaud & Brunelle 2016), and Dane-zaa (i.e. Beaver, Schwiertz 2009).

Moreover, Ixcatec combines a focus marker with prosodic marking, unlike other documented “particle languages” (Büring 2009) that rely on a specialized focus marker (i.e. a focus marker with no additional meaning, whether restrictive, additive or scalar) \textit{rather than} prosody to express focus, including Navajo (McDonough 2003), Chickasaw (Gordon 2007), Gürüntüm (Hartmann & Zimmerman 2009), and Bole and Fon (Fiedler et al. 2010).
The paper is organized as follows. In 2, we give an overview of how focus is expressed cross-linguistically. We then provide information on Ixcatec in 3. Prosody and its interaction with morphological marking of focus are presented in 4. We conclude with a summary and discussion of the results in 5.

2. Focus cross-linguistically

Following the theory of focus elaborated by Rooth (1992), a widely-accepted definition of focus refers to constituents introducing alternatives into the discourse. Cross-linguistically focus may be expressed through syntactic, morphological, and/or prosodic means. A central component of the prosodic marking of focus is increased prominence, which is variably realized across languages through manipulation of F0 (and its perceptual correlate pitch), duration and/or intensity. However, studies on individual languages have shown that such prominence is not necessary, as is the case in Wolof, a Niger-Congo language that uses no prosodic marking for focus (Rialland & Robert 2001). On a theoretical level, the use of prosodic means, especially intonation, to express focus has been questioned for tone languages (Cruttenden 1986), in which fundamental frequency – the acoustic dimension along which both tone and intonation are manifested – serves the crucial functional role of conveying lexical contrasts in tone. In support of this view, intonation seems not to be used for focus marking at least in Tamang (Tibeto-Burman, Mazaudon 2003), Yucatec Maya (Mayan, Kügler & Skopeteas 2007), and Gürüntüm (Chadic, Hartmann & Zimmerman 2009). Nevertheless, even if fundamental frequency is less available as a phonetic exponent of focus marking in a tone language, there are other prosodic means that could potentially be exploited to cue focus in a language with lexical tone. Duration thus serves to signal focus in many languages, both non-tonal languages such as English (e.g. Turk & Sawusch 1997; Katz & Selkirk 2011) and Mexican Spanish (Kim & Avelino 2003; De la Mota, Butragueño & Prieto 2010), as well as tonal languages such as Standard Chinese (Chen & Gussenhoven 2008). Furthermore, a growing body of work shows that F0 range manipulation is used even in tone languages for the expression of focus. Xu (1999), for example, shows that in Mandarin Chinese, a language with four contrastive tones, the F0 contour of the focused constituent (in situ) is expanded and the F0 profile of the remaining utterance is compressed (also see Chen & Gussenhoven 2008, Chen 2010 for further evidence and discussion). Similar strategies have been described for the tone languages Yongning Na (Sino-Tibetan) and Vietnamese (Austroasiatic) in Michaud & Brunelle (2016), and Dane-zaa (Athabaskan) in Schwiertz (2009). Hyman (1999) surveys focus mechanisms in Bantu languages, showing that tone
may be influenced indirectly through shifts in prosodic phrasing employed in focus marking. Prosodic rephrasing and suspension of downdrift are thus used to cue narrow focus in the tone Bantu language Nkhotakota Chichewa (Kanerva 1990; Downing, Mtenje & Pompino-Marschall 2004).

Syntactic marking of focus is also widespread across the languages of the world. It has been shown that in several languages a change of the canonical word order is required for an item to be focused. For example in Spanish, a non-rigid, verb-medial language, clause-final position is preferred for focus due to its intrinsic prosodic prominence (Zubizarreta 1998). Such languages fall under the category of “edge languages” in Büring’s (2009) tentative typology, together with languages that mark focus in a position close to the edge. This is the case for verb-final languages such as Turkish, which has a specific, preverbal, focus position (Erguvanlı 1984).

A theoretical debate to which Ixcatec potentially contributes is the extent to which languages employing a specialized focus marker (i.e. a morpheme with no additional scalar, restrictive or additive meaning) also redundantly use prosodic means to express focus. It has been observed that languages with specialized focus markers characteristically do not rely on prosody to signal focus, as in Navajo (Athabaskan, McDonough 2003), Chickasaw (Muskogean, Gordon 2007), Gùrùntùm (Chadic, Hartmann & Zimmerman 2009), Bole (Chadic) and Fon (Gbe) (Fiedler et al. 2010). Examples of languages such as Moroccan Arabic (Benkirane 1998) with both prosodic and morphological marking of focus are rare in the literature.

Büring (2009) proposes the category of “particle languages” for languages that make use of a specialized focus marker with no additional prosodic marking. In order to account for the fact that prosodic marking of focus is rarely encountered in particle languages, two analyses are suggested: “It seems straightforward to analyze the focus morpheme as a direct spell-out of the syntactic feature F […]. Alternatively, one could hypothesize that the focus morpheme marks prominence of prosodic units” (Büring 2009: 201). Similarly, Féry (2013) considers that “a focus marker often has an additional prosodic role: it delimits the focus, even if it does not have another prosodic correlate like a boundary tone or duration” (Féry 2013: 720). In evaluating the hypothesis of mutual exclusivity of morphological and prosodic marking of focus, it is important to consider evidence from more “particle languages”, such as Ixcatec.
3. Some background on Ixcatec

ʃʰwa²ni³ is an Otomanguean language better known in the literature under the name of Nahuatl origin, Ixcatec (ISO code: ixc). Ixcatec belongs to the Popolocan branch, alongside Chocholtec (or Ngiba and Ngigua), Popoloc (or Ngiga), and Mazatec (see map in Figure 1).

Ixcatec¹ is only spoken in the municipality of Santa María Ixcatlán in the State of Oaxaca, in Mexico. Today the village has some 400 inhabitants but at the time of the arrival of the Spaniards in 1522 it was an important center for the Mixteca zone with an estimated population of 10,000 people (Hironymous 2007).

Unlike other Popolocan languages, such as Mazatec and Popoloc, which are spoken by several thousands of speakers, Ixcatec is nearly extinct. There are only ten identified Ixcatec speakers; most of them are in their 80s, and only four of them are fluent in Ixcatec. Although Ixcatec has been in contact with the socially and economically dominant language of Spanish for hundreds of years, the current moribund status of the language is a relatively recent development attributed to a rapid shift to Spanish beginning in the early 20th century.

Previous linguistic research on Ixcatec consists of the phonology and dictionary of Fernández de Miranda (1959, 1961) and the work of Veerman-Leichsenring on pronouns (Veerman-Leichsenring 2000), noun phrases (Veerman-Leichsenring 2001a), and word order (Veerman-Leichsenring 2001b). More recently, work has been done on Ixcatec phonetics and phonology (Alarcón Montero 2010; DiCanio 2011, 2012) as well as on syntax and morphology (Adamou & Costaouec 2013; Adamou 2014; Adamou 2017), codeswitching and language contact (Adamou 2016), and spatial language and cognition (Adamou 2017).

Ixcatec is a verb-initial language with accusative alignment. It makes a clear distinction between nouns and verbs. Several adjectives can be used as non-verbal predicates and there are existential/attributional, locative, and possessive predicates that are distinct from verbs in that they do not carry any person morphemes.

Ixcatec’s phonology is complex and not yet well understood. It has five vowels, which may be oral or nasalized. Depending on the analysis of consonant clusters, it has 24 or 52 consonants, with pre- and post-aspirated as well as pre- and post-glottalized consonants (Fernández de Miranda 1959; Alarcón Montero 2010; DiCanio 2011, 2012). Syllables are all open but may contain a complex onset.

Ixcatec has three tones that contrast in word-final syllables: high (H), mid (M), and low (L). In non-final syllables, the contrast between mid and low tone is neutralized in favor of the mid tone (see DiCanio 2011 for a preliminary phonetic analysis of Ixcatec tone). Stress in Ixcatec typically falls on the penultimate syllable and is associated with increased duration and intensity and higher F0 (DiCanio 2012).

¹ Ixcatec is not to be confused with the Mazatec variety of San Pedro Ixcatlán.
4. The expression of focus through prosodic and morphological means

4.1 Predictions

In keeping with results from other languages discussed in Section 2, we hypothesize that focus will be expressed through one or more of the following acoustic properties: higher F0, greater duration, and/or increased intensity. We also hypothesize that focus may have an asymmetric realization dependent on the stress level and tone of a vowel. This prediction is based on the fact that both stress and tone are conveyed through acoustic properties also used in the signaling of focus: F0 in the case of tone and F0, duration, and intensity in the case of stress. As mentioned in Section 3, unstressed vowels in Ixcatec typically have reduced duration, intensity and F0 relative to their stressed counterparts (DiCanio 2012). We might thus expect unstressed syllables to be poorer sites than stressed syllables for realizing focus phonetically.

Moreover, an optional focus marker -\textit{na}² is encountered in Ixcatec, and a similar marker is reported for other closely-related languages such as Metzontla Popoloc...
To identify focus in Ixcatec, a language for which we have no native speaker’s intuitions, we rely on a discourse approach to focus and apply the Question-Answer Congruence principle (Büring 2012), where questions may be explicit or implicit.

Examples in (1) illustrate the use of the Ixcatec focus marker in a corrective focus condition, i.e. involving two exclusive alternatives (Büring 2012). However, we note that the use of the focus marker is optional and that the last Ixcatec speakers frequently omit it.

(1a) [tʰ'i²nə¹nə³]F-na²
1st-foc
'It is ME.'

(1b) [sɛ²ku²nə²ɗə³rja³]F-na²
middle-school-foc
'(At the) MIDDLE SCHOOL!'

The Ixcatec focus marker is also used in the contrastive focus condition, as can be seen in (2a). Similar to the corrective condition, the use of the focus marker is optional. Compare (2b), where the focus marker –na² is used, with (2c), where it is omitted. Note that focus is not marked syntactically, as there are no changes in the canonical word order Numeral-Noun.

(2a) ku² ju¹hu² la² [tʃe³]F-na²
coord two rel spicy-foc
'...and two that are SPICY.'

(2b) ti³la² [nɪ¹hɛ³]F-na¹ tʰi²
until three-foc day
'In THREE days...' [Fernández de Miranda 1961: 183, our glosses and translation from Spanish]²

2. The examples from the texts of Fernández de Miranda (1961) have been adapted to IPA for the sake of consistency with our newly collected data. Note that de Miranda uses both a mid and a high tone for the suffix -na.
Chapter 3. Prosodic and morphological focus marking in Ixcatec

[answer to the question: How many pairs of flowers does the *mayordomo* buy?]

(2c) \[\text{ʔu²te³ ni¹hẽ²}_F \text{ pa¹re²} \]

ten three pairs

‘THIRTEEN pairs (of flowers).’

Example (3a) illustrates the use of the Ixcatec focus suffix to mark contrastive topics \((\text{CT})\), i.e. constituents that are related to alternative questions (see a recent account in Büring 2016). Example (3b) shows that the use of the focus marker is optional in this condition, at least among the last Ixcatec speakers.

[answer to an implicit question: Who does what?]

(3a) \[\text{mi²-}[\text{ʔa²}]_\text{CT-na² ?u¹ ni²}\text{ɲu² sa¹ la² [nda²wa²]}_\text{CT-na²} \]

cls-woman- foc mill tortillas def rel male- foc

\[\text{ba²tu²-b-e²re²-fi² hŋu² ka²la² ju¹hu² ?u³} \]

ipfv.3pl- ipfv-give-appl one or two rope

‘…the WOMENCT mill tortillas […] the MENCT are collaborating with one or two ropes.’  

{Fernández de Miranda 1961: 181, 182 sentences 25, 26, our glosses and translation from Spanish}

[context: Niki and her mother are in front of the speaker. Niki is dancing, but her mother is falling asleep.]

answer to the question: What are we doing?]

(3b) \[\text{sa¹ [kʰa²-ni¹ki²]_CT ki¹=ʃte²-kʰa²} \]

[ne²ʔe¹e¹]_\text{CT} 

def clf.f-Niki prog.3sg-dance-co.3sg.f mother.poss.3sg

\[\text{ki¹=tsu²-kʰa²} \]

\[\text{ʃe²-kʰa²} \]

prog.3sg-want-co.3sg.f sleep-co.3sg.f

‘NIKICT is dancing, her MOTHERCT wants to sleep.’

{speaker F1_elic in context}

A further hypothesis relevant to the variable realization of morphological marking of focus is that the acoustic marking of focus will be more pronounced in tokens without an overt focus morpheme than in tokens with focus morphology expressed. This prediction is driven by the suggestion raised in the literature (e.g. Büring 2009) that a focus morpheme precludes, or at least discourages, phonetic exponents of focus.

To test these hypotheses, we conducted a phonetic study of the effects of corrective and contrastive focus on the acoustic parameters of duration, mean F0, and mean intensity. In 4.2 we discuss the methodology employed in this study. Section 4.3 presents the results and Section 4.4 a discussion of those results.
4.2 Methodology

A controlled experiment was designed to investigate which kinds of phonetic expression distinguish different kinds of focus by adapting the Animal Game task (Skopeteas et al. 2006; Swerts & Zerbian 2010) to the specificities of Ixcatec.

4.2.1 Participants

This study involved three of the four fluent Ixcatec speakers. The participants, two female and one male, are all in their 80s. Only one of them was brought up monolingual in Ixcatec and acquired Spanish at the age of six at school. The other two speakers acquired both Spanish and Ixcatec in their childhood. All of them have received little formal education. They all reside in the municipality of Santa María Ixcatlán (State of Oaxaca) and their everyday language is Spanish.

4.2.2 Stimuli

Thirty-two words were selected based on phonetic properties including tone and number of syllables, semantic field, and origin (only two Spanish-origin words were retained); see Table 4 in the Appendix. To avoid problems with picture-recognition, we used real-life, culturally-adapted objects and simple drawings (e.g. for colors or certain objects); see Figure 2.

Figure 2. Picture of a local preparation of corn for the target word *nixtamal*

---

3. The fourth fluent speaker was excluded due to a hearing impairment.
Procedure

Speakers were recorded separately, at home, in the village of Santa María Ixcatlán. During three sessions, participants were shown real-life objects, pictures, and drawings. In order to obtain enough tokens (two were targeted for each word under each condition for each speaker) to allow for robust generalizations while minimizing the risk that information structure effects would be washed out over the course of multiple repetitions, participants were shown real-life objects in the first session. During the second and third session, they were presented with photographs of the stimuli on a computer screen.

The objects were presented in a specific order designed to manipulate their discourse status. Three conditions were targeted: non-focus, contrastive focus, and corrective focus, where ‘contrastive focus’ refers to a constituent that introduces alternatives in the discourse, and ‘corrective focus’ offers two exclusive alternatives (Katz & Selkirk 2011; Büring 2012). Objects were grouped together by semantic field (colors, objects, food, animals, and numbers) and each series was introduced by an object that was not analyzed.

Note that participants always used isolated words. In the corrective condition, when they used the negative answer particle, ‘no’, the tokens were discarded from the phonetic analysis. Elicited words in isolation allowed for control of asymmetric declination effects in the various conditions.

For the contrastive condition, the speakers were instructed to name what they saw. For example, pink color would be followed by yellow color on a sheet of a paper, in which case the color term ‘yellow’ was assumed to be contextually contrastive. See an example in (4).

(4) \[\text{[The participants were shown a drawing of yellow color. They had just seen a drawing of pink color and described it as ‘(it is) pink].} \]

\[\text{[‘sa²ne²\text{\textsubscript{F}}\}}\]

\text{yellow}

\text{‘(It is) YELLOW.’}\]

For the corrective condition, the interviewer described the picture in Ixcatec using an inappropriate noun or color term. The speakers then corrected the interviewer, as if they were in a classroom, and proposed the correct term. See an example in (5).

(5) \[\text{[The participants were shown a drawing of yellow color. The interviewer suggested a wrong color name in Ixcatec: ‘ka\textsuperscript{\textsubscript{t}}se³ ‘red’.} \]

\[\text{[‘sa²ne²\text{\textsubscript{F}}\}}\]

\text{yellow}

\text{‘(It is) YELLOW.’}\]
Lastly, to elicit the non-focus condition and obtain comparable single-word tokens, the interviewer asked the translation of the target words from Spanish; see an example in (6). To avoid the contrastive focus effect, each word was introduced by a relatively long question with discussion of various related and unrelated topics. This procedure induced speakers to produce words with the default statement-final terminal fall in intonation.

\[ \text{[The participants were shown a drawing of yellow color. The interviewer asked in Spanish: } \textit{Como se dice en idioma?} 'how do you call (this) in your language (Ixcatec)?'] \]

(6) `sa²ne²
yellow
`Yellow.'

In total, the task elicited 576 tokens: 32 words \( \times \) 3 conditions (contrastive focus, corrective focus, and non-focus) \( \times \) 3 speakers \( \times \) 2 repetitions.

4.2.4 Recordings
Elicitation sessions were recorded using a Tascam DR-100 solidstate recorder at a 44.1kHz sampling rate via two microphones, a supercardioid head-worn microphone (Shure beta 54) and an AKGC480b handheld condenser microphone, the former of which provided the signal submitted to acoustic analysis.

4.2.5 Measurements
In order to assess the acoustic realization of different types of focus in Ixcatec, a series of measurements were made of all the vowels in the data set using Praat (Boersma & Weenink 2010). Based on a waveform in conjunction with a time-aligned spectrogram, the beginning and end of each vowel was demarcated. For vowels following a consonant other than a glide, the onset of a visible second formant was taken as the start of the vowel. For vowels preceding a consonant other than a glide, the offset of the vowel as determined by the second formant served as the end point. The start of a steady state second formant was taken as the beginning point for a vowel following a glide and the end of a steady state second formant served as the end point for a vowel preceding a glide. For word-final vowels, the primary criterion for delimiting the right edge of a vowel was the start of non-modal phonation (i.e. breathiness or creakiness) associated with sufficiently irregular pitch pulses to result in a failure of the pitch tracking algorithm.

A script was run within Praat to collect a series of measurements for all the segmented vowels. Measurements included duration, mean intensity and mean
F0. For a relatively small number of tokens (most commonly vowels with low tone), F0 values could not be extracted by the pitch-tracking algorithm and were consequently excluded from the analysis. In order to minimize microprosodic effects attributed to adjacent consonants, the window over which F0 was calculated excluded the 10 milliseconds at the beginning and the 10 milliseconds at the end of each vowel.

4.2.6 Statistical analysis and evaluation

For the statistical analysis, we employ multinomial regression modelling in order to assess which properties (i.e. variables) most reliably distinguish between the three different levels of focus with some predictive power. The dependent variable is therefore focus, a categorical variable with 3 levels: contrastive, corrective, and none. The independent variables (i.e. the variables potentially distinguishing between/predicting focus levels) fall broadly into three groups. First, there are three continuous variables, mean intensity, mean F0, and duration, all of which are possible phonetic exponents of focus.

The second group of independent variables are categorical and include a number of properties that might be expected to interact with the continuous variables. These include tone and stress level, both of which are involved in hypotheses to be tested, as well as other properties that have been shown in studies of other languages to potentially correlate and/or interact with one or more of the continuous variables: vowel quality, location of the syllable relative to the left edge of the word (equivalent in this study to the utterance given the isolation context in which the words appeared), word length measured in number of syllables, and speaker. An additional variable reflecting the presence vs. absence of the focus suffix is relevant only for the subset of data characterized by corrective focus since the suffix did not occur with contrastive focus in our data. The relationship between the focus suffix and the phonetic realization of focus is statistically explored in Section 4.3.3.

Finally, the variable speaker, which has one level for each of our three speakers, was included in the analysis to be able to evaluate speaker-specific differences.

4. Other F0 values taken at different time points (the beginning, middle and end of each vowel) as well as maximum F0 values were also collected but were not included in the statistical analyses reported in 5.2 after inspection suggested that they were not more effective at differentiating focus than a simpler measure of mean F0.

5. All statistical analysis were carried out with the open-source programming language and environment R 3.2.1 (R Core Team 2015), relying most heavily on the base package, but also the packages effects (Fox 2003; Fox & Hong 2009) and nnet (Venables & Ripley 2002).
The predictors (i.e. independent variables) employed in the analysis are summarized in Table 1.

Table 1. Independent variables employed in multinomial logistic regression model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Abbreviation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration (in seconds)</td>
<td>DURATION</td>
<td>0.026 to 0.471 seconds</td>
</tr>
<tr>
<td>Mean F0 (in Hz)</td>
<td>F0MEAN</td>
<td>103 to 334 Hz</td>
</tr>
<tr>
<td>Mean intensity (in dB)</td>
<td>INTENSMEAN</td>
<td>40.2 to 83.44 dB</td>
</tr>
<tr>
<td>Vowel quality</td>
<td>VOWEL</td>
<td>a, e, i, u</td>
</tr>
<tr>
<td>Tone</td>
<td>SYLLTONE</td>
<td>low, medium, high, unspecified*</td>
</tr>
<tr>
<td>Stress</td>
<td>STRESS</td>
<td>stressed, unstressed</td>
</tr>
<tr>
<td>Syllable location</td>
<td>SYLLFROMLEFT</td>
<td>1st, 2nd, 3rd, 4th</td>
</tr>
<tr>
<td>Length of Root (in no. of syllables)</td>
<td>ROOTLENGTH</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Speaker</td>
<td>SPEAKER</td>
<td>F1, F2, M1</td>
</tr>
<tr>
<td>Focus Suffix (see Section 4.3.3)</td>
<td>SUFFIXATROOT</td>
<td>Yes(=suffix), No(=no suffix)</td>
</tr>
</tbody>
</table>

* In order not to bias the analysis, the tonal category unspecified was employed for the optional suffix -na, which has been variously transcribed as mid or high tone by Fernández de Miranda (1961).

A key virtue of including all the potential predictors of focus level together in the regression model is that it provides a means for assessing the relative efficacy of the predictors all at the same time, thereby avoiding the possibility of misjudging the predictive capacity of one or more variables, which happens when independent variables are considered in isolation. For example, if duration emerged as a reliable predictor of focus level in one analysis and f0mean were significant in a separate analysis, it would be unclear how much of the success of each predictor was in fact covertly attributed to the other one. Another benefit to a single regression model encompassing all independent variables is its ability to discover potential interactions between variables that would be missed if variables were evaluated in separate analyses.

As an initial exploratory step, the distribution of data points was inspected. The histogram in Figure 3 provides a visual display of the duration (Figure 3a), f0mean (Figure 3b), and intensmean (Figure 3c) data. To provide a better sense of the degree to which the distributions are normal, the data are divided into ten quantiles, each representing 10% of the data points.

As Figure 3 shows, even if one abstracts away from the outliers that broaden the distribution, all of the parameters display a wide range of values. More importantly, none of the phonetic parameters have a perfectly normal distribution. F0 and especially duration are skewed leftward such that the bulk of the data points occupy the lower half of the plot, whereas intensity shows the opposite pattern of rightward skewing.

In order to make its distribution more normal, we logged the values of the variable duration. In addition, we z-standardized duration as well as intensmean...
a. The distribution of data for the continuous variable **duration**

![Histogram for duration](image)

b. The distribution of data for the continuous variable **f0mean**

![Histogram for f0mean](image)

c. The distribution of data for the continuous variable **intensmean**

![Histogram for intensmean](image)

Figure 3.
and f0mean in order to protect ourselves against collinearity (the fact that predictors might be highly correlated with each other, which can give rise to highly unstable regression coefficients) and to be able to evaluate their effects all on the same scale. We then proceeded to explore to what degree the independent variables and their pairwise interactions would predict the variable focus using a multinomial regression analysis. Given that the variable suffixatroot was deterministically correlated with focus (i.e. all forms with an overt focus suffix were associated with corrective focus), our initial regression approach was only applied to the 772 unsuffixed cases (277 with contrastive focus, 205 with corrective focus, 290 with no focus).

In a first step, we generated the null model, equivalent to a model with only speaker as a predictor (to immediately allow for speaker-specific differences). Then, we employed an automatic stepwise and bidirectional model selection procedure using AIC (Akaike Information Criterion). This approach entails enlisting an algorithm that begins from the smallest possible model – the one not including only speaker – and iteratively adds or subtracts predictors to improve the fit of the model. Goodness of fit is defined in terms of AIC, a criterion that evaluates the fit of a model against its number of parameters (i.e. it effectively integrates Occam’s razor into the model selection process). Note that for all numeric predictors we did not merely include the predictor per se, but also implemented it as an orthogonal polynomial to the second degree; this is rarely done but is in fact very useful because it allows the regression algorithm to identify whether (some of) the trends of the numeric predictors in the data exhibit curvature rather than the traditional ‘straight regression lines only’ approach. This process stops with the discovery of a so-called ‘minimal adequate’ model, a model that cannot be improved by either adding or deleting a predictor to the model. In other words, predictors that are not encompassed by this final model did not contribute enough to be included. For this minimal adequate model, we then provide summary statistics (to assess overall model quality), classification accuracy (to assess model accuracy), as well as visualizations of the model’s effects.

4.3 Results

As a first interim result, the automatic model selection process returned a model that was significantly better than the null model. The final, minimal adequate model (summarized in Table 2; see Table 5 in the appendix for further details) is a highly significant improvement over the null model ($LR$-statistic = 190.007, $df = 32$, $p < 0.001$). This model comes with a classification accuracy of 51%, which, according to binomial tests, is highly significantly better than either just picking the most frequent focus category or choosing focus categories randomly; $\lambda^*$ improvement prediction accuracy = 0.216.
Table 2. Overall results of the minimal adequate model (highest-order effects)

<table>
<thead>
<tr>
<th>Effect</th>
<th>LR-statistic (type II)</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0mean</td>
<td>21.308</td>
<td>2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>speaker : duration</td>
<td>15.432</td>
<td>4</td>
<td>&lt;0.004</td>
</tr>
<tr>
<td>speaker : poly(intensmean, 2)</td>
<td>90.5</td>
<td>8</td>
<td>&lt;10^{-15}</td>
</tr>
<tr>
<td>speaker : stress</td>
<td>20.254</td>
<td>4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>vowel</td>
<td>16.5</td>
<td>6</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Sections 4.3.1 (Main Effects) and 4.3.2 (Interactions) discuss the relevant significant highest-order effects of this model (see Table 6 in the appendix for all coefficients of this model); we are not discussing the effect of vowel, which was merely included as a control. Section 4.3.3 presents the relationship between the morphological and phonetic expression of focus.

4.3.1 The main effect of F0
Figure 4 visualizes the effect of F0 on FOCUS: the x-axis represents the (for ease of understanding, again unstandardized) values of F0mean, the y-axis represents the predicted probability of an outcome (i.e. a kind or absence of focus), and the three regression lines with shaded confidence intervals represent the predicted probabilities of the different outcomes under the three levels of FOCUS. The rugs on
the x-axis indicate the actually observed values of f0mean, the vertical dashed line represents the median of the f0mean values, and the segmented line at y = 0.95 represents 10% quantiles of f0mean. The nature of this effect is plain to see: (i) When f0mean is lower than average, focus is none; (ii) when f0mean is higher than average, then focus is contrastive.

4.3.2 The interactions
The next three effects all involve the predictor speaker; these effects thus involve predictors whose effect on focus is not constant across speakers. Consider first the relationship between duration and focus, which is different across speakers. Figure 5 is a visualization of this interaction.

a. The effect of the interaction of duration:speaker on focus for speaker F1
b. The effect of the interaction of duration:speaker on focus for speaker F2

The effect of duration for speaker F2 on the predicted probabilities of focus (p < 0.004)

Predicted probability of levels of focus

<table>
<thead>
<tr>
<th>Duration</th>
<th>contrastive</th>
<th>corrective</th>
<th>none</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.033</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.077</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.182</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.279</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.427</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.

c. The effect of the interaction of duration:speaker on focus for speaker M1

The effect of duration for speaker M1 on the predicted probabilities of focus (p < 0.004)
The plots in Figure 5a and 5b show the results for the two female speakers (F1 and F2), the plot in Figure 5c the corresponding effect for the male speaker (M1). The interaction shows that the two female speakers pattern very similarly to each other and very differently from the male speaker. Specifically, the female speakers are more likely to mark contrastive focus with increased duration and less likely to mark no focus with increased duration, whereas the male speaker uses increased duration for corrective focus.

The plots in Figure 6 are an analogous representation of the interaction INTENSITY:SPEAKER.

a. The effect of the interaction of INTENSITY:SPEAKER on FOCUS for speaker F1
b. The effect of the interaction of INTENSITY:SPEAKER on FOCUS for speaker F2

The effect of INTENSITY for SPEAKER: F2
on the predicted probabilities of FOCUS (p < 0.00001)

Predicted probability of levels of FOCUS
contrastive corrective none

Intensity

Figure 6.

c. The effect of the interaction of INTENSITY:SPEAKER on FOCUS for speaker M1

The effect of INTENSITY for SPEAKER: M1
on the predicted probabilities of FOCUS (p < 0.00001)

Predicted probability of levels of FOCUS
contrastive corrective none

Intensity

Figure 6.
The nature of this effect defies easy characterization: low intensity is associated with corrective focus for F1 and M1 but contrastive focus for F2; medium intensity is associated with no focus for all speakers (though less so for F2); finally, high intensity is associated with contrastive focus and no focus for F1, no focus for F2, and corrective focus for M1. In general, the confidence intervals are highly overlapped at both ends of the scale with one exception: the relationship between high intensity and corrective focus for M1.

Finally, the plots in Figure 7 represent the interaction stress : speaker (with predicted probabilities and their confidence intervals). On the whole, this interaction is weak: in most cases, the change from unstressed to stressed results in only small changes of predicted probabilities (and most of the confidence intervals of the predicted values overlap); the main source of significance is that, for the male speaker, stressed is associated significantly more with no focus than unstressed and significantly less with corrective focus than unstressed.

a. The effect of the interaction of stress:speaker on focus for speaker F1
b. The effect of the interaction of stress:speaker on focus for speaker F2

The effect of STRESS for SPEAKER: F2
on the predicted probabilities of FOCUS (p < 0.001)

Predicted probability of levels of FOCUS

/zero.oldstyle./two.oldstyle
/zero.oldstyle./four.oldstyle
/zero.oldstyle./six.oldstyle
/zero.oldstyle./eight.oldstyle
/one.oldstyle./zero.oldstyle

contrastive corrective none

Stress unstressed stressed

Figure 7.

c. The effect of the interaction of stress:speaker on focus for speaker M1

The effect of STRESS for SPEAKER: M1
on the predicted probabilities of FOCUS (p < 0.001)

Predicted probability of levels of FOCUS

/zero.oldstyle./two.oldstyle
/zero.oldstyle./four.oldstyle
/zero.oldstyle./six.oldstyle
/zero.oldstyle./eight.oldstyle
/one.oldstyle./zero.oldstyle

contrastive corrective none

Stress unstressed stressed

Figure 7.
4.3.3 The relationship between the morphological and phonetic expression of focus

In the experimental task, -na² was encountered only in the corrective condition and not in the contrastive condition (e.g. [ju²wa³]F-na² ‘green–foc’). In the analyzed data, the focus suffix appeared in 46.4% \((N = 26/56)\) of corrective focus tokens for one female speaker, 22% \((N = 11/50)\) for the other female speaker, and none for the male speaker. In order to test the hypothesis that phonetic expression of focus will be stronger in tokens without morphological marking of focus, a regression model was fit to the data for only the tokens with corrective focus. Recall that only corrective focus had the option of morphological marking in our data. In the analysis, SUFFIXATROOT served as the dependent variable and the predictor variables were otherwise the same as those employed in the original model. It should be noted that the coding for the variable STRESS was the same in the suffixed as in the unsuffixed forms since the suffix falls outside the domain of stress and thus does not trigger a rightward shift of stress from the penultimate syllable of the root. Similarly, the coding for SYNTONE also did not vary between suffixed and unsuffixed forms as the contrast between mid and low tone, which is lost in root non-final syllables (see Section 3), is preserved in the final syllable of a root appearing before a suffix.

The minimal adequate model we arrived at indicated a highly significant correlation between the predictors and SUFFIXATROOT \((LR\text{-statistic}=66.48, df = 9, p < 10^{-10})\), and a medium-sized correlation \((Nagelkerke R^2 = 0.45)\), a classification accuracy 81.7%, which is highly significantly better than either baseline and a good C-value of 0.851. The highest-order effects are summarized in Table 3.

<table>
<thead>
<tr>
<th>Effect</th>
<th>LR-statistic (type II)</th>
<th>df</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPEAKER</td>
<td>2.303</td>
<td>1</td>
<td>0.13</td>
</tr>
<tr>
<td>STRESS</td>
<td>2.418</td>
<td>1</td>
<td>0.12</td>
</tr>
<tr>
<td>VOWEL</td>
<td>6.365</td>
<td>3</td>
<td>0.095</td>
</tr>
<tr>
<td>poly(intensmean.zstand, 2)</td>
<td>8.401</td>
<td>2</td>
<td>&lt;0.015</td>
</tr>
<tr>
<td>poly(duration.zstand, 2)</td>
<td>20.53</td>
<td>2</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Before considering the results of primary interest, those involving the continuous phonetic variables of intensity, and duration, there are three other less interesting effects that were included in the model. These were insignificant but included because our model selection process used AIC as a selection criterion. First, there was an effect of vowel quality, such that the non-high vowels /a, e/ were slightly more predictive than the high vowels of the presence of the focus suffix. Second, there
was a difference between the speakers such that F1 exhibited a higher occurrence of the focus suffix. Finally, there was a tendency for lack of stress to predict the presence of the focus suffix.

Turning to the effects that are significant and of particular interest, Figures 8 and 9 depict the predicted probabilities of the focus suffix surfacing as a function of DURATION (Figure 8) and INTENSMEAN (Figure 9); for ease of interpretation, values on the x-axes are unstandardized in both figures.

Figure 8 shows a very salient effect of duration on the probability of the focus suffix being realized: an increase in duration of vowels in the root (as suffixal vowels were not included in the analysis) results in a sharp increase in the likelihood of the suffix occurring.

![Graph showing the main effect of duration on the predicted probabilities of suffix at root](image)

**Figure 8.** The effect of duration (polynomial, 2) on SUFFIXATROOT

As Figure 9 shows, there is a strongly inverse U-shaped trend. However, it is instructive again to pay attention to the quantiles: the increase on the left side of the plot is supported by a mere 10% of the INTENSMEAN data – what remains once those are weighted by their relatively low frequency is a much more pronounced and robust decreasing trend such that, with increasing INTENSMEAN the suffix becomes less likely.
4.4 Discussion

Results of the acoustic study indicate that focused words are phonetically differentiated from their unfocussed counterparts and, further, that corrective and contrastive focus are also acoustically distinguished. This result corroborates the primary hypothesis tested in the phonetic study: that focus has phonetic exponents in Ixcatlec, in keeping with the lack of focus marking through word order changes and the only sporadic morphological marking of focus. Most consistently across speakers, higher F0 values are associated with contrastive focus whereas lowered F0 values are associated with lack of focus. Furthermore, decreased duration is predictive of lack of focus for all speakers.

There are, however, other properties that vary between speakers. At the upper end of the spectrum of duration values an increase in duration triggers a greater likelihood of contrastive focus for the two female speakers. The male speaker, on the other hand, employs increased duration as a marker of corrective focus.

Intensity displays the greater interspeaker variation in its behavior and is also generally the least reliable predictor of focus as reflected in its characteristically very broad confidence intervals. For the male speaker (the speaker with the narrowest
confidence bands), corrective focus is associated with greater intensity to go along with the increase in duration associated with corrective focus. One of the female speakers (F1) displays greater intensity under the contrastive focus condition in keeping with the increased duration also observed under contrastive focus. The other female speaker (F2) has a divergent pattern characterized by decreased intensity under contrastive focus and increased intensity under lack of focus.

Synthesizing the results by speaker, the male speaker appears to display the clearest phonetic distinctions between focus levels. In his speech, contrastive focus is associated with higher F0, while corrective focus is associated with increased duration and intensity. A reduction in any of the three patterns is predictive of lack of focus. For the two female speakers, increased duration and F0 are both predictive of contrastive focus. For one of the female speakers, increased intensity is also associated with greater probability of contrastive focus. For the other speaker, the increase in duration under contrastive focus is paradoxically accompanied by a decrease in intensity. Interestingly, for both of the female speakers, there is no phonetic dimension along which an increase in the relevant property distinguishes corrective focus from the other two focus conditions. Rather, it is the absence of a reduction in F0 and duration that differentiates corrective focus from lack of focus (and from contrastive focus as well). This result may be due to a ceiling effect, whereby the exploitation of increased F0 and duration to signal contrastive focus renders these phonetic properties less available as markers of corrective focus. In any case, corrective focus is still phonetically distinguishable from a lack of focus by virtue of possessing (relative to the unfocused condition) greater duration and F0.

With respect to stress, its association with lack of focus, on the one hand, supports the claim that Ixcatec possesses stress in addition to tone (DiCanio 2012). On the other hand, its lack of efficacy in predicting focus is plausibly attributed to a ceiling effect whereby the properties used to signal stress are less available to signal focus. These patterns were most apparent for the male speaker, for whom lack of stress was predictive of corrective focus and stress was predictive of the no focus condition. This finding contradicts the hypothesis that focus would have a more robust realization in stressed syllables, which are inherently conducive to supporting the same prominence-lending properties characteristically associated with focus. One interpretation of the present findings is that there is actually more freedom to implement focus phonetically in unstressed syllables relative to stressed syllables, the latter of which already possess certain of the same features that mark corrective focus for the male speaker. In a cross-linguistic study of the acoustic interaction between focus and stress in four languages (Greek, Hungarian, Spanish, and Turkish), Vogel et al. (2016) find differences between languages in the realization of stress under focus, even observing a reduction in the prominence of stress under focus.
The hypothesis that focus may have a more pronounced realization for certain tones was not corroborated. Phonemic tone did not predict focus condition in the main analysis excluding suffixed forms and also did not reliably predict whether corrective focus was realized with a suffix or not.

Examining the results from a typological lens, the marking of focus through a (speaker-dependent) change in duration, intensity and/or F0 is unremarkable typologically, though it has not been quantitatively documented for any Otomanguean languages. There has been very little quantitative work comparing the acoustic reflexes of different types of focus within a single study. Even though at present the main assumption is that all the pragmatic contexts (contrastive focus, corrective focus, answer focus, etc.) trigger a single type of grammatical focus, the possibility that the various pragmatic types of focus could also be considered as distinct grammatical types of focus remains open (Büring 2009: 180). The differences in the phonetic marking of corrective and contrastive focus observed in Ixcatec thus inform the broader typology and contribute to the ongoing discussion in focus theory. The association of lack of focus with a reduction along one or more phonetic dimensions (varying on a speaker-specific basis) conforms, however, to cross-linguistic patterns.

Perhaps somewhat unexpected is the role of F0 in marking focus in Ixcatec. A priori one might expect F0 to be less available as a phonetic correlate of focus in Ixcatec due to its use on a lexical level to distinguish different words (see Chen & Gussenhoven 2008: 726 on Standard Chinese). Conversely, one might predict intensity to be a more reliable marker of focus by analogy with the important role of intensity in marking stress in tone languages, such as Thai (Potisuk et al. 1996) and Pirahã (Everett 1998). Our data, however, suggest that F0 does play a role in predicting focus in Ixcatec, especially contrastive focus. F0, in fact, is a more consistent predictor of different focus conditions in Ixcatec than either intensity or duration. The Ixcatec results thus indicate that, although the use of F0 may be constrained in the marking of focus in a tone language, it is certainly not precluded.

Intensity played a more robust role in predicting whether corrective focus was expressed morphologically through a suffix or not: as intensity increased, the likelihood of the focus suffix surfacing decreased, a finding that is consistent with the hypothesis that morphological and phonetic exponents of focus are in complementary distribution. This hypothesis was only partially confirmed, however, since an increase in duration was associated with a greater likelihood of the focus suffix being realized.

One suggestive finding that cannot be explored further is the divergence between the male speaker and the two female speakers, most conspicuously in the relationship between duration and focus, such that increased duration was predictive
of contrastive focus for the females but corrective focus for the male. Because there was only one male speaker in our study, it is unclear whether this discrepancy is a function of gender or of idiolect. Sadly, this confound can never be teased apart as the three consultants studied for this paper are the last fluent speakers of Ixcatec.

5. Conclusions

The present study of Ixcatec contributes to our understanding of focus in several ways both from a theoretical and a descriptive standpoint. On a descriptive level, the current paper broadens the typological database on focus by providing the first comprehensive analysis of phonetic and morphological features of focus in an Otomanguean language. Furthermore, our work contributes to the extremely meager descriptive literature on focus in severely threatened languages by examining through varied methodological approaches a moribund language that has only a handful of fluent speakers.

In our experimental data, a focus suffix occurs only with tokens associated with corrective focus. More striking is the relationship between the optional focus suffix and the phonetic exponents of focus in Ixcatec. The results of the second regression analysis limited to the corrective focus condition showed that prosodic marking can be used concurrently with the focus marker in corrective conditions. An increase in duration of the root was thus predictive of the occurrence of the focus suffix. This finding appears to contradict the hypothesis that the morphological and prosodic marking of focus is parameterized on a language-specific basis (Büring 2009). An extreme interpretation of this position is that the morphological and acoustic expression of focus are mutually exclusive and that a language may employ one but not the other. This view seems unlikely to be true on the basis of not only the Ixcatec data but also based on studies of other languages with both morphological and prosodic marking of focus, such as Moroccan Arabic (Benkirane 1998). A less categorical and a priori more defensible version of the hypothesis of mutual exclusivity would be that the degree of reliance on morphological vs. prosodic marking of focus is inversely related: the more a language relies on morphology to cue focus, the less it depends on acoustic cues, and vice versa. Even this position, however, is contradicted by the Ixcatec results, in which morphological and prosodic marking of focus appear to act synergistically, such that the acoustic expression of focus is more salient, at least along the phonetic dimension of duration, in conjunction with the focus suffix than without it.

To complicate matters, intensity displayed a different pattern from duration in our data: morphological marking of focus was associated with lower intensity. The
overall picture is thus not one of an unambiguously synergistic relationship between prosody and morphology but rather that focus has a different acoustic realization depending on whether it co-occurs with a suffix or not.

In contextualizing the Ixcatec results, it is important to note that the set of languages subjected to a comprehensive study of both the acoustic and morphosyntactic exponents of focus is still relatively small. Indeed, future typological work on focus might reveal that, similar to Ixcatec, other languages employing morphological marking of focus may also use prosody redundantly to signal focus, in which case it will be necessary to rethink the relationship between prosody and morphology in the expression of focus.

Acknowledgements

We are grateful to the two anonymous reviewers for their many helpful and insightful comments. We are profoundly indebted to the speakers of Ixcatec who generously shared their language with us. The data collection was funded by the Hans Rausing Endangered Languages Programme “Textual and Lexical Documentation of Ixcatec, a highly endangered Otomanguean language of Oaxaca, Mexico” (MDP 0214, 2010–2013), awarded to Denis Costaouec, and the French National Research Agency through the program “Investments for the Future” (ANR-10-LABX-0083), Strand 3 GD1 “The Typology and Corpus Annotation of Information Structure and Grammatical Relations”.

Abbreviations

1, 3 first and third person
APPL applicative
CLF classifier
CLS class
CO cross-reference
COORD coordinator
DEF definite article
FOC focus
IPFV imperfective
PL plural
POSS possessive
PROG progressive
REL relative
SG singular
References


Fernández de Miranda, María Teresa. 1959. *Fonémica del ixcateco*. Mexico: INAH.


Appendix

Table 4. Target words of the experiment

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>'ru¹wa¹</td>
<td>white</td>
</tr>
<tr>
<td>'me¹sa¹</td>
<td>table</td>
</tr>
<tr>
<td>'ju¹hu²</td>
<td>two</td>
</tr>
<tr>
<td>'ka¹tse³</td>
<td>red</td>
</tr>
<tr>
<td>'tju²jĩ¹</td>
<td>lemon</td>
</tr>
<tr>
<td>'la²jẽ¹</td>
<td>sweet</td>
</tr>
<tr>
<td>'sa²ne²</td>
<td>yellow</td>
</tr>
<tr>
<td>'tsa²ku²</td>
<td>leg</td>
</tr>
<tr>
<td>'ni³nu²</td>
<td>tortilla</td>
</tr>
<tr>
<td>'tʃa²kũ³</td>
<td>sun</td>
</tr>
<tr>
<td>'nu²ngu³</td>
<td>church</td>
</tr>
<tr>
<td>'ka³ne³</td>
<td>nixtamal</td>
</tr>
<tr>
<td>'ka³tse³</td>
<td>red</td>
</tr>
<tr>
<td>'tʃu² (hʃa)</td>
<td>pumpkin</td>
</tr>
<tr>
<td>'la²ʃe¹</td>
<td>flower</td>
</tr>
<tr>
<td>'me¹sa¹</td>
<td>table</td>
</tr>
<tr>
<td>'ju¹hu²</td>
<td>two</td>
</tr>
<tr>
<td>'ka¹tse³</td>
<td>red</td>
</tr>
<tr>
<td>'tʃu² (hʃa)</td>
<td>pumpkin</td>
</tr>
</tbody>
</table>

Table 5. The minimal adequate multinomial regression model (Focus ~ ...)

<table>
<thead>
<tr>
<th>Term</th>
<th>Contrastive → Corrective</th>
<th>Contrastive → None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef</td>
<td>SE</td>
<td>Z</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>-1.187</td>
<td>0.310</td>
</tr>
<tr>
<td>SPEAKER: female → male</td>
<td>-1.809</td>
<td>0.490</td>
</tr>
<tr>
<td>SPEAKER F1 → F2</td>
<td>-1.743</td>
<td>0.751</td>
</tr>
<tr>
<td>DURATION.zstand</td>
<td>0.536</td>
<td>0.144</td>
</tr>
<tr>
<td>f0MEAN.zstand</td>
<td>0.211</td>
<td>0.143</td>
</tr>
<tr>
<td>STRESS: unstressed -&gt; stressed</td>
<td>-0.401</td>
<td>0.258</td>
</tr>
<tr>
<td>VOWEL: a → e</td>
<td>-0.127</td>
<td>0.379</td>
</tr>
<tr>
<td>VOWEL: a → i</td>
<td>0.207</td>
<td>0.289</td>
</tr>
<tr>
<td>VOWEL: a → u</td>
<td>0.856</td>
<td>0.279</td>
</tr>
<tr>
<td>poly(intensMEAN.zstand,2)1</td>
<td>9.902</td>
<td>7.016</td>
</tr>
<tr>
<td>poly(intensMEAN.zstand,2)2</td>
<td>5.720</td>
<td>5.247</td>
</tr>
<tr>
<td>SPEAKER: female → male : STRESS: unstressed → stressed</td>
<td>1.671</td>
<td>0.506</td>
</tr>
<tr>
<td>SPEAKER: F1 → F2 : STRESS: unstressed → stressed</td>
<td>0.677</td>
<td>0.645</td>
</tr>
<tr>
<td>SPEAKER: female → male : DURATION.zstand</td>
<td>-1.043</td>
<td>0.299</td>
</tr>
</tbody>
</table>

(continued)
Table 5. (continued)

<table>
<thead>
<tr>
<th>contrastive → corrective</th>
<th>contrastive → none</th>
</tr>
</thead>
<tbody>
<tr>
<td>speaker: F1 → F2: duration.zstand</td>
<td>coef se z</td>
</tr>
<tr>
<td>-0.280 0.352 0.796</td>
<td>-0.008 0.274 -0.029</td>
</tr>
<tr>
<td>speaker: female → male: poly(intensmean.zstand,2)</td>
<td>coef se z</td>
</tr>
<tr>
<td>-16.304 12.671 -1.287</td>
<td>29.635 9.935 2.983</td>
</tr>
<tr>
<td>speaker: F1 → F2: poly(intensmean.zstand,2)</td>
<td>coef se z</td>
</tr>
<tr>
<td>speaker: female → male: poly(intensmean.zstand,2)</td>
<td>coef se z</td>
</tr>
<tr>
<td>-38.064 10.880 -3.499</td>
<td>-5.191 9.850 -0.527</td>
</tr>
</tbody>
</table>

Table 6. The minimal adequate binary logistic regression model (SuffixatRoot ~ …)

<table>
<thead>
<tr>
<th>coef</th>
<th>se</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.6806 0.5268</td>
<td>1.292</td>
<td>0.1964</td>
</tr>
<tr>
<td>speaker: F1 → F2</td>
<td>-0.8648 0.5713</td>
<td>-1.514</td>
<td>0.1301</td>
</tr>
<tr>
<td>poly(duration.zstand,2)</td>
<td>14.4240 3.7028</td>
<td>3.895</td>
<td>9.8e-05</td>
</tr>
<tr>
<td>poly(duration.zstand,2)</td>
<td>4.4426 4.0209</td>
<td>1.105</td>
<td>0.2692</td>
</tr>
<tr>
<td>poly(intensmean.zstand,2)</td>
<td>-8.9616 4.8473</td>
<td>-1.849</td>
<td>0.0645</td>
</tr>
<tr>
<td>poly(intensmean.zstand,2)</td>
<td>-8.2454 3.6944</td>
<td>-2.232</td>
<td>0.0256</td>
</tr>
<tr>
<td>stress unstressed → stressed</td>
<td>-0.7495 0.4886</td>
<td>-1.534</td>
<td>0.1250</td>
</tr>
<tr>
<td>vowel: a → e</td>
<td>0.2916 0.7048</td>
<td>0.414</td>
<td>0.6791</td>
</tr>
<tr>
<td>vowel: a → i</td>
<td>-1.1932 0.5961</td>
<td>-2.002</td>
<td>0.0453</td>
</tr>
<tr>
<td>vowel: a → u</td>
<td>-0.5656 0.5576</td>
<td>-1.014</td>
<td>0.3104</td>
</tr>
</tbody>
</table>

Author queries

Please provide a citation for the reference id “CIT0116 (Bolinger, Dwight L. 1961), CIT0143 (Malchukov, Andrej, Haspelmath, Martin & Comrie, Bernard (eds). 2010)” since citation is missing in the article.