



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Transphonologization of voicing in Chru

**Citation for published version:**

Brunelle, M, Tan, TT, Kirby, J & Giang, L 2020, 'Transphonologization of voicing in Chru: Studies in production and perception', *Laboratory Phonology*, vol. 11, no. 1, pp. 1-33.  
<https://doi.org/10.5334/LABPHON.278>

**Digital Object Identifier (DOI):**

[10.5334/LABPHON.278](https://doi.org/10.5334/LABPHON.278)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Laboratory Phonology

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



JOURNAL ARTICLE

# Transphonologization of voicing in Chru: Studies in production and perception

Marc Brunelle<sup>1</sup>, Tạ Thành Tấn<sup>1</sup>, James Kirby<sup>2</sup> and Đinh Lư Giang<sup>3</sup>

<sup>1</sup> Department of Linguistics, University of Ottawa, CA

<sup>2</sup> School of Philosophy, Psychology and Language Sciences, University of Edinburgh, UK

<sup>3</sup> Department of Linguistics and Department of Spanish Studies, University of Social Sciences and Humanities, Hồ Chí Minh City, VN

Corresponding author: Marc Brunelle ([mbrunell@uottawa.ca](mailto:mbrunell@uottawa.ca))

Chru, a Chamic language of south-central Vietnam, has been described as combining contrastive obstruent voicing with incipient registral properties (Fuller, 1977). A production study reveals that obstruent voicing has already become optional and that the voicing contrast has been transphonologized into a register contrast based primarily on vowel height (F1). An identification study shows that perception roughly matches production in that F1 is the main perceptual cue associated with the contrast. Structured variation in production suggests a sound change still in progress: While younger speakers largely rely on vowel height to produce the register contrast, older male speakers maintain a variety of secondary properties, including optional closure voicing. Our results shed light on the initial stages of register formation and challenge the claim that register languages must go through a stage in which breathiness or aspiration is the primary contrastive property (Haudricourt, 1965; Wayland & Jongman, 2002; Thurgood, 2002). This article also complements several recent studies about the transphonologization of voicing in typologically diverse languages (Svantesson & House, 2006; Howe, 2017; Coetzee, Beddor, Shedden, Styler, & Wissing, 2018).

**Keywords:** Register; voicing; Chru; transphonologization; production; perception; cue weights

## 1. Introduction

Recent studies have revealed several cases of transphonologization of laryngeal contrasts in languages as diverse as Kammu, Malagasy, and Afrikaans (Svantesson & House, 2006; Howe, 2017; Coetzee et al., 2018). In these languages, low-level f<sub>0</sub> perturbations induced by onset voicing or aspiration have become contrastive as VOT-based contrasts were neutralized, providing apparent-time evidence for a process that was diachronically and phonetically well-established (House & Fairbanks, 1953; Haudricourt, 1954; Hyman, 1976; Hombert, Ohala, & Ewan, 1979). In many Southeast Asian languages, however, the same type of onset voicing is transphonologized into a bundle of properties called *register*, characterized by contrastive differences in phonation type,<sup>1</sup> vowel quality, and/or duration in addition to f<sub>0</sub> (Henderson, 1952).

In this paper, we explore the development of register in Chru, a Chamic language (Austronesian) of Vietnam. Chru was described by Fuller (1977) as having a voicing contrast accompanied by redundant register, and was chosen as it could inform us about the earliest stages of register formation, when voicing is still contrastive but is gradually enhanced by non-automatic, or *extrinsic*, register properties (Hyman, 1976). However,

<sup>1</sup> We follow Abercrombie (1967) and Laver (1980) in using the term ‘phonation’ rather than ‘voice quality.’ The former is limited to laryngeal settings, while the latter also includes supralaryngeal modulations.

it became obvious as we were conducting our production study that Chru voicing has become optional and that its register system is already contrastive. The results of our production and perception experiments further reveal that this contrastive register is largely structured around F1, contrasting with other cases of instrumentally-described transphonologization in which it is f0 that takes on the functional role of voicing. Our results also show that there is socially structured variation in the realization of Chru voicing and register: We argue that this structured variation is evidence for a possible sound change in progress.

### 1.1. Voicing and the diachronic development of register

Register is a common phonological contrast in the Austroasiatic<sup>2</sup> and Chamic languages of Mainland Southeast Asia (Henderson, 1952; Haudricourt, 1965; Gregerson, 1976; Huffman, 1976; Ferlus, 1979). It consists in a bundle of acoustic properties—the most important being f0, phonation type, and vowel quality—that are realized on rhymes but originate from a voicing contrast in onsets (no voicing contrast in sonorants is reconstructed in Proto-Austroasiatic and Proto-Chamic). The *low* register (also *second* or *lax* register) derives from former voiced stops, while the *high* register (also *first* or *tense* register) stems from original voiceless stops. Low register syllables typically have a lower f0 and a laxer/breathier phonation and they often have more close vowels (or vowels with close onglides). Note that register languages do not necessarily combine all of these properties, and that others, like VOT and vowel duration, are often associated with the contrast. This characterization of register, summarized in **Table 1**, is supported by acoustic evidence from Austroasiatic (Lee, 1983; L.Thongkum, 1987, 1989, 1991; Wayland, 1997; Watkins, 2002; Wayland & Jongman, 2003; Abramson, Luangthongkum, & Nye, 2004; Abramson, Nye, & Luangthongkum, 2007; DiCanio, 2009; Abramson, Tiede, & Luangthongkum, 2015; Tạ, Brunelle, & Nguyễn, 2019) and Austronesian languages (Fagan, 1988; Edmondson & Gregerson, 1993; Hayward, 1995; Thurgood, 2004; Brunelle, 2005, 2009a, 2010; Matthews, 2015).

The connection between onset voicing and most of the individual acoustic properties of register is relatively well understood. There is ample acoustic and perceptual evidence that vowels have a higher f0 following voiceless than voiced obstruents (Hombert et al., 1979; Ohde, 1984; Lisker, 1986; Hanson, 2009; Dmitrieva, Llanos, Shultz, & Francis, 2015; Kirby & Ladd, 2016). Several compatible mechanisms have been proposed to account for this effect: 1) the higher f0 after voiceless obstruents could be attributable to

**Table 1:** Typical properties of register (adapted from Brunelle & Kirby, 2016).

<b>High register</b> <b>(also tense, clear, or first register)</b> <b>(&lt;voiceless stops, [*pa])</b>	<b>Low register</b> <b>(also lax, breathy, or second register)</b> <b>(&lt;voiced stops, [*ba])</b>
Higher pitch	Lower pitch
Tense/modal voice	Lax/breathy voice
Raised F1 (whole vowel or beg. of vowel)	Lowered F1 (whole vowel or beg. of vowel)
More peripheral F2?	More centralized F2?
Shorter VOT	Longer VOT
Shorter vowels	Longer vowels

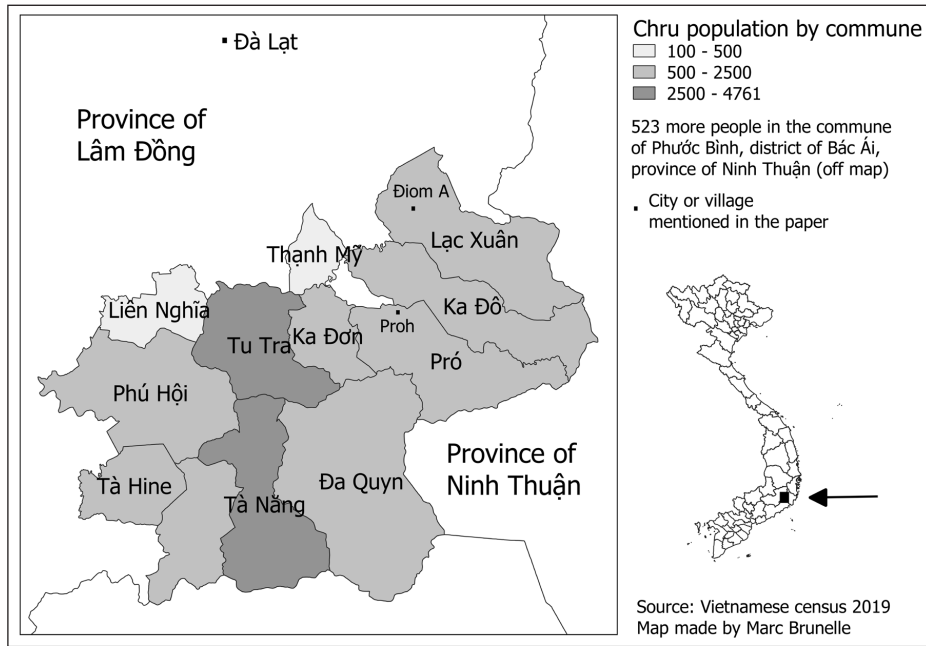
<sup>2</sup> The traditional division of Austroasiatic into Munda and Mon-Khmer has recently been abandoned by most authors (Sidwell, 2009; Sidwell & Rau, 2015). The register languages to which we refer here were previously classified as Mon-Khmer are thus treated as such in most of our references.

an increased longitudinal tension of the vocal folds to suppress voicing (Löfqvist, Baer, McGarr, & Story, 1989; Hoole & Honda, 2011), 2) the lower  $f_0$  after voiced obstruents could be a secondary effect of lowering of the larynx to increase the transglottal pressure differential (Ohala, 1972; Ewan & Krones, 1974; Honda, Hirai, Masaki, & Shimada, 1999; Proctor, Shadle, & Iskarous, 2010; Hoole & Honda, 2011; Solé, 2018), or 3) there could be an auditory association between voicing and a ‘low frequency effect’ (Kingston & Diehl, 1994; Kingston, Diehl, Kirk, & Castleman, 2008). There is also acoustic and perceptual evidence that a lower  $F_1$  follows voiced stops (House & Fairbanks, 1953; Stevens & House, 1956; Stevens & Klatt, 1974; Lisker, 1975; Hillenbrand, Clark, & Nearey, 2001; Esposito, 2002), which could be attributed, again, to tongue-root advancement or larynx lowering to increase the transglottal pressure differential or the low frequency effect (Bell-Berti, 1975; Lindau, 1979; Kingston & Diehl, 1994; Hillenbrand, et al., 2001; Kingston et al., 2008; Brunelle, 2010; Ahn, 2018). Finally, the lengthening of vowels in the vicinity of voiced stops has been explained as an auditory strategy to create the impression of a shorter closure, thus favoring the perception of voicing (Kluender, Diehl, & Wright, 1988).

The association between onset voicing and laxness/breathiness, however, is not as direct. In languages where phonologically voiceless stops are typically aspirated, such as in German, English, and Swedish, vowels tend to be breathier, to variable extents, following voiceless stops than phonologically voiced stops (Löfqvist & McGowan, 1992; Ní Chasaide & Gobl, 1993). In languages in which phonologically voiceless stops are not canonically aspirated, like Italian and French, voiced and voiceless stops do not appear to condition spectral differences on following vowels (Ní Chasaide & Gobl, 1993). By contrast, voiced stops are followed by moderately breathy vowels in Eastern Armenian, a language that has a three-way voicing contrast involving voiced, voiceless, and aspirated stops (Seyfarth & Garellek, 2018). The nature of the articulatory relation between voicing and breathiness is also ill-understood. As far as we know, the only plausible mechanism that has been proposed is that speakers may expand their pharynx to lower supraglottal pressure during voiced obstruents, which would in turn stretch the aryepiglottal folds, forcing a rotation of the arytenoid cartilages and a slight glottal opening (Kingston, Macmillan, Dickey, Thorburn, & Bartels, 1997). Despite the absence of an uncontroversial mechanism linking voicing and breathiness, most diachronic models of registrogenesis (i.e., register formation) propose a stage at which devoiced stops are weakly aspirated or followed by breathy phonation (Haudricourt, 1965; Huffman, 1976; Ferlus, 1979). Some authors go a step further and see this aspiration/breathiness as a necessary precursor to the development of other acoustic properties of register, like  $f_0$ , and vowel quality (Thurgood, 2002; Wayland & Jongman, 2002). However, in the absence of detailed acoustic studies of languages thought to be undergoing registrogenesis, such proposals remain to be tested.

## **1.2. Optional voicing and register in Chru**

Chru (ISO-693: cje) is a language of the Chamic branch of the Austronesian family spoken in the Vietnamese provinces of Lâm Đồng and Ninh Thuận (**Figure 1**). In 2019, there were 23,242 ethnic Chru in Vietnam (Central population and housing census steering committee, 2020). It is probably safe to assume that they almost all speak Chru as a first language and use it in most of their daily interactions, despite widespread bilingualism in Vietnamese. During our two stays in Chru-speaking community, all the interactions we witnessed between community members were conducted in Chru. According to our participants, Chru is also the preferred contact language in interactions with speakers of Koho, a neighboring Austroasiatic language, even if Vietnamese is gradually becoming the lingua franca. Vietnamese is the primary medium of written communication; written Chru is used in church by both Catholics and Protestants, but the proportion of Chru who can read and write fluently in their native language nonetheless remains low.



**Figure 1:** Geographical distribution of ethnic Chru in Lâm Đồng.

**Table 2:** Chru onsets.

p	t	c	k	ʔ
p <sup>h</sup>	t <sup>h</sup>		k <sup>h</sup>	
b	d	ɟ	g	
β	ɗ			
	s		h	
m	n	ɲ	ŋ	
w	l, r	j		

Chru onsets (**Table 2**) have been described as preserving the Proto-Chamic voicing contrast (Phạm, 1955; Lee, 1966; Fuller, 1977). According to previous sources, there is a contrast between plain voiceless and plain voiced stops (there are also implosives, aspirated stops, and sonorants, but they do not contrast in voicing). However, Fuller (1977, p. 85) mentions that Chru “seems to have a non-contrastive feature of register in which the vowel and sometimes the syllable has a lax, breathy quality or a tense, clear quality. Often the breathy quality is concomitant of length in the vowel and voicing in the syllable initial stop.” Based on this description, we set out to investigate the language with the goal of ascertaining if voiced stops condition the expected acoustic properties on following vowels in a language at what may be an early stage of registrogenesis. The complicated relationship between voicing and phonation type mentioned in Section 1.1 is of special interest.

Fuller’s observation that Chru has a non-contrastive register feature also raises the possibility that register has already started to phonologize without becoming contrastive (cf. Hyman, 1976). If it is the case, it could be generalized to non-contrastive contexts (Huffman, 1976). Such a generalization is in fact attested in Cham dialects closely related to Chru, in which there is a process of *register spreading*, i.e., a rightwards propagation of register in sonorant-initial syllables (Friberg & Hor, 1977; Thurgood, 1999). In Formal Eastern Cham, for instance, the second syllable of /ala/ ‘snake’ is realized with a high register, but in /kila/ ‘stupid,’ a word whose initial consonant bears a low register (marked with a subscript dot), the second syllable takes on a low register as well, yielding surface

[kɪla] (Brunelle, 2009b). Register spreading is not limited to Chamic: It is also attested in the orthography of Khmer and has evolved into long-distance vowel alternations in Madurese (Cohn, 1993; Cohn & Lockwood, 1994; Misnadin & Kirby, 2020).

### **1.3. Structured variation and change**

Also important for understanding registrogenesis and transphonologization more generally is to understand how the acoustic distributions and perceptual salience of the individual phonetic cues evolve over time. In classical treatments (e.g., Hyman, 1976), a secondary cue such as  $f_0$  becomes phonologized first, giving rise to a period where the previously primary cue is still produced, but is perceptually redundant. The transphonologization of  $f_0$  in Malagasy (Howe, 2017) and Afrikaans (Coetzee et al., 2018) appears to have followed this pattern. However, it is also possible that the acoustic and/or perceptual relevance of one cue increases proportionally as another decreases, as appears to have been the case in Seoul Korean (Kang, 2014; Bang, Sonderegger, Kang, Clayards, & Yoon, 2018), or even that listeners' attention shifts to a secondary cue in perception before changing in their production (Kuang & Cui, 2018). Here, we focus on three factors: covariation between cues; correlation between acoustic and perceptual salience; and individual differences in the production-perception relationship.

First, is there a compensatory relationship between the presence of onset voicing and the vocalic properties of register? In particular, does there exist an inverse relationship between the frequency and/or temporal extent of closure voicing and spectral properties of register in vowels? Within-category covariation in production between primary and secondary cues has been proposed for other contrasts (Shultz, Francis, & Llanos, 2012; Kang, 2014; Kirby & Ladd, 2016; Howe, 2017; Bang et al., 2018), but has not always been found (Kirby & Ladd, 2016; Clayards, 2018).

Second, is there a correlation between the acoustic distribution and the perceptual salience of the phonetic properties associated with voicing and register? We expect from previous work that the distribution of acoustic and perceptual properties should be closely related: Acoustic properties that have sharp bimodal distributions should be perceptually more salient (Newman, Clouse, & Burnham, 2001; Holt & Lotto, 2006; Clayards, Tanenhaus, Aslin, & Jacobs, 2008; Goudbeek, Cutler, & Smits, 2008; Schreiber, Onishi, & Clayards, 2013). However, cues can be acoustically distinct but still perceptually integrated (Garner & Felfoldy, 1970; Kingston & Diehl, 1994; Kingston & Macmillan, 1995; Kingston et al., 2008; Brunelle, 2009a). For instance, variations in  $F_1$  cannot be entirely disentangled from variations in phonation type in some ranges (Kingston et al., 1997).

Finally, recent studies suggest that one can date a cue shift by comparing the production and perception of individuals. Structured relations between production and perception are not always found (Shultz et al., 2012; Schertz, Cho, Lotto, & Warner, 2015; Brunelle, Hą, & Grice, 2016), but the picture that emerges is that at an early stage, coarticulation and reduction biases alter the relative salience of the phonetic properties associated to a contrast (Ohala, 1989; Beddor, 2009; Bang et al., 2018). Cues that were ancillary then gain perceptual salience in some listeners (Ohala, 1981; Harrington, Kleber, & Reubold, 2008; Beddor, 2009; Kleber, Harrington, & Reubold, 2012; Ohala, 2012; Kuang & Cui, 2018), and in turn trigger a production shift in some individuals. At late stages, both innovative and conservative speakers exhibit some degree of sensitivity to all the relevant cues, until the production shift is completed in the entire community (Pinget, Kager, & Van de Velde, 2016; Howe, 2017; Kuang & Cui, 2018; Pinget, Kager, & Van de Velde, 2019). Although we did not specifically set out to study structured sociolinguistic variation in the speech community, we attempted to balance our speaker sample to the extent possible in order to see what differences, if any, obtain between men and women and between older and younger speakers.

## 1.4. Research questions

We therefore set out to answer the following research questions:

- Q1) How is the voicing/register contrast of Chru realized acoustically? How much individual variation is there in the speech community?
- Is there any evidence that Chru has already developed a register system, as suggested by Fuller (1977)? If register properties are already present, do they correspond to the secondary acoustic properties of voicing expected to be phonologized in register systems?
  - Is there any evidence that the voicing contrast has been neutralized? If register properties are already present, is there evidence that they are, or are becoming, contrastive?
  - If voicing and register still coexist, are they in a compensatory relation? Are the acoustic properties of register more or less distinct when prevoicing is absent?
- Q2) What are the perceptual cues used by Chru listeners to identify the voicing/register contrast? Is there variation across listeners?
- Q3) Do the weights of individual acoustic properties and perceptual cues of voicing/register correlate at the individual or group level? If there is structured variation across individuals, what does it reveal about registrogenesis in Chru and sound change in general?

In order to answer these questions, we undertook both production and perception studies in the Chru community. For practical reasons, data collection was staggered: We gathered the production data in June 2018 and analyzed it first, in order to have sufficient time to design and pilot sensible stimuli for the perception study, which was conducted in June, 2019. We present the production study in Section 2 and the perception study in Section 3.

## 2. The production experiment

### 2.1. Methods

#### 2.1.1. Data collection

Twenty-six speakers of Chru (15 women and 11 men) were recorded in the villages of Điom A and Proh, in the province of Lâm Đồng (about 50 km south of Đà Lạt). We chose to work in Điom A, a village where Fuller also worked in the 1960–70s, to ensure that our data are comparable with his (Fuller, 1977). Proh was selected because we had a reliable contact there. Our speakers were born between 1951 and 2000 (between 18 and 67 at the time of recording), were all highly proficient in Vietnamese, and were all born and raised in the district of Đơn Dương, home to the majority of Vietnam's Chru population.

Participants were presented with selected target words in Vietnamese, had to translate them into Chru, and to produce them four times in the frame sentence in (1).

- (1) /kəw də:m bəh ak<sup>h</sup>a:r \_\_ tə: saʔa:j paŋ/  
 I say CLF word \_\_ for older.sibling hear  
 “I say the word \_\_ for you.”

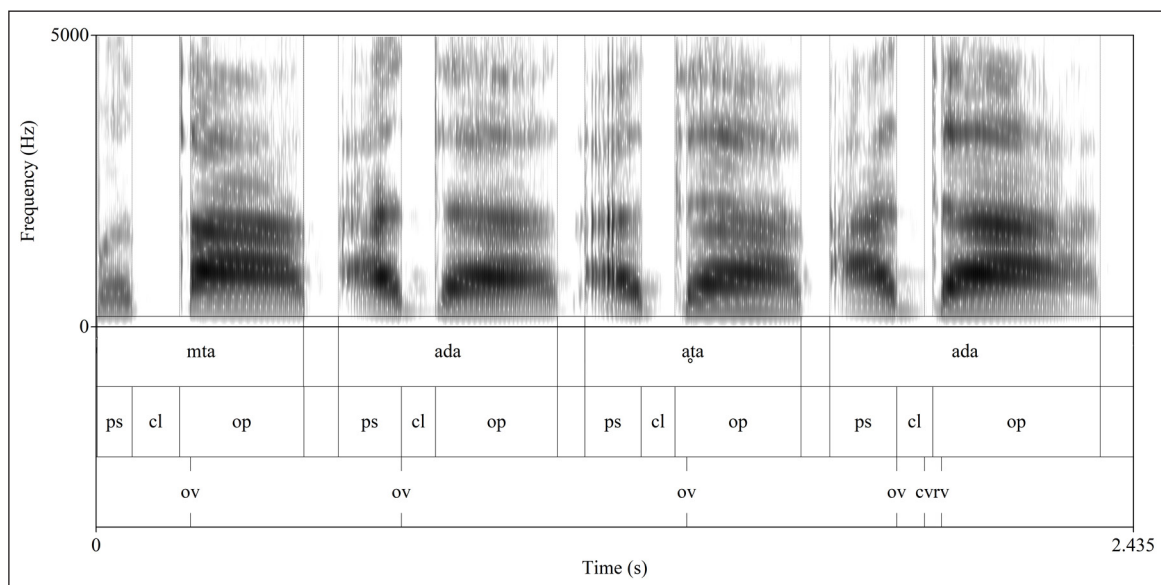
The 60 target words (Appendix A) were chosen because their final stressed syllable (the *main syllable*) contained the coronal or velar onsets /t, d, ɗ, t<sup>h</sup>, s, n, l, r, k, g, k<sup>h</sup>, ŋ/ and the vowels /i:, ɛ:, a:, ɔ:, u:/. We selected open monosyllables when possible (like /ki:/ ‘to comb’), but in most words (44/60), an unstressed syllable (the *presyllable*) preceded the word-final target syllable (as in /pəga:/ ‘hedge, fence’). When codas could not be avoided,

sonorants were preferred (as in /kɔ:ŋ/ ‘bronze’). After excluding 190 production errors and noisy tokens, a total of 6,050 target words were kept. Individual recording sessions took an average of 40 minutes.

Simultaneous audio and electroglottographic (EGG) signals were recorded. Audio recordings were made with a Beyerdynamic 55.18 Mk II microphone connected to a Marantz PMD-660 digital recorder. EGG recordings were acquired through the MATLAB data acquisition toolbox with a Glottal Enterprise EG2-PCX laryngograph connected to a laptop through a National Instrument USB6210 data acquisition device. Three signals were acquired with the EGG: an electroglottograph signal, a larynx height signal, and an audio channel. This paper focuses exclusively on the high-quality audio recordings, which are available from the Pangloss collection ([https://pangloss.cnrs.fr/corpus/list\\_rsc\\_en.php?lg=Chru&name=Chru](https://pangloss.cnrs.fr/corpus/list_rsc_en.php?lg=Chru&name=Chru)) and from Cocoon (<https://doi.org/10.24397/pangloss-0005939>).

### 2.1.2. Acoustic and statistical analysis

Each target word was annotated in a Praat Textgrid. As illustrated in **Figure 2**, three major acoustic landmarks were used to make our measurements: the beginning and endpoint of onset stop closures, onset fricatives or onset sonorants, the beginning and endpoints of the open phase ranging from the burst to the endpoint of the vowel, and the voiced onset time (VOT) associated with stops. VOT was calculated by subtracting the time at the onset of voicing (ov) from the time at the beginning of the open phase (op). When there was no bleeding—i.e., no progressive voicing extending from a previous sonorant without reaching the burst (Davidson, 2016)—ov was set at the beginning of the voice bar or, in the absence of a voice bar, of vowel phonation. If bleeding covered less than the first 50% of the closure, ov was positioned after it ([aʔa] in **Figure 2**). Finally, in a few tokens,



**Figure 2:** Examples of annotation of plain stops in sample recordings made with an 18-year-old woman (born in 2000). Three intervals are labelled in the second tier: ps (preceding sonorant), cl (closure), and op (open phase). Three types of points are labelled in the third tier: ov (onset of voicing), cv (cessation of voicing), and rv (resumption of voicing). The /t/ of the first word, /mta/ ‘eye’, has a voiceless closure and a positive VOT. The stop of the second word, /ada/ ‘duck,’ has three possible realizations: 1) the first instance of [ada] has full closure voicing, 2) [aʔa] exhibits moderate bleeding from the previous sonorant and is thus treated as having a positive VOT, and 3) the second instance of [ada] has a strong voicing that is interrupted just before the closure and is treated as having a negative VOT. More details in the text.



most of the closure was voiced, but voicing stopped shortly before the burst because of the aerodynamic voicing constraint (second instance of [ada] in **Figure 2**). In such cases, *ov* was marked as soon as closure voicing began, but additional annotations were used to mark cessation of voicing (*cv*) and resumption of voicing (*rv*). As only 27 words contained *cv* and *rv* labels, they will not be reported here.

Several types of acoustic measurements were obtained from these landmarks with PraatSauce, a Praat-based application for spectral measures based on VoiceSauce (Shue, Keating, Vicenik, & Yu, 2011; Kirby, 2018): The most relevant are the duration of onset stop closures and vowels, stop VOT, and the  $f_0$ , first two formants, cepstral peak prominence (CPP), and H1-H2 (H1-A1 and H1-A3 were also measured but will not be reported as they do not distinguish registers as clearly as H1-H2 at each 1 ms of the vowel). H1-H2 measures were corrected for formant frequencies and bandwidths, and will thus appear as H1\*-H2\* (Hawks & Miller, 1995; Iseli & Alwan, 2004).

The data and the R script used for data processing (Script 1) are available as Supplementary Material, but some decisions need to be mentioned here. First, since 25 ms windows were used for acoustic measures, the first sampling point reported for each vowel corresponds to a window centered on its 12th ms, thus excluding any influence from the onset. Second, two algorithms were used to remove measurements errors. In order to remove sudden jumps in tracking,  $f_0$ , F1, and F2 were z-normalized per speaker. Derivatives were then computed for consecutive sampling points and all measures with derivatives of  $\pm 0.5$  standard deviations were erased. Then, in order to remove tracking errors over longer time spans, we excluded all  $f_0$ , F1, and F2 values deviating by more than 3 standard deviations from means computed for combinations of subject, vowel, and register. In total, 2.4% of  $f_0$  values, 4.5% of F1 values, and 4.4% of F2 values were removed. All H1\*-H2\* measures derived from excluded  $f_0$ , F1, and F2 values were also deleted.

To facilitate the comparison of acoustic measurements across participants and ensure convergence of statistical models,  $f_0$ , formants, CPP, and H1\*-H2\* measurements were z-normalized by speaker a second time, after removing outliers. As z-scales are difficult to interpret, z-scores were converted back to familiar scales based on means and standard deviations obtained for all speakers in the groups under investigation (mean of all speakers + z-score \* standard deviations for all speakers). These normalized scales are used in figures where data is pooled over groups of speakers and in statistical analyses.

The statistical strength of linguistically meaningful differences was tested with mixed models using the R package *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017). Dependent variables will be indicated where relevant. Unless indicated otherwise, *register*, *place*, and *vowel* were used as fixed factors. All two-way interactions were included; three-way interactions were excluded to avoid overfitting as combinations of the three fixed effects often contained a single word. Random effects included by-*subject* and by-*word* random intercepts. Random slopes were not included as this often resulted in singular (overfitted) models. Maximal models were simplified by dropping non-significant fixed effects if doing so yielded a significantly lower Akaike information criterion (AIC) score. Interactions were dropped before main effects (by decreasing order of F-values in ANOVA model comparisons) and effects were not dropped if they were a subset of a significant interaction. In the main text, we focus on visual displays and discussion of the most relevant model parameters; see Appendix B for the fixed effect estimates and the Supplementary Material for the R code and data files.

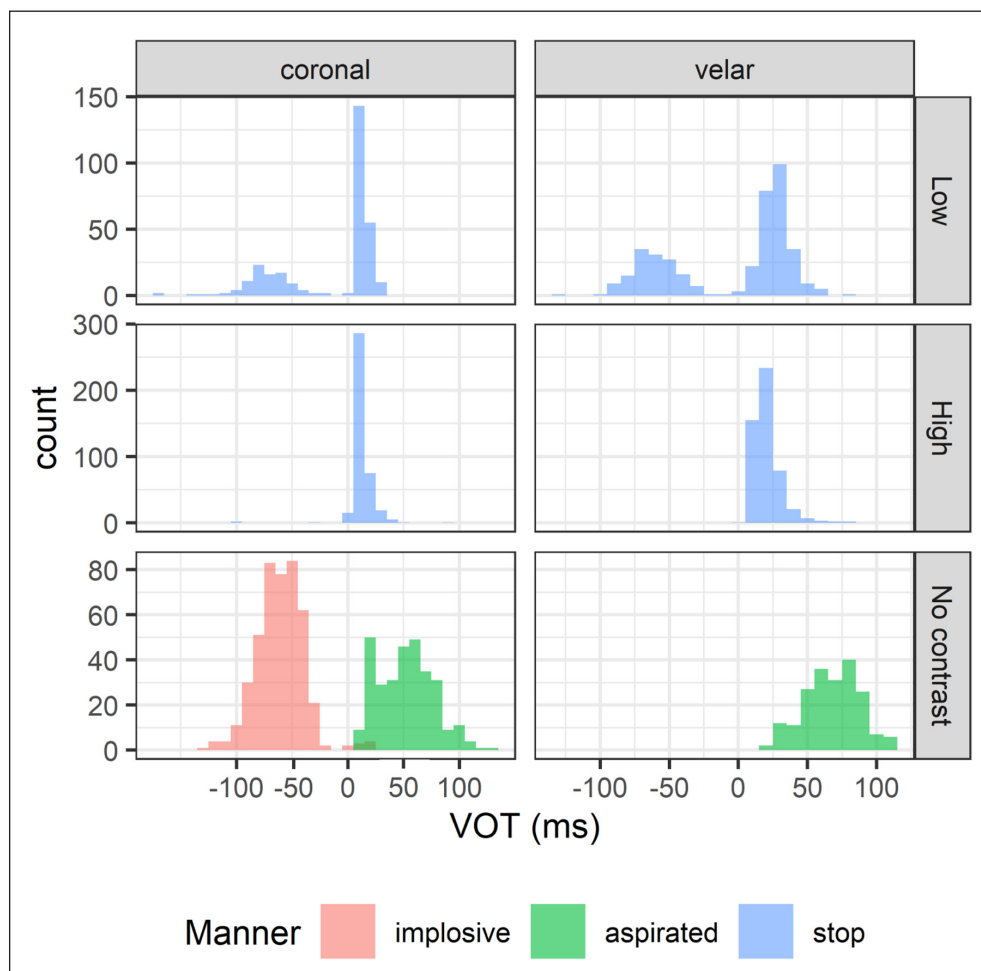
## 2.2. Results

### 2.2.1. Onsets

The top row of **Figure 3** reports VOT values for the series of stops described as ‘voiced’ in previous sources (Phạm, 1955; Lee, 1966; Fuller, 1977). Contrary to expectations, they have a bimodal VOT distribution: They are sometimes voiced, but more often voiceless. To avoid confusion, we will relabel them *low-register stops*, and will characterize them as *prevoiced* when they have a negative VOT, and as *devoiced* when they have a positive VOT. The second row corresponds to the series previously described as voiceless stops, which we will refer to as *high-register stops*. High-register stops have a unimodal distribution, centered around a 13 ms positive VOT.

A closer look at low-register stops reveals that the VOT of devoiced coronal stops is comparable to that of high-register coronal stops, but that devoiced velar stops have a slightly longer VOT than their high-register counterparts (RegisterLow:  $\beta = -1$  ms,  $t = -.799$ ,  $p = .507$ ; RegisterLow:Placevelar:  $\beta = 5$  ms,  $t = 3.714$ ,  $p = .062$ ; see Table 1 in Appendix B).

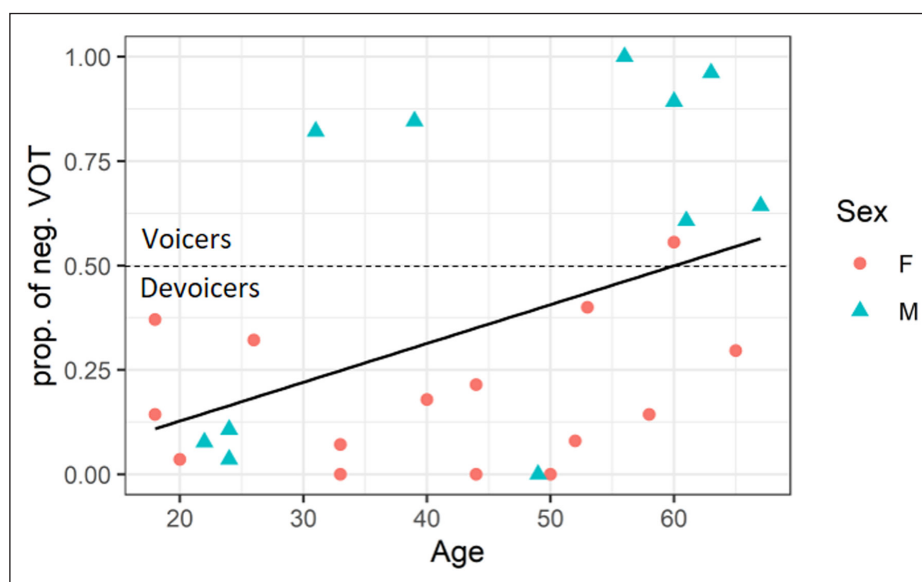
The distribution of VOT in **Figure 3** hides important interspeaker variation in low-register stops. A breakdown by speaker is given in **Figure 4**, where the proportion of prevoiced low-register stops is plotted by age and sex. Seven out of 11 men prevoice more than 50% of their low-register stops, but only one out of 15 women does. There also seems



**Figure 3:** VOT of onset stops. Plain stops are split into high register (<voiceless) and low register (<voiced). Implosives and aspirates are given for comparison.

to be an age gradient, illustrated by the significant regression line in **Figure 4** ( $r = 0.45$ ,  $p = 0.02$ ): Younger speakers have lower proportions of prevoicing than their elders. Since this could be evidence for a change in progress, differences between two groups will be tracked in the rest of Section 2.2: Speakers will be split into *Voicers* (8/26 participants), who prevoice 50% or more of their low-register stops, and *Devoicers* (18/26 participants), who devoice them more than 50% of the time.

Mean closure duration for high- and low-register stops is reported in **Table 3**. Devoiced velar stops have a shorter closure duration than their high-register counterparts in both Voicers and Devoicers (see statistical models in Table 2 of Appendix B). The last row of **Table 3** also shows that Devoicers have comparable proportions of devoicing in low-register coronals and velars (87.6% versus 84%), but that Voicers devoice coronals more often than velars (33.3% versus 17%). Temporal measures obtained from onset stops thus suggest that the original voicing contrast is no longer consistently realized in Chru onset stops: Closure voicing has become optional (even for Voicers), and closure duration differences between high- and low-register stops are now restricted to velars.



**Figure 4:** Proportion of low-register stops produced with closure voicing per speaker, by sex and age.

**Table 3:** Mean closure duration in plain stops (ms); standard deviations in parentheses. Low-register stops can preserve closure voicing (negative VOT) or be realized with a positive VOT.

Voicers (8/26)

Place	Coronals			Velars		
Register	High	Low		High	Low	
VOT	Positive	Positive	Negative	Positive	Positive	Negative
Mean (ms)	74 (20)	76 (43)	75 (24)	93 (44)	55 (20)	61 (16)
n	125	35	70	150	21	102

Devoicers (18/26)

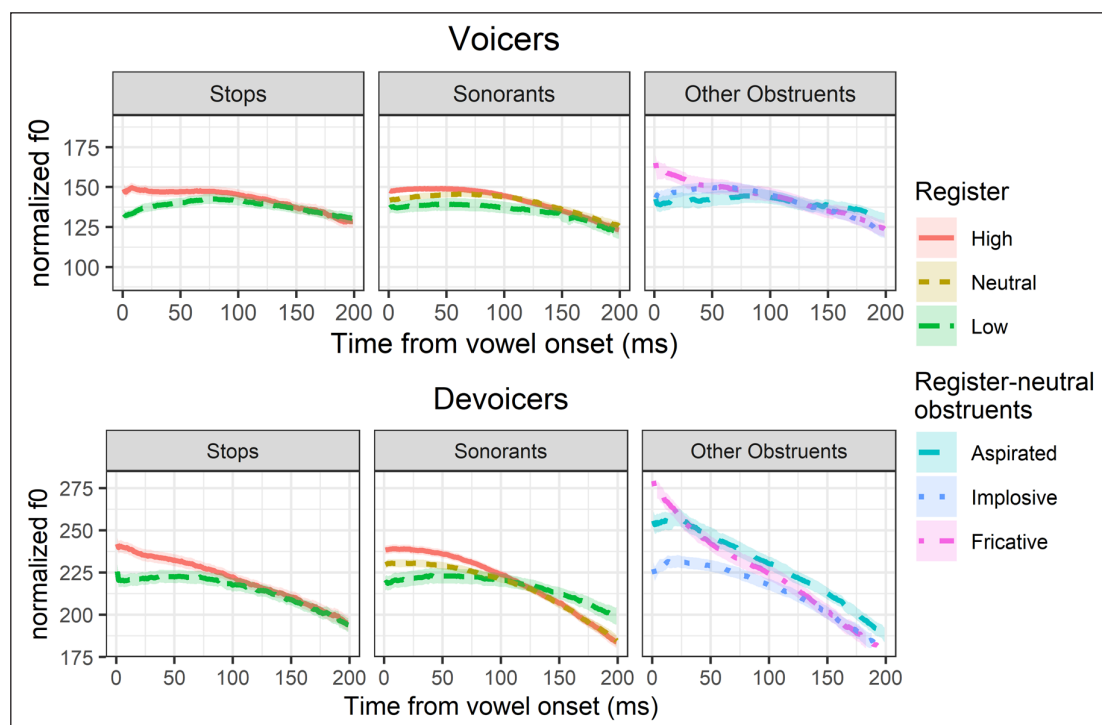
Place	Coronals			Velars		
Register	High	Low		High	Low	
VOT	Positive	Positive	Negative	Positive	Positive	Negative
Mean (ms)	90 (41)	86 (51)	86 (22)	113 (62)	65 (28)	64 (20)
n	276	185	26	352	231	44

### 2.2.2. Vocalic properties of register

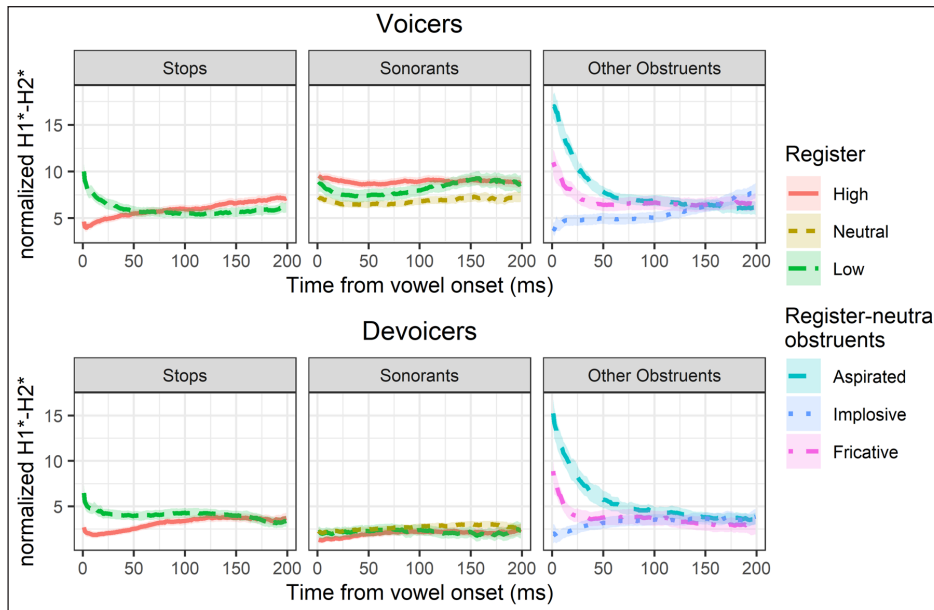
Vowels following low-register stops are longer than those following high-register stops in both Voicers and Devoicers (Voicers: 308 ms versus 251 ms; Devoicers: 269 ms versus 249 ms). However, this effect is weak in Voicers and negligible in Devoicers (Voicers: RegisterLow  $\beta = 75$  ms,  $t = 3.059$ ,  $p = .095$ ; Devoicers: RegisterLow  $\beta = 20$  ms,  $t = .417$ ,  $p = .714$ ; Table 3 in Appendix B).

F0 is not a reliable acoustic indicator of register either. In **Figure 5**, it appears higher after high-register than after low-register stops, but in Voicers, this effect is largely circumscribed to the vowel /i:/ (RegisterLow:Vowel1  $\beta = -43$  Hz,  $t = -3.156$ ,  $p = .016$ ), while in Devoicers, the effect of Register is not robust enough to be included in the final statistical model (see Table 4 in Appendix B). The vowels of sonorant-initial syllables following presyllables headed by high-register stops also have a higher f0 than those following presyllables headed by low-register stops (sonorant-initial monosyllables, which are labelled as ‘register neutral’ in **Figure 5** as they do not contrast in voicing and should not be affected by register spreading, fall in between), but this coarticulatory effect is again weak and limited to some combinations of vowels and places (see Table 5 in Appendix B). Finally, vowels headed by obstruents that do not contrast for register (implosive, fricative, and aspirated) all have high f0 contours, f0 following fricative /s/ being the highest. The fact that the f0 effects visible in **Figure 5** are statistically weak can be attributed to individual variation in the f0 patterns of both Voicers and Devoicers (discussed further in Section 2.2.3 below).

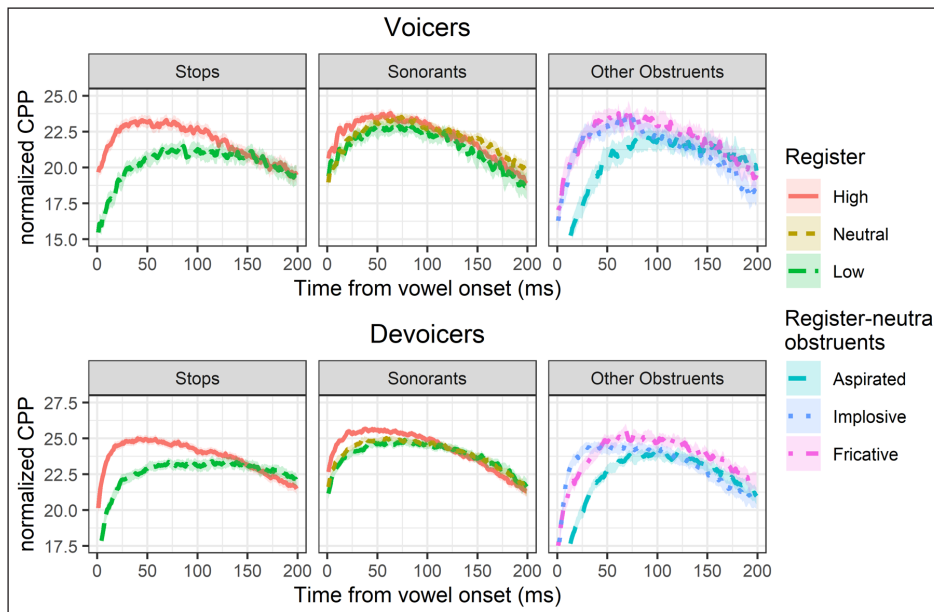
Out of the three spectral slope indicators that were analyzed, only normalized H1\*-H2\* will be reported, as it shows the strongest effect. In **Figure 6**, register conditions a robust H1\*-H2\* difference at vowel onset, indicating a breathier phonation in the low register (Voicers: RegisterLow  $\beta = 3.25$  dB,  $t = 4.248$ ,  $p = .027$ ; Devoicers: RegisterLow  $\beta = 9.58$  dB,  $t = 5.151$ ,  $p = .002$ , with much weaker effects in close vowels; see Table 6 in Appendix B). This effect of register on H1\*-H2\* is limited to the beginning of the vowel



**Figure 5:** Normalized f0 in the first 200 ms of the vowel, after different classes of onsets. Ribbons represent the 95% CI of the mean.



**Figure 6:** Normalized H1\*-H2\* in the first 200 ms of the vowel, after different classes of onsets. Ribbons represent the 95% CI of the mean.



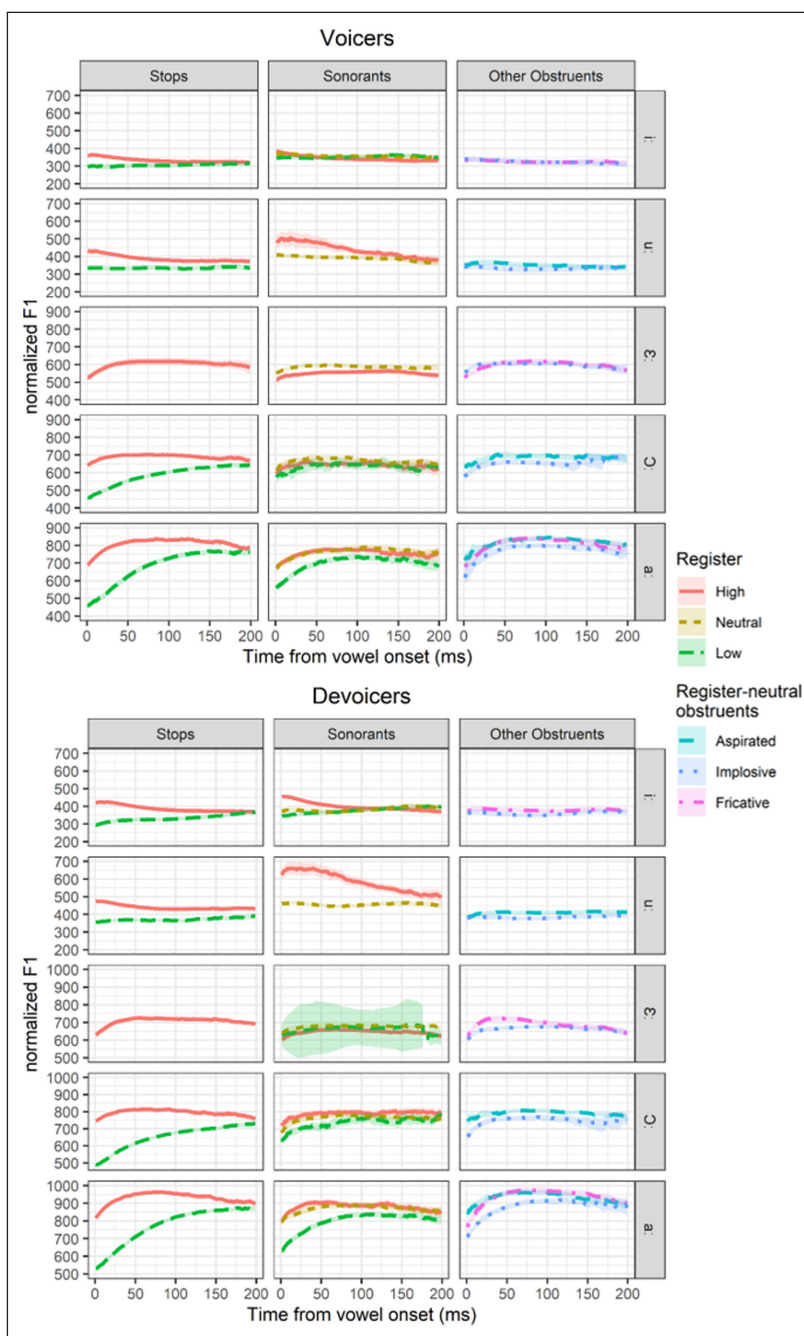
**Figure 7:** Normalized CPP in the first 200 ms of the vowel, after different classes of onsets. Ribbons represent the 95% CI of the mean.

and does not extend to sonorant-initial syllables (see Table 7 in Appendix B). Finally, while fricatives and aspirates are associated with a high H1\*-H2\* because of their glottal opening, implosives are followed by a low H1\*-H2\* caused by their glottal closure.

Normalized CPP, shown in **Figure 7**, is another indicator of phonation type. It is expected to be high when phonation is modal, and to be low when phonation is non-modal. As such, the systematic rise in CPP at the beginning of vowels is an indication that phonation is perturbed by onsets. CPP is consistently lower after low-register than high-register stops in Voicers (RegisterLow  $\beta = -3.95$  dB,  $t = -8.286$ ,  $p < .001$ ), but this effect is not as robust in Devoicers (RegisterLow  $\beta = -3.74$  dB,  $t = -2.324$ ,  $p = .137$ ). CPP does not start as low in sonorants as in stops because they are not produced with a spread or constricted

glottis, but the register of presyllables nonetheless exerts a weak coarticulatory effect on sonorant-initial syllables (see Table 7.2.9). Vowels following register-neutral obstruents all start with a low CPP because their onsets are either produced with an open (aspirates and fricatives) or a closed glottis (implosives), two settings that favor non-modal phonation.

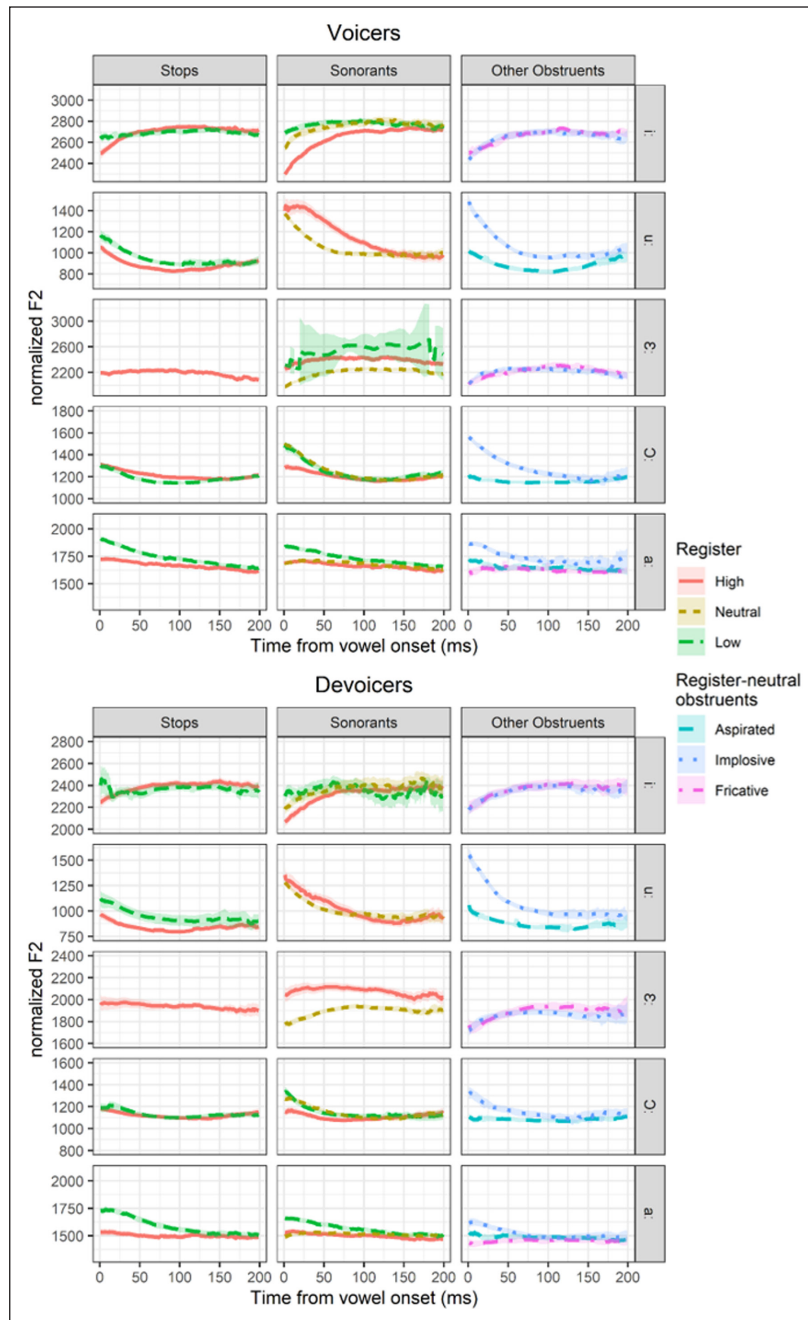
F1 is plotted by vowel in **Figure 8**. Vowels have a lower F1 immediately after low-register than high-register stops (Voicers: RegisterLow  $\beta = -166$  Hz,  $t = -9.743$ ,  $p < 0.001$ . Devoicers: RegisterLow  $\beta = -240$  Hz,  $t = -9.220$ ,  $p < 0.006$ ). However, the effect of register is much smaller in close vowels (see interactions of Register and



**Figure 8:** Normalized F1 in the first 200 ms of the five vowels, after different classes of onsets. The range of the y-axis is kept constant across vowels (500 Hz). Ribbons represent the 95% CI of the mean. The large confidence interval for /ɛ:/ sonorants in Devoicers is due to the small number of tokens as the intended target word /bəŋɛ/ was produced with the high register by most speakers.

Vowel in Table 10 in Appendix B). The difference between registers diminishes during the production of the vowel, but is maintained for at least 100 ms, longer than for any other spectral property. The coarticulatory influence of presyllables on the F1 of sonorant-initial syllables is weak at best (see Table 11 in Appendix B). There is little difference between vowels following register-neutral obstruents.

In **Figure 9**, the effect of register on F2 following stops is weak (statistical results are given in Table 12 of Appendix B). The formants of sonorant-initial syllables and syllables headed by register-neutral obstruents are highly variable and show no robust statistical pattern (see Table 13 in Appendix B).



**Figure 9:** Normalized F2 in the first 200 ms of five vowels, after different classes of onsets. The range of the y-axis is kept constant across vowels (700 Hz). Ribbons represent the 95% CI of the mean. The large confidence interval for /e:/ sonorants in Voicers is due to the small number of tokens as the intended target word /bəŋɛ/ was produced with the high register by most speakers.

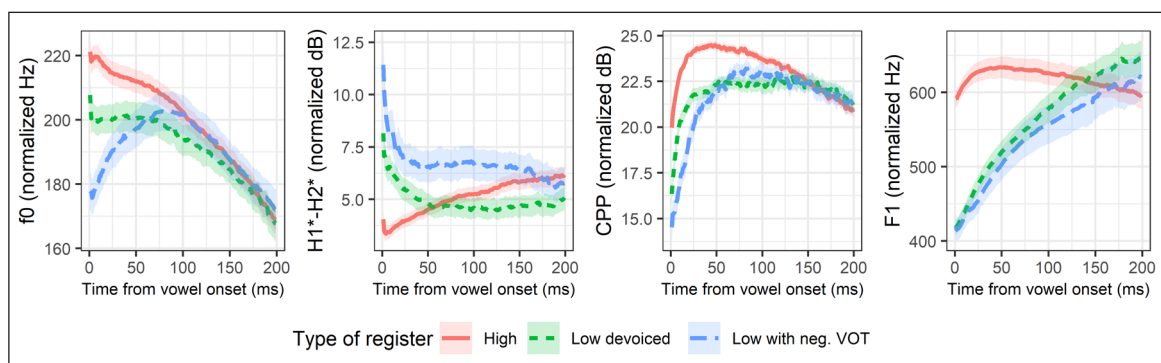
Inspection of the spectral properties of register does not reveal important qualitative differences between Voicers and Devoicers, contrary to what was found for VOT. The register contrast is primarily realized by differences in F1, especially (but not exclusively) in open and open-mid vowels. Phonation ( $H1^*$ - $H2^*$  and CPP) is also a consistent indicator of register. On the other hand, registers do not differ consistently in terms of  $f_0$  or F2. The register of presyllables has a coarticulatory effect on syllables headed by sonorants, but this effect is smaller than that found in syllables headed by stops, suggesting that it is not categorical register spreading.

### 2.2.3. Relation between closure voicing and vocalic properties of register

A breakdown of syllables headed by low-register plain stops by phonetic voicing (devoiced versus prevoiced), plotted in **Figure 10**, shows that out of the four acoustic properties that are conditioned by register,  $f_0$  and  $H1^*$ - $H2^*$  (and possibly CPP) are more distinct from the high register when onsets are *prevoiced* than when they are *devoiced*. If speakers enhanced register cues to compensate for the lack of prevoicing, we would expect these properties to be enhanced to a greater extent when prevoicing is absent. Crucially, F1, which was arguably the most robust register property in Section 2.2.2, behaves differently, in that it seems equivalent after prevoiced and devoiced stops. As there are relatively few tokens of low-register stops in the dataset (an average of 26.2 per speaker) and as the frequency of closure voicing is highly variable across speakers (see **Figure 4**), there is too little data for a meaningful statistical analysis comparing Voicers and Devoicers, but we note that similar results are obtained when the same figure is only plotted with the four speakers that have the most balanced proportions of prevoiced and devoiced stops.

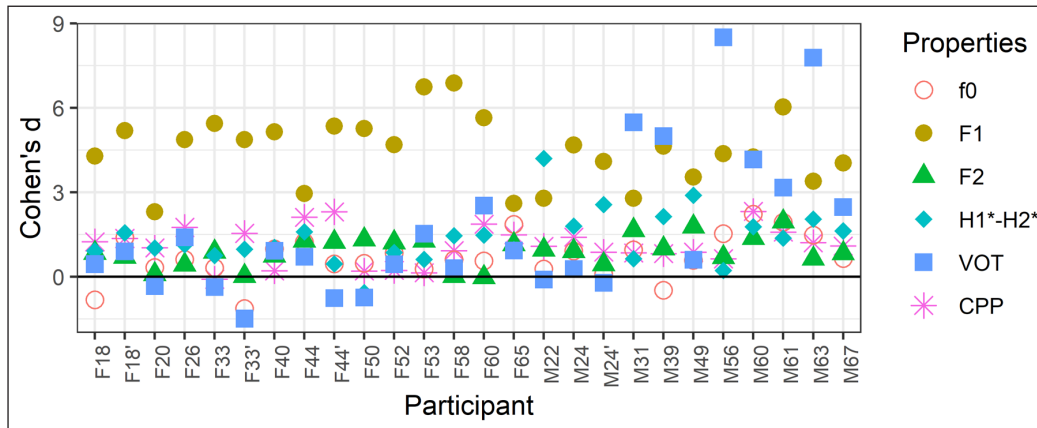
### 2.2.4. Individual variation

Although we did not find significant qualitative differences between Voicers and Devoicers in Section 2.2.2 at the group level, the normalized mean values presented therein conceal non-negligible individual variation in each group. It is therefore essential to consider individual behavior, and more especially the magnitude of the difference between the two registers for each speaker and acoustic property. We did this by computing Cohen's *d* (Cohen, 1988), an effect size indicator, for individual speakers and relevant acoustic properties. We calculated Cohen's *d* as the vowel-weighted difference between the means of the two registers at the first sampling point after plain stops, divided by the pooled register-weighted mean of their standard deviations. While Cohen's *d* is simple to compute, we note that the scores must be interpreted with some caution, as this measure does not take into account possible correlations between cues.



**Figure 10:** Effect of closure voicing on some vocalic properties of the low register. High register given as a reference. Ribbons represent the 95% CI of the mean. Values are speaker-normalized.





**Figure 11:** Individual variation in the use of various properties of register after plain stops. Cohen’s *d* scores express distance between registers. Cohen’s *d* for H1\*-H2\* and F2 have been multiplied by -1 so that negative scores represent differences going in the unexpected direction. Subject labels encode sex and age in 2018. The prime symbol is used when two participants have the same age and sex (as in M24’).

**Figure 11** plots Cohen’s *d* scores for the six properties of register reported in Section 2.2.2. Scores for H1\*-H2\* and F2 have been multiplied by -1 so that positive scores represent differences going in the expected direction (in light of the literature on register and of the results of Section 2.2.3). For instance, six speakers have Cohen’s *d* scores below 0 for VOT, indicating that they have a longer positive VOT after low-register stops than after high-register stops, contrary to the majority of speakers.

The most striking regularity in **Figure 11** is that women distinguish their registers primarily in terms of F1, only using other properties to a limited extent. Men’s productions, on the other hand, tend to be distinguished by a wider range of cues. They largely base their register contrast on F1, like women, but many also have high Cohen’s *d* for VOT, as well as slightly higher scores for H1\*-H2\*.

As was already foreshadowed in Section 2.2.2, Cohen’s *d* scores for f0 vary unexpectedly across speakers. They tend to be close to 0, but while some speakers have a distinctly higher f0 in the high register (like most older men on the right of **Figure 11**), others have a consistently higher f0 in the low register, like F18, F33’, and M39.

**2.3. Summary of acoustic results**

Our acoustic results confirm Fuller’s (1977) intuition that Chru has already developed a register contrast and that the vocalic properties of this register system are analogous to those found in other Austroasiatic and Chamic register languages. After stops, F1 differences between registers are robust, last over the first 100–150 ms of the vowel, and tend to be greater in non-close vowels. Non-close vowels have a moderate falling on-glide in the low register: In the clearest cases, low register /a:/ and /ɔ:/ sound like [ʔa] and [ʔɔ]. Close vowels, on the other hand, have a slightly higher F1 at vowel onset in the high register, but this effect is never large enough to be heard as an onglide. H1\*-H2\* and CPP results also indicate that there is a moderate but consistent lax to breathy phonation in the initial 50 ms of low register vowels. Differences in f0 are also found, but are weaker and vary significantly across speakers. Besides the greater diphthongization of open vowels in the low register, which is common in Austroasiatic (Jenner, 1974; Ferlus, 1979; Huffman, 1985), the realization of register does not seem to vary significantly across vowels, contrary to what was found in Southern Yi (Kuang

& Cui, 2018). H1\*-H2\* effects appear weaker in close vowels in Devoicers, but no such effect is found in Voicers. Register also seems to have a greater effect on f0 in vowel /i:/, but not in other vowels.

What Fuller seems to have overlooked is that for many speakers, register may have already taken over the contrastive role of voicing. Most speakers (18/26) have prevoicing in less than 50% of ‘voiced stops,’ and only one speaker preserves it in all low-register stops. Synchronically, Chru therefore seems to have a register system combined with optional prevoicing, rather than a voicing contrast with redundant register. Since some speakers with significant devoicing were already in their mid-20s when Fuller conducted his field research in the late 1960s, the fact that he did not describe this distribution cannot be attributed to a dramatic devoicing in the past half-century; it is probable that devoicing was already present in the community, but that it was overlooked because Fuller was mostly working with older men.

Another important observation concerning onsets is that the large majority of speakers show no evidence of aspiration in devoiced low-register stops. There is evidence that low-register stops have a longer VOT than high-register stops when they lose their closure voicing, but this effect is limited to velar stops and is probably too small to be audible (5.5 ms). Moreover, a small number of speakers (6/26) have a marginally longer VOT in low-register than high-register stops, as can be seen in **Figure 11**. While this is reminiscent of the aspiration (or ‘breathy release’) that is a central keystone of most models of register formation (Haudricourt, 1965; Huffman, 1985; Thurgood, 2002; Wayland & Jongman, 2002), the small VOT difference found in some Chru speakers is certainly weaker and less systematic than expected. That F1 appears to be the primary acoustic correlate of register for all Chru speakers in our sample, despite the fact that only a minority shows signs of lengthened VOT, leads us to doubt the claim that aspiration is a necessary step in registrogenesis.

Speaker averages hide a certain amount of individual variation in the realization of register. In women, F1 is the best-defined acoustic property of register, with Cohen’s *d* scores about five times greater than those of any other property. F1 is also a strong distinctive property among men, but it is less dominant: Many men also maintain large VOT and H1\*-H2\* differences, and a few have non-negligible Cohen’s *d* scores for f0, CPP, and F2. However, when we look at individuals rather than groups, there is no significant compensatory relation between the weight of the various acoustic properties (for instance, there is no inverse relation between the use of voicing and the use of F1 to contrast registers). Furthermore, with the notable exception of F1, differences in f0, H1\*-H2\*, and CPP are more pronounced when stops are prevoiced than when they are devoiced. We interpret this as evidence that F1 is the primary, obligatory, acoustic property of the register contrast, but that other properties can be enhanced in clear speech contexts, where prevoicing is also most likely to be present.

Additional evidence about the phonological status of Chru register can be gathered from syllables in which register is non-contrastive. An inspection of the rightwards propagation of register properties from presyllables to syllables headed by sonorants reveals that there appears to be weak coarticulatory effects in f0, CPP, and F1, but that these effects are not indicative of the type of categorical spreading reported in Cham dialects or in Khmer. This shows that even if register has become contrastive in Chru, it is neither involved in productive phonological alternations nor generalized to sonorants. The second type of evidence comes from the acoustic properties of register-neutral obstruents (aspirated stops /t<sup>h</sup>, k<sup>h</sup>/, implosive /d/, and fricative /s/). The acoustic properties

of vowels following these obstruents do not clearly pattern with a specific register. Vowels following aspirated stops, for example, have a high initial  $f_0$  reminiscent of the high register, but their high  $H1^*-H2^*$  and low CPP at voicing onset indicate that they are breathier than the low register. Along the same lines, vowels following the implosive /ɖ/ start with a high  $f_0$  and a low  $H1^*-H2^*$ , just like vowels following high-register plain stops, but are initiated with a low CPP and a high  $F_2$ , which is more similar to the low register. This suggests that register-neutral obstruents are not forced into phonologized register categories, but maintain their own idiosyncratic phonetic properties.

### 3. The perception experiment

In order to determine if the acoustic differences between registers uncovered in the production experiment are meaningful to Chru listeners, we returned to Ìiom A and Proh to conduct a perception experiment in June 2019.

#### 3.1. Methods

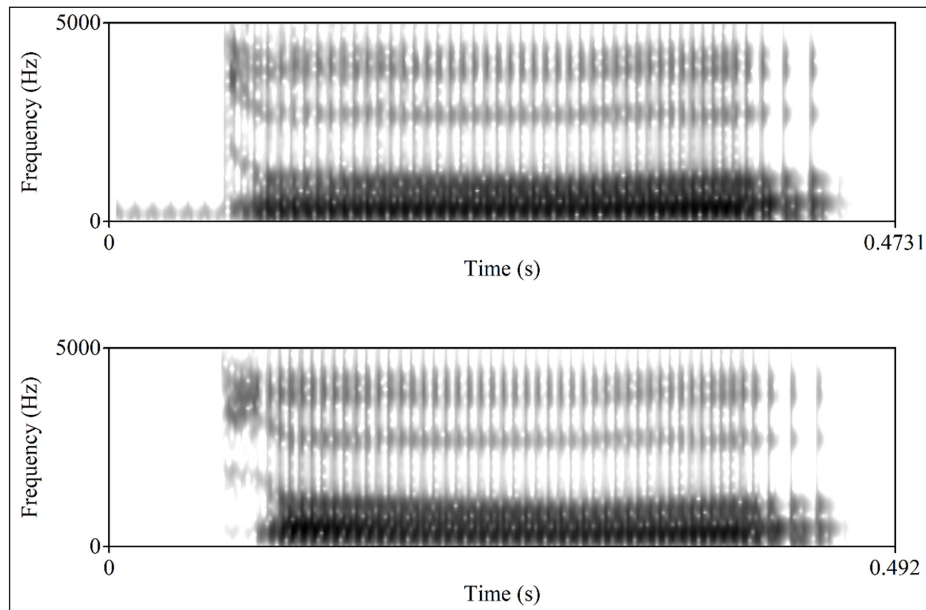
The perception experiment was conducted with 41 listeners (21 women); two additional participants were excluded because they could not complete the task. They were all born between 1950 and 2001 (18 to 69 years old at the time of the experiment) in the district of Đơn Dương and raised there. Nineteen of these listeners had participated in the production experiment the previous year.

All listeners took part in a forced choice identification task in which they had to listen to synthesized stimuli varying in  $F_1$ , phonation type,  $f_0$ , and VOT, and to identify them by pressing one of two keyboard buttons associated to images presented on a computer screen. For one set of stimuli, they had to choose between /mta/ ‘eye’ and /mda/ ‘rich’ (where /a/ does not contrast for length but is phonetically long, like all Chru vowels in open syllable) while for the other, the choices were /tu:ʔ/ ‘bamboo joint, section’ and /du:ʔ/ ‘honey bee’ (where /d/ can be realized as [d] or low-register [t]). The experiment was run in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). Instructions were largely visual because few of the participants were fully literate in Chru.

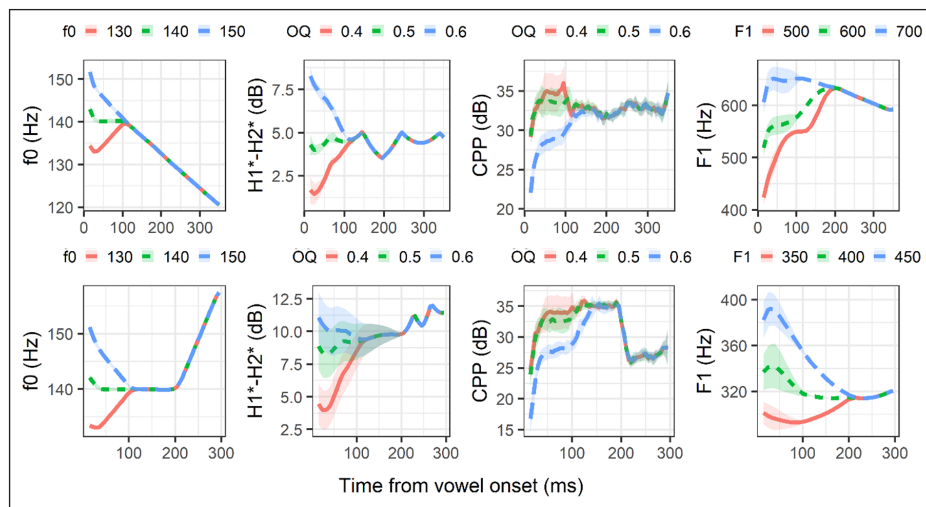
Ideal stimulus pairs would have consisted of open syllables without presyllables. However, such minimal pairs are rare in Chru, and can be difficult to represent visually (one has to exclude function words and most abstract lexical words). The selection of our two minimal pairs is thus based on the assumption, backed up with acoustic evidence, that the final glottal stop of /tu:ʔ~du:ʔ/ only affects spectral balance towards the end of the vowel and that the effect of the nasal presyllable of /mta~mda/ on the spectral tilt of the vowel is blocked by the onset of the main syllable (see Styler, 2017, for an overview of the effects of nasality of spectral balance).

Stimuli were synthesized using KlattGrid synthesis in Praat (see Scripts 3 and 4 in the Supplementary Material). Parameters were set in such a way as to imitate natural tokens without including superfluous low-level variation. Spectrograms of sample resynthesized stimuli are given in **Figure 12**. The four parameters that were manipulated across stimuli are  $f_0$ , open quotient,  $F_1$ , and VOT. For each parameter, maximum and minimum values were selected based on the acoustic results presented in Section 3, and fine-tuned to maximize naturalness based on the natural productions of a middle-aged male speaker.<sup>3</sup>

<sup>3</sup> As a reviewer pointed out, it is possible that the male-sounding voice favored the use of cues associated with men by listeners. If this is correct, VOT, the property most affected by gender in production, would be expected to have a smaller perceptual weight with a female voice.



**Figure 12:** Examples of resynthesized stimuli. Top panel: Resynthesized low register syllable /du:ʔ/, with a voiced onset, an initial F1 of 350 Hz, an initial f0 of 130 Hz, an initial open quotient of 0.6, and 60 dB of added breathiness amplitude. Bottom panel: Resynthesized high register syllable /tu:ʔ/, with a 20 ms VOT, an initial F1 of 450 Hz, an initial f0 of 150 Hz, an initial open quotient of 0.4, and no added breathiness amplitude.



**Figure 13:** Mean values of the key acoustic parameters of the stimuli in the identification experiment. Top panel: /mta~mda/; Bottom panel: /tu:ʔ~du:ʔ/. The ribbons correspond to one standard deviation above and below the mean (due to interactions between the various parameters, the acoustic properties of stimuli were not always exactly on target).

The four parameters were crossed so as to obtain all possible combinations of acoustic properties (**Figure 13**, plotted with Script 5 provided in the Supplementary Material). The four parameters were manipulated as follows:

- VOT: Three VOTs were synthesized: -50 ms, 10 ms, and 20 ms. The 20 ms VOT is exaggerated compared to values observed in Chru and was included to test the cross-linguistic hypothesis that there is a relation between obstruent devoicing and aspiration.

- f0: An initial pitch target was set at the beginning of the vowel, with three possible values: 130, 140, and 150 Hz. A fixed target was set to 140 Hz at 100 ms into the vowel for all stimuli.
- Phonation type (Open Quotient): an initial open quotient was set at the beginning of the vowel, with three possible values: 0.4, 0.5, and 0.6. A fixed target was set to 0.5 at 100 ms into the vowel. In order to increase naturalness and to modulate the CPP variation found in the production experiment, breathiness amplitude (BA) was added to the first 100 ms of the vowel: 60 dB of BA were added to the tokens with an initial open quotient of 0.6 and 30 dB to those with an open quotient of 0.5. No BA was added to tokens with an initial open quotient of 0.4.
- F1: An initial F1 target was set at the beginning of the vowel, with three possible values: 350, 400, and 450 Hz for the vowel /u:/ and 500, 600, and 700 Hz for the vowel /a:/. The differences in F1 steps between the two vowels mirror those found in natural tokens. Fixed targets were set to 400 Hz for /u:/ and 700 Hz for /a:/ at 200 ms into the vowel. Other formants were kept constant across tokens.

The resulting stimuli sounded fairly natural, even if they sometimes combined acoustic parameter values that do not naturally cooccur in Chru. As far as we know, no participant came to realize that the stimuli had not been recorded from a real speaker and many participants asked us who the speaker was.<sup>4</sup>

Before proceeding to the testing phase proper, all participants had to undergo 1) a training phase with the two tokens closest to natural productions (6 tokens per word pair), 2) a first test phase with the same stimuli (10 tokens for each word pair) and 3) a second test phase with random stimuli (10 tokens for each word pair). They had time to rest and ask for clarification between each block. The real testing phases then started. Stimuli were presented in six blocks. There were three blocks per word pair, alternating between /mta~mda/ and /tu:~du:~/ (henceforth a- and u-stimuli). Each block contained all 81 randomized stimuli for the relevant word pair, for a total of 486 tokens per listener. The entire experiment took 15–25 minutes per participant. Responses with reaction times above 2 seconds were not recorded.

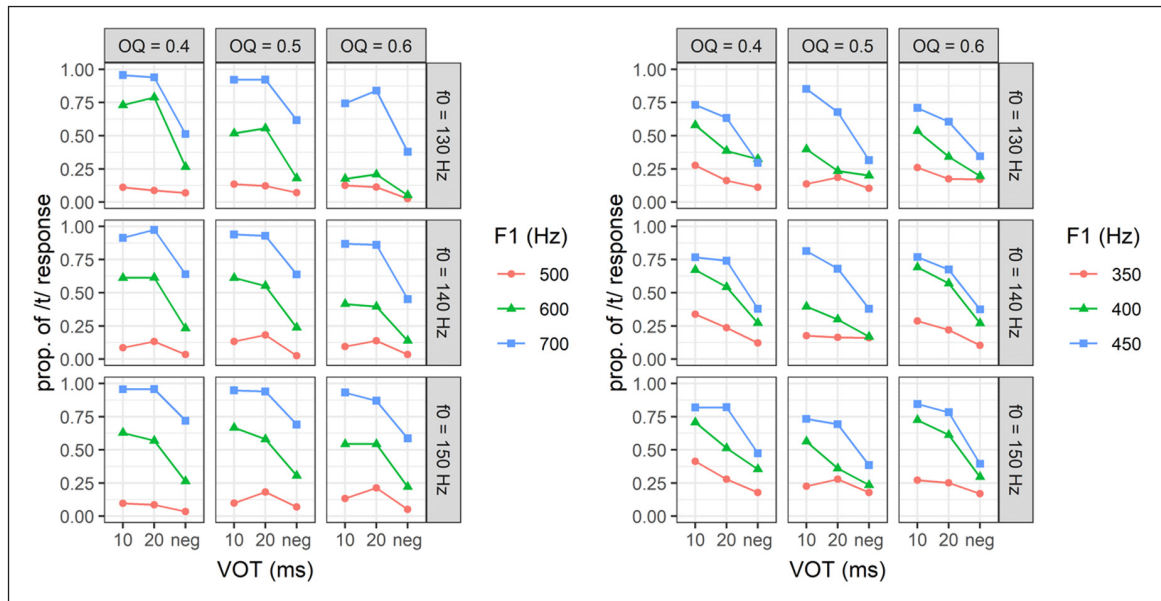
Responses were analyzed using mixed logistic regressions with the R package lmerTest (Kuznetsova et al., 2017). The fixed effects were the f0, F1, OQ, and VOT of the synthesized stimuli, where f0, F1, and OQ were treated as centered and ordered categorical factors, and VOT was treated as an unordered categorical factor. Maximal models included all two-way interactions of fixed effects, and random effects included random intercepts for Subject and random slopes combining Subject and main fixed effects. Models were then simplified in a stepwise manner by dropping the fixed effect or interaction with the lowest F-value as long as this did not significantly increase the Akaike information criterion (AIC) score of the model.

## 3.2. Results

### 3.2.1. Identification experiment

The pattern of responses obtained from the listeners are reported in **Figure 14**, and statistically summarized in **Tables 4** and **5** (response data are available in the Supplementary Material). Note that the intercepts of the final models are not significantly different from 0, which means that we were successful in generating stimuli scales in which middle values have neither a high or low register bias. For both synthesized

<sup>4</sup> As pointed out by a reviewer, this is not a guarantee of naturalness, but it indicates that the signal was probably as good as that of the signals that participants are used to hearing from their cell phones and TV and radio receptors.



**Figure 14:** Proportion of high register /t/ responses for each type of stimulus, by F1, VOT, f0, and open quotient (OQ), averaged over all listeners. Left panel: a-stimuli. Right panel: u-stimuli.

**Table 4:** Table of estimates of the final logistic regression model for a-stimuli. Estimates represent the log odds of high register responses. F0, F1, and OQ are centered.

	Estimate	SE	z value	Pr(> z )
(Intercept)	0.162	0.171	0.946	0.344
F1	2.884	0.176	16.433	<0.001
f0	0.273	0.058	4.678	<0.001
OQ	-0.457	0.056	-8.227	<0.001
VOT 20 ms	0.043	0.091	0.469	0.639
VOT neg	-2.190	0.315	-6.950	<0.001
F1:f0	0.175	0.058	3.013	0.003
F1:OQ	-0.317	0.058	-5.423	<0.001
F1:VOT 20 ms	-0.173	0.114	-1.519	0.129
F1:VOT neg	-0.325	0.132	-2.454	0.014
f0:OQ	0.358	0.048	7.522	<0.001

**Table 5:** Table of estimates of the final logistic regression model for u-stimuli. Estimates represent the log odds of high register responses. F0, F1, and OQ are centered.

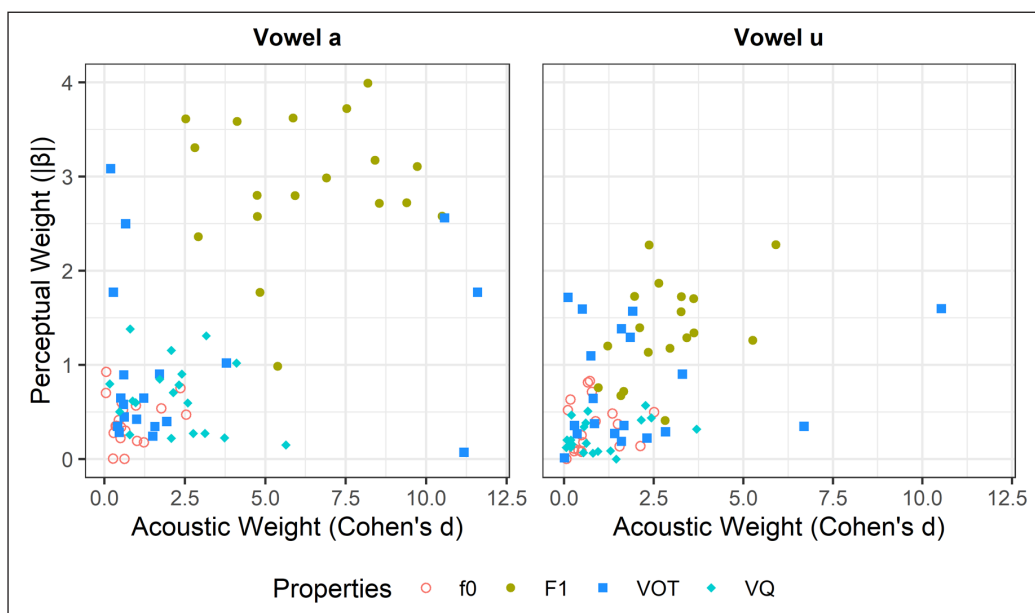
	Estimate	SE	z value	Pr(> z )
(Intercept)	0.209	0.210	0.997	0.319
F1	1.500	0.115	12.989	<0.001
VOT 20 ms	-0.600	0.129	-4.653	<0.001
VOT neg	-2.000	0.297	-6.725	<0.001
f0	0.345	0.048	7.214	<0.001
OQ	-0.062	0.051	-1.210	0.226
F1:VOT 20 ms	-0.138	0.084	-1.634	0.102
F1:VOT neg	-0.561	0.092	-6.115	<0.001

syllables, the factors that have the strongest effect on the results are F1 and the presence of a negative VOT. F1 is positively correlated with high register responses (i.e., syllables with /t/), while a negative VOT prompts more low register responses (i.e., a lower rate of /t/ responses). F1 weighs more than negative VOT for a-stimuli ( $\beta$ 's of 2.88 and  $-2.19$ , respectively in **Table 4**), while the opposite is observed for u-stimuli ( $\beta$ 's of 1.50 versus 2.00 in **Table 5**). Positive VOTs (10 and 20 ms) elicited similar responses in a-stimuli, while a 20 ms long VOT seems to slightly favor low register responses in u-stimuli.

Other factors play a more limited, yet still significant role. F0 is positively correlated with high register responses in both sets of stimuli, but the magnitude of the effect is fairly small. Open quotient is correlated with low-register responses in a-stimuli, but is not significant in u-stimuli. Several weak interactions are also observed. In both sets of stimuli, an increase in F1 favors high-register responses to a more limited extent in stimuli with a negative than a positive VOT (F1:VOT neg). Moreover, in a-stimuli, simultaneous increases in F1 and breathiness (F1:OQ) favor low-register responses more than independent increments in F1 and breathiness, and an increase in breathiness yields more high-register responses when combined with a high f0 (f0:OQ).

### 3.2.2. Relation between production and perception

Finally, we explore the relation between production and perception for the 19 participants who completed both production and perception experiments. Identification weights were computed by fitting logistic regressions for each listener with the independent variables f0, F1, OQ, and VOT on the two sets of stimuli. Interactions were not included due the small number of tokens tested with each individual participant. Our proxy for production weights are Cohen's  $d$  scores computed for the f0, F1, H1\*-H2\*, and VOT of the registers as realized on vowels /a:/, u:/ at the first sampling point after plain stops for each participant. Results are plotted in **Figure 15**. The weight of each relevant acoustic property (Cohen's  $d$ ) is reported on the x-axis, while the weight of each perceptual cue ( $|\beta|$ ) is reported on the y-axis. Each speaker is represented by five different symbols corresponding to the



**Figure 15:** Correlation between perceptual weights ( $|\beta|$ ) and acoustic weights (Cohen's  $d$ ) in individual participants, for vowels /a:/ and /u:/. VQ stands for phonation type and represents the acoustic property H1\*-H2\* on the x-axis and the OQ of the synthesized stimuli on the y-axis. A token of F1 with a  $\beta$  of 22.8 is omitted from the left panel.

five phonetic properties of interest. F1 is the main acoustic property and perceptual cue of register across speakers, and distinguishes registers more efficiently in open vowels (/a:/) than close vowels (/u:/). Although VOT also plays a role in the production and perception of register for most participants, the production and perception weights of VOT do not always correlate at the individual level. Four women and two men who have a low production weight for VOT (Cohen's  $d$  below 2.5) in either /a:/ or /u:/ nevertheless attribute it a large perceptual weight ( $|\beta|$  above 1): In other words, while they recognize VOT to be a correlate of register, they do not use it to distinguish registers in their own production (cf. Coetzee et al., 2018). Also notable is a 63-year-old male who maintains a clear VOT distinction in production, but does not rely on it for identification (the blue square at the bottom right of both panels in **Figure 15**). Phonation type and  $f_0$  play a more limited role in both production and perception.

Globally, the only significant correlation between the Cohen's  $d$ 's and  $\beta$  estimates of any of phonetic properties across participants is found in  $f_0$  in the vowel /u:/ ( $r = .3$ ,  $p = .008$ ; see Script 2 in the Supplementary Material). This means that overall, the weight of participants' perceptual cues cannot be predicted from the weight of their corresponding production properties, which is not surprising as listeners have to accommodate speakers who produce the contrast differently from them. Moreover, no gender differences similar to those found in the weight of production properties (**Figure 11**) are apparent.

### 3.3. Summary of perception results

Pooled results show that low F1 and negative VOT are the primary cues associated with the low register. However, two noteworthy differences between a- and u- stimuli must be mentioned. First, the weight of F1 is greater for a- than u- stimuli. Although this may be due in part to the more restricted range of F1 in our u- stimuli, it mirrors the stronger acoustic weight of F1 in open vowels in the production experiment. Second, there is an association between the exaggerated positive VOT (+20 ms) and the low register in u-, but not in a- stimuli. While this could be further evidence of a weak association between a long positive VOT and low register, why it is only found in u- stimuli remains unclear.

Phonation (OQ) and  $f_0$  play a more limited role in perception, but both show an effect in the expected direction. That  $f_0$  is perceptually relevant despite its weak acoustic distinctiveness and individual variation in production probably indicates that listeners are aware that most speakers have a higher  $f_0$  in the high register, even if it is not a very reliable cue across speakers. It is also possible that the weak perceptual role of phonation type, which is not even significant in u- stimuli, is due to imperfect synthesis.

The  $\beta$  estimates obtained from individual logistic regressions (**Figure 15**) show that F1 is a crucial identification cue for all speakers. VOT is important for a significant minority of listeners, while other acoustic parameters have a much weaker effect on perception, which is in general accordance with the production results. However, there is no evidence that the weight of perceptual cues is structured along gender or age lines, and there are no systematic correlations between individual participants' production and perception weights. In fact, while a significant minority of participants are perceptually sensitive to VOT, they do not systematically use it to distinguish between the registers in production.

## 4. General discussion

Our production results clearly establish that the Proto-Chamic onset voicing contrast has been transphonologized into register in Đờn Dượng Chru (Q1a). All our participants now realize the original contrast by means of register properties, while the voicing contrast is now optional for all but one speaker (Q1b). The register distinction is primarily realized



as a modulation of F1 over the first 100 ms of the vowel: There is a clear F1 rise at the beginning of open vowels in the low register (sometimes strong enough to yield an audible falling onglide), and a more moderate F1 fall at the beginning of close vowels in the high register. These vowel trajectories are in line with what is found in other register languages (Jenner, 1974; Huffman, 1976; Ferlus, 1979; Huffman, 1985; Wayland & Jongman, 2002). Other acoustic properties are associated with the register contrast: Some laxness/breathiness is systematically present at the very beginning of low register vowels across speakers, and weak  $f_0$  and F2 differences are attested in most, but not all, speakers.

Chru could be categorized as a language in which there is ‘phonemic vowel register’ and optional ‘retention of sub-phonemic differentiation in the stops vis-à-vis register’ (Huffman, 1976, p. 587). However, in all but a handful of Chru speakers, this optional sub-phonemic differentiation in stops is realized not as an increased VOT, the typical scenario in Austroasiatic register, but rather as prevoicing. That prevoicing and register coexist in Chru suggests that a weak aspiration of devoiced stops may not always be a step in the development of the register contrast, as generally assumed (Haudricourt, 1965; Huffman, 1976; Wayland & Jongman, 2002). The fact that F1 distinguishes registers better than phonation type or positive VOT for all of our participants, including those with high rates of prevoicing, further challenges the idea that  $f_0$  and vowel quality differences must *necessarily*<sup>5</sup> develop out of breathiness or laxness (Thurgood, 2002; Wayland & Jongman, 2002).

If not mediated through breathy phonation, how could a system like Chru come about? Several authors have proposed that the phonetic properties of register are phonologized consequences of articulatory strategies for circumventing the aerodynamic voicing constraint by increasing the transglottal pressure differential (Gregerson, 1976; Ferlus, 1979; Thurgood, 2002; Brunelle, 2010). Two such strategies that would have a direct impact on F1 are tongue root advancement, which causes a raising and a forward movement of the tongue body, and larynx lowering, which lengthens the back cavity responsible for F1 resonances (Bell-Berti, 1975; Lindau, 1979; Tiede, 1996; Fulop, Kari, & Ladefoged, 1998; Ahn, 2018). The spectral slope and  $f_0$  differences between registers would be additional, parallel, consequences of these two gestures (Ohala, 1972; Bell-Berti, 1975; Lindau, 1979; Fulop et al., 1998; Honda et al., 1999; Hoole & Honda, 2011), rather than being directly responsible for the development of F1 differences. Whether a given register language attributes more weight to spectral slope, F1 or  $f_0$  would then be a consequence of this multidimensionality, as individual listeners may potentially assign each cue a perceptual weight differing from those of the speaker (Beddor, 2009).

A direct route from voicing to register signaled primarily by vowel height is in fact supported by our comparison of the acoustic properties of vowels following prevoiced and devoiced low-register stops (Section 2.2.3). We have seen that there is evidence for the *opposite* of a compensatory relation between prevoicing, and  $f_0$ ,  $H1^*-H2^*$ , and CPP: Low-register stops pattern closer to high-register stops when they are devoiced than prevoiced (Q1c). This suggests that pitch and phonation variation in the register system of Chru remain to some extent automatic effects of prevoicing, a phenomenon also seen in Tibeto-Burman languages such as Dzongkha (Kirby & Hyslop, 2019). On the other hand, F1, the primary registral property for all speakers, does not seem to differ after prevoiced

<sup>5</sup> There is solid evidence that the emergence of contrastive tone *can* be mediated by onset-conditioned breathy phonation, at the very least in Sino-Tibetan (Cao & Maddieson, 1992; Watters, 2002; Mazaudon & Michaud, 2008; Mazaudon, 2012; Shi et al., 2020).

and devoiced low-register stops, indicating that it has been phonologized. What remains unclear, as pointed out in Section 1.1, is the articulatory or auditory mechanism that links voicing and laxness/breathiness, especially in light of the fact that closure voicing does not typically condition this type of phonation on following vowels.

Although Chru seems to have developed a contrastive register system in the obstruent sub-system, there is no evidence that it has been generalized to non-contrastive contexts (sonorants, fricatives, implosives, aspirated stops), contrary to what has been observed in many Austroasiatic languages (Huffman, 1976). The acoustic properties following these onsets have not been forced into specific registers: /s-/, for example, is followed by a high  $f_0$  characteristic of the high register, but a steep spectral slope normally associated with the low register. There is also no indication that register is involved in phonological alternations like register spreading. This could be interpreted as a sign that Chru still has a relatively conservative form of register, but it should be emphasized that the phonologization of register does not entail its generalization to new environments: This can be compared to the evolution of tone splits, which can lead to a larger inventory of contrastive tones on just a subset of syllable types (unchecked syllables in Vietnamese: Haudricourt, 1954; obstruent-initial syllables in Kra-Dai: Pittayaporn, 2009; sonorant-initial syllables in Wu: Gao, 2015).

In terms of perception (Q2), F1 and VOT are the main cues used in register identification, but while F1 is a significant cue across listeners, VOT is weaker and appears more variable across individuals. F1 also weighs more in the identification of /a:/ than /u:/, which reflects its greater role in the register contrast in non-close vowels. Other cues bias responses in expected directions, but have smaller weights. The results of our perception study thus roughly mirror those of the production study. A comparison of the weights of perceptual and production cues (Q3) reveals no direct correlation between the acoustic weight (Cohen's  $d$ ) and perceptual weight ( $\beta$ ) of the participants who underwent both experiments. The only structured pattern that emerges is that some participants seem sensitive to VOT in perception even if they do not themselves consistently produce different VOTs for the two registers.

In light of these results, can we claim that the Chru register contrast is stable? A definitive answer to this question requires a more systematic sociophonetic study, with a larger participant sample and better controls for factors other than age and sex. However, we seem to be dealing with a situation in which all speakers have a contrastive register based on F1, but in which male speakers, especially older ones, maintain optional closure voicing in the low register and idiosyncratic use of other cues. On the other hand, female speakers, especially younger ones, appear to have mostly dropped the closure voicing cue and assign a systematically lesser weight to acoustic cues other than F1, thus leading the change in reinterpreting register as an exclusively vocalic contrast. Younger men seem similar to young women, suggesting that the attrition of prevoicing may be nearing completion. This is similar to what was recently described in Afrikaans, where the loss of voicing was transphonologized into an  $f_0$  contrast in utterance-initial position (Coetzee et al., 2018). Our perception results are compatible with this interpretation: Almost all participants primarily rely on F1 for identification, which should be enough for communicative purposes given that all speakers use F1 in production. However, several participants are making a more significant use of VOT in perception than production, which could be interpreted as evidence that some innovators preserve conservative perceptual cues to accommodate less advanced speakers (Pinget et al., 2016; Howe, 2017; Kuang & Cui, 2018; Pinget et al., 2019).

## Additional Files

The additional files for this article can be found as follows:

- **Appendix A.** Wordlist. DOI: <https://doi.org/10.5334/labphon.278.s1>
- **Appendix B.** Mixed models. DOI: <https://doi.org/10.5334/labphon.278.s2>
- **Supplementary Material 1.** Acoustic\_results zip files contain data for each of the speakers (15 female and 11 male). DOI: <https://doi.org/10.5334/labphon.278.s3>
- **Supplementary Material 2.** Acoustic\_measures\_synthesized\_data contain tabulated acoustic measures of the synthesized stimuli. DOI: <https://doi.org/10.5334/labphon.278.s4>
- **Supplementary Material 3.** Identification\_results contain tabulated results of the identification experiment. DOI: <https://doi.org/10.5334/labphon.278.s5>
- **Supplementary Material 4.** Scripts.zip contains five scripts: two R markdown files containing the code used to analyze and plot tabulated acoustic measures and identification responses (Chru Script 1 Acoustic measures and charts.R, Chru Script 2 Identification results, charts and perception-production.R), two Praat-Klattgrid scripts used to synthesis identification stimuli (Chru\_Script\_3\_mta\_synthesis.praat, Chru\_Script\_4\_tuq\_synthesis.praat), and one R script used to plot tabulated acoustic measures of the identification stimuli (Chru Script 5 Acoustics of synthesized identification stimuli.R). DOI: <https://doi.org/10.5334/labphon.278.s6>

## Acknowledgements

We would like to thank our Chru consultants, especially Touneh Mabio, Gia Đăm Kai, and Touneh Hàn Mai, who helped us with logistics in Đôn Dương, and the People's Committees of the province and district, who graciously granted us field work authorizations. We also want to thank our RAs, Sue-Anne Richer, Belén López, and Sabrina McCullough, and to acknowledge Ryan Gehrman and Suzy Ahn for useful discussions of issues discussed in the paper. This work was made possible by grants from the Social Sciences and Humanities Research Council of Canada [435-2017-0498] and the UK Arts and Humanities Research Council [AH/P014879/1].

## Competing Interests

The authors have no competing interests to declare.

## References

- Abercrombie, D.** (1967). *Elements of general phonetics*: Aldine Publishing Company.
- Abramson, A. S., Luangthongkum, T., & Nye, P. W.** (2004). Voice register in Suai (Kuai): An analysis of perceptual and acoustic data. *Phonetica*, *61*, 147–71. DOI: <https://doi.org/10.1159/000082561>
- Abramson, A. S., Nye, P. W., & Luangthongkum, T.** (2007). Voice register in Khmu': Experiments in production and perception. *Phonetica*, *64*, 80–104. DOI: <https://doi.org/10.1159/000107911>
- Abramson, A. S., Tiede, M. K., & Luangthongkum, T.** (2015). Voice Register in Mon: Acoustics and Electroglottography. *Phonetica*, *72*, 237–56. DOI: <https://doi.org/10.1159/000441728>
- Ahn, S.** (2018). The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics*, *71*, 451–67. DOI: <https://doi.org/10.1016/j.wocn.2018.10.003>
- Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T.-J.** (2018). The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics*, *66*, 120–44. DOI: <https://doi.org/10.1016/j.wocn.2017.09.005>

- Beddor, P. S.** (2009). A coarticulatory path to sound change. *Language*, 85, 785–821. DOI: <https://doi.org/10.1353/lan.0.0165>
- Bell-Berti, F.** (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *The Journal of the Acoustical Society of America*, 57, 456–61. DOI: <https://doi.org/10.1121/1.380468>
- Brunelle, M.** (2005). *Register in Eastern Cham: Phonological, Phonetic and Sociolinguistic approaches*. Doctoral dissertation, Cornell University. [http://aix1.uottawa.ca/~mbrunell/Brunelle\\_Dissertation.pdf](http://aix1.uottawa.ca/~mbrunell/Brunelle_Dissertation.pdf)
- Brunelle, M.** (2009a). Contact-induced change? Register in three Cham dialects. *Journal of Southeast Asian Linguistics*, 2, 1–22.
- Brunelle, M.** (2009b). Diglossia and Monosyllabization in Eastern Cham: A Sociolinguistic Study. J. Stanford & D. Preston (Eds.), *Variation in Indigenous Minority Languages* (pp. 47–75). Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/impact.25.04bru>
- Brunelle, M.** (2010). The role of larynx height in the Javanese tense ~ lax stop contrast. R. Mercado, E. Potsdam & L. Travis (Eds.), *Austronesian Contributions to Linguistic Theory: Selected Proceedings of AFLA* (pp. 7–24). Amsterdam/Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/la.167.03bru>
- Brunelle, M., Hà, K. P., & Grice, M.** (2016). Inconspicuous coarticulation: A complex path to sound change in the tone system of Hanoi Vietnamese *Journal of Phonetics* (pp. 23–39). DOI: <https://doi.org/10.1016/j.wocn.2016.08.001>
- Brunelle, M., & Kirby, J.** (2016). Tone and phonation in Southeast Asian languages. *Language and Linguistics Compass*. DOI: <https://doi.org/10.1111/lnc3.12182>
- Cao, J., & Maddieson, I.** (1992). An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics*, 20, 77–92. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30255-4](https://doi.org/10.1016/S0095-4470(19)30255-4)
- Central population and housing census steering committee.** (2020). *The 2019 Vietnam population and housing census*. Hanoi: Central Population and Housing Census, Steering Committee.
- Clayards, M.** (2018). Differences in cue weights for speech perception are correlated for individuals within and across contrasts. *The Journal of the Acoustical Society of America*, 144, EL172–EL177. DOI: <https://doi.org/10.1121/1.5052025>
- Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A.** (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, 108, 804–809. DOI: <https://doi.org/10.1016/j.cognition.2008.04.004>
- Coetzee, A. W., Beddor, P. S., Shedden, K., Styler, W., & Wissing, D.** (2018). Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis. *Journal of Phonetics*, 66, 185–216. DOI: <https://doi.org/10.1016/j.wocn.2017.09.009>
- Cohen, J.** (1988). *The effect size index: d*. New York: Routledge. DOI: <https://doi.org/10.4324/9780203771587>
- Cohn, A. C.** (1993). Consonant-Vowel Interactions in Madurese: The Feature Lowered Larynx. *Papers from the regional meeting of the Chicago Linguistic Society*, 29, 105–119.
- Cohn, A. C., & Lockwood, K.** (1994). A phonetic description of Madurese and its phonological implications. *Working Papers of the Cornell Phonetics Laboratory*, 9, 67–92. DOI: <https://doi.org/10.1017/S0025100318000257>
- Davidson, L.** (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, 54, 35–50. DOI: <https://doi.org/10.1016/j.wocn.2015.09.003>
- DiCanio, C.** (2009). The Phonetics of Register in Takhian Thong Chong. *Journal of the International Phonetic Association*, 39, 162–88. DOI: <https://doi.org/10.1017/S0025100309003879>

- Dmitrieva, O., Llanos, F., Shultz, A. A., & Francis, A. L.** (2015). Phonological status, not voice onset time, determines the acoustic realization of onset f<sub>0</sub> as a secondary voicing cue in Spanish and English. *Journal of Phonetics*, 49, 77–95. DOI: <https://doi.org/10.1016/j.wocn.2014.12.005>
- Edmondson, J., & Gregerson, K.** (1993). Western Cham as a Register Language. J. Edmondson & K. Gregerson (Eds.), *Tonality in Austronesian Languages* (pp. 61–74). Honolulu: U of Hawaii Press. [www.jstor.org/stable/20006748](http://www.jstor.org/stable/20006748)
- Esposito, A.** (2002). On vowel height and consonantal voicing effects: Data from Italian. *Phonetica*, 59, 197–231. DOI: <https://doi.org/10.1159/000068347>
- Ewan, W., & Krones, R.** (1974). Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics*, 2, 327–35. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31302-6](https://doi.org/10.1016/S0095-4470(19)31302-6)
- Fagan, J. L.** (1988). Javanese Intervocalic Stop Phonemes. *Studies in Austronesian Linguistics*, 76, 173–202. <https://www.ohioswallow.com/book/Studies+in+Austronesian+Linguistics>
- Ferlus, M.** (1979). Formation des Registres et Mutations Consonantiques dans les Langues Mon-Khmer. *Mon Khmer Studies*, VIII, 1–76. <http://sealang.net/archives/mks/pdf/8:1-76.pdf>
- Friberg, T., & Hor, K.** (1977). Register in Western Cham phonology. D.D. Thomas, E.W. Lee & Đ.L. Nguyễn (Eds.), *Papers in Southeast Asian Linguistics No.4* (pp. 17–38). Canberra: Pacific Linguistics. <https://openresearch-repository.anu.edu.au/bitstream/1885/145081/1/PL-A48.pdf>
- Fuller, E.** (1977). Chru phonemes. D.D. Thomas, E.W. Lee & Đ.L. Nguyễn (Eds.), *Papers in South East Asian Linguistics No. 4: Chamic Studies* (pp. 105–24). Canberra: Pacific Linguistics. <https://openresearch-repository.anu.edu.au/bitstream/1885/145081/1/PL-A48.pdf>
- Fulop, S. A., Kari, E., & Ladefoged, P.** (1998). An Acoustic Study of the Tongue Root Contrast in Degema Vowels. *Phonetica*, 55, 80–98. DOI: <https://doi.org/10.1159/000028425>
- Gao, J.** (2015). *Interdependance between Tones, Segments and Phonation types in Shanghai Chinese: Acoustics, articulation, perception and evolution* Paris: Sorbonne Nouvelle – Paris III <http://www.theses.fr/2015USPCA057>
- Garner, W. R., & Felfoldy, G. L.** (1970). Integrality of Stimulus Dimensions in Various Types of Information Processing. *Cognitive Psychology*, 1, 225–41. DOI: [https://doi.org/10.1016/0010-0285\(70\)90016-2](https://doi.org/10.1016/0010-0285(70)90016-2)
- Goudbeek, M., Cutler, A., & Smits, R.** (2008). Supervised and unsupervised learning of multidimensionally varying non-native speech categories. *Speech Communication*, 50, 109–25. DOI: <https://doi.org/10.1016/j.specom.2007.07.003>
- Gregerson, K.** (1976). Tongue-root and Register in Mon-Khmer. P.N. Jenner, L. Thompson & S. Starosta (Eds.), *Austroasiatic Studies* (pp. 323–69). Honolulu: University Press of Hawaii. <http://sealang.net/sala/archives/pdf8/gregerson1976tongue.pdf>
- Hanson, H. M.** (2009). Effects of obstruent consonants on fundamental frequency at vowel onset in English. *The Journal of the Acoustical Society of America*, 125, 425–41. DOI: <https://doi.org/10.1121/1.3021306>
- Harrington, J., Kleber, F., & Reubold, U.** (2008). Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America*, 123, 2825–35. DOI: <https://doi.org/10.1121/1.2897042>
- Haudricourt, A.** (1954). De l'origine des tons en vietnamien. *Journal Asiatique*, 242, 69–82.

- Haudricourt, A.** (1965). Les mutations consonantiques et les occlusives initiales en mon-khmer. *Bulletin de la Société Linguistique de Paris*, 60, 160–72.
- Hawks, J. W., & Miller, J. D.** (1995). A formant bandwidth estimation procedure for vowel synthesis [43.72. Ja]. *The Journal of the Acoustical Society of America*, 97, 1343–44. DOI: <https://doi.org/10.1121/1.412986>
- Hayward, K.** (1995). /p/ vs. /b/ in Javanese: the Role of the Vocal Folds. *Working Papers in Linguistics and Phonetics*, 5, 1–11.
- Henderson, E.** (1952). The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies*, 14, 453–76. DOI: <https://doi.org/10.1017/S0041977X00084251>
- Hillenbrand, J. M., Clark, M. J., & Nearey, T. M.** (2001). Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America*, 109, 748–63. DOI: <https://doi.org/10.1121/1.1337959>
- Holt, L. L., & Lotto, A. J.** (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America*, 119, 3059–71. DOI: <https://doi.org/10.1121/1.2188377>
- Hombert, J.-M., Ohala, J. J., & Ewan, W. G.** (1979). Phonetic Explanation for the Development of Tones. *Language*, 55, 37–58. DOI: <https://doi.org/10.2307/412518>
- Honda, K., Hirai, H., Masaki, S., & Shimada, Y.** (1999). Role of Vertical Larynx Movement and Cervical Lordosis in F0 Control. *Language and Speech*, 42, 401–11. DOI: <https://doi.org/10.1177/00238309990420040301>
- Hoole, P., & Honda, K.** (2011). Automaticity vs. feature-enhancement in the control of segmental F0. *Where do phonological features come from* (pp. 131–71). DOI: <https://doi.org/10.1075/lfab.6.06hoo>
- House, A. S., & Fairbanks, G.** (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America*, 25, 105–13. DOI: <https://doi.org/10.1121/1.1906982>
- Howe, P. J.** (2017). *Tonogenesis in Central dialects of Malagasy: Acoustic and perceptual evidence with implications for synchronic mechanisms of sound change*. Houston: Rice University. <https://scholarship.rice.edu/bitstream/handle/1911/96031/HOWE-DOCUMENT-2017.pdf?sequence=1&isAllowed=y>
- Huffman, F.** (1976). The register problem in fifteen Mon-Khmer languages. *Oceanic Linguistics special publication Austroasiatic Studies, part 1*, 575–89. <https://www.jstor.org/stable/20019172>
- Huffman, F. E.** (1985). Vowel permutations in Austroasiatic languages: Papers presented to Paul K. Benedict for his 71st birthday. G. Thurgood, J. Matisoff & D. Bradley (Eds.), *Linguistics of the Sino-Tibetan Area: The State of the Art* (pp. 141–45). Canberra: Pacific Linguistics Series C 87, Australian National University. DOI: <https://doi.org/10.15144/PL-C87>
- Hyman, L. M.** (1976). Phonologization. A. Juilland (Ed.), *Linguistic studies presented to Joseph H. Greenberg* (pp. 407–18). Saratoga: Anima Libri.
- Iseli, M., & Alwan, A.** (2004). An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04)*, 72, 669–72. DOI: <https://doi.org/10.1109/ICASSP.2004.1326074>
- Jenner, P. N.** (1974). The Development of Registers in Standard Khmer. N.Đ. Liêm (Ed.), *Southeast Asian Linguistic studies* (pp. 47–60). Canberra: Australian National University. <http://www.sealang.net/sala/archives/pdf8/jenner1974development.pdf>

- Kang, Y.** (2014). Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 76–90. DOI: <https://doi.org/10.1016/j.wocn.2014.03.005>
- Kingston, J., Diehl, R., Kirk, C., & Castleman, W.** (2008). On the internal perceptual structure of distinctive features: The [voice] contrast. *Journal of Phonetics*, 36, 28–54. DOI: <https://doi.org/10.1016/j.wocn.2007.02.001>
- Kingston, J., & Diehl, R. L.** (1994). Phonetic Knowledge. *Language*, 70, 419–54. DOI: <https://doi.org/10.1353/lan.1994.0023>
- Kingston, J., & Macmillan, N. A.** (1995). Integrality of nasalization and F1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Garner paradigm. *Journal of the Acoustical Society of America*, 97, 1261–85. DOI: <https://doi.org/10.1121/1.412169>
- Kingston, J., Macmillan, N. A., Dickey, L. W., Thorburn, R., & Bartels, C.** (1997). Integrality in the perception of tongue root position and voice quality in vowels. *Journal of the Acoustical Society of America*, 101, 1696–709. DOI: <https://doi.org/10.1121/1.418179>
- Kirby, J.** (2018). Praatsauce: Praat-based tools for spectral analysis. <https://github.com/kirbyj/praatsauce>
- Kirby, J., & Hyslop, G.** (2019). An acoustic analysis of onset voicing in Dzonkha obstruents. S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 3607–11). [https://assta.org/proceedings/ICPhS2019/papers/ICPhS\\_3656.pdf](https://assta.org/proceedings/ICPhS2019/papers/ICPhS_3656.pdf)
- Kirby, J., & Ladd, D. R.** (2016). Effects of obstruent voicing on vowel F0: Evidence from “true voicing” languages. *J Acoust Soc Am*, 140, 2400. DOI: <https://doi.org/10.1121/1.4962445>
- Kleber, F., Harrington, J., & Reubold, U.** (2012). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech*, 55, 383–405. DOI: <https://doi.org/10.1177/0023830911422194>
- Kluender, K. R., Diehl, R. L., & Wright, B. A.** (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics*, 16, 153–69. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30480-2](https://doi.org/10.1016/S0095-4470(19)30480-2)
- Kuang, J., & Cui, A.** (2018). Relative cue weighting in production and perception of an ongoing sound change in Southern Yi. *Journal of Phonetics* (pp. 194–214). DOI: <https://doi.org/10.1016/j.wocn.2018.09.002>
- Kuznetsova, A., Brockhoff, P., & Christensen, R.** (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82, 1–26. DOI: <https://doi.org/10.18637/jss.v082.i13>
- Laver, J.** (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.
- Lee, E. W.** (1966). *Proto-Chamic phonologic word and vocabulary*. Bloomington: Indiana University microform.
- Lee, T.** (1983). An acoustical study of the register distinction in Mon. *UCLA Working Papers in Phonetics*, 57, 79–96. <https://escholarship.org/uc/item/1kq6011w>
- Lindau, M.** (1979). The feature expanded. *Journal of Phonetics*, 7, 163–76. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31047-2](https://doi.org/10.1016/S0095-4470(19)31047-2)
- Lisker, L.** (1975). Is it VOT or a first-formant transition detector? *The Journal of the Acoustical Society of America*, 57, 1547–51. DOI: <https://doi.org/10.1121/1.380602>
- Lisker.** (1986). “Voicing” in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech*, 29, 3–11. DOI: <https://doi.org/10.1177/002383098602900102>

- Löfqvist, A., Baer, T., McGarr, N. S., & Story, R. S.** (1989). The cricothyroid muscle in voicing control. *The Journal of the Acoustical Society of America*, *85*, 1314–21. DOI: <https://doi.org/10.1121/1.397462>
- Löfqvist, A., & McGowan, R. S.** (1992). Influence of consonantal environment on voice source aerodynamics. *Journal of Phonetics*, *20*, 93–110. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30256-6](https://doi.org/10.1016/S0095-4470(19)30256-6)
- Mathôt, S., Schreij, D., & Theeuwes, J.** (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–24. DOI: <https://doi.org/10.3758/s13428-011-0168-7>
- Matthews, M.** 2015. An acoustic investigation of Javanese stop consonant clusters. Y. Otsuka, C. Stabile & N. Tanaka (Eds.), *Proceedings of AFLA 21* (pp. 201–17). Canberra: Australian National University. <https://users.clas.ufl.edu/potsdam/papers/AFLA21.pdf>
- Mazaudon, M.** (2012). Paths to tone in the Tamang branch of Tibeto-Burman (Nepal). *The dialect laboratory: Dialects as a testing ground for theories of language change* (pp. 139–77). DOI: <https://doi.org/10.1075/slcs.128.07maz>
- Mazaudon, M., & Michaud, A.** (2008). Tonal Contrasts and Initial Consonants: A Case Study of Tamang, a ‘Missing Link’ in Tonogenesis. *Phonetica*, *65*, 231–56. DOI: <https://doi.org/10.1159/000192794>
- Misnadin, & Kirby, J.** (2020). Acoustic correlates of plosive voicing in Madurese. *The Journal of the Acoustical Society of America*, *147*, 2779–90. DOI: <https://doi.org/10.1121/10.0000992>
- Newman, R. S., Clouse, S. A., & Burnham, J. L.** (2001). The perceptual consequences of within-talker variability in fricative production. *The Journal of the Acoustical Society of America*, *109*, 1181–96. DOI: <https://doi.org/10.1121/1.1348009>
- Ní Chasaide, A., & Gobl, C.** (1993). Contextual variation of the vowel voice source as a function of adjacent consonants. *Language and Speech*, *36*, 303–30. DOI: <https://doi.org/10.1177/002383099303600310>
- Ohala, J.** (1981). The Listener as a Source of Sound Change. C. Masek, R. Hendrik & M.F. Miller (Eds.), *Paper from the Parasession on Language and Behavior: Chicago Linguistic Society* (pp. 178–203). Chicago: The University of Chicago.
- Ohala, J.** (2012). The listener as a source of sound change: An update. M.-J. Solé & D. Recasens (Eds.), *The Initiation of Sound Change* (pp. 21–36). Amsterdam/Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/cilt.323>
- Ohala, J. J.** (1972). How is pitch lowered? Paper presented to the The 83rd meeting of the Acoustical Society of America, Buffalo, 1972.
- Ohala, J. J.** (1989). Sound change is drawn from a pool of synchronic variation. L.E. Breivik & E.H. Jahr (Eds.), *Language change: Contributions to the study of its causes* (pp. 173–98. Berlin; New York: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110853063.173>
- Ohde, R. N.** (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America*, *75*, 224–30. DOI: <https://doi.org/10.1121/1.390399>
- Phạm, X. T.** (1955). *Đa-Ng<sup>~</sup> Tiểu T<sup>~</sup>-Điền (Lexique Polyglotte)*. Đà Lạt: Nhà in Langbian.
- Pinget, A.-F., Kager, R., & Van de Velde, H.** (2016). Regional differences in the perception of a consonant change in progress. *Journal of Linguistic geography*, *4*, 65–75. DOI: <https://doi.org/10.1017/jlg.2016.13>
- Pinget, A.-F., Kager, R., & Van de Velde, H.** (2019). Linking Variation in Perception and Production in Sound Change: Evidence from Dutch Obstruent Devoicing. *Language and Speech*. DOI: <https://doi.org/10.1177/0023830919880206>




- Pittayaporn, P.** (2009). *The Phonology of Proto-Tai*. Ithaca, NY: Cornell. <https://core.ac.uk/display/79023573>
- Proctor, M. I., Shadle, C. H., & Iskarous, K.** (2010). Pharyngeal articulation in the production of voiced and voiceless fricatives. *The Journal of the Acoustical Society of America*, 127, 1507–18. DOI: <https://doi.org/10.1121/1.3299199>
- Schertz, J., Cho, T., Lotto, A., & Warner, N.** (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics*, 52, 183–204. DOI: <https://doi.org/10.1016/j.wocn.2015.07.003>
- Schreiber, E., Onishi, K., & Clayards, M.** (2013). Manipulating phonological boundaries using distributional cues. *Proceedings of Meetings on Acoustics*, 19, 060298. DOI: <https://doi.org/10.1121/1.4801082>
- Seyfarth, S., & Garellek, M.** (2018). Plosive voicing acoustics and voice quality in Yerevan Armenian. *Journal of Phonetics*, 71, 425–50. DOI: <https://doi.org/10.1016/j.wocn.2018.09.001>
- Shi, M., Chen, Y., & Mous, M.** (2020). Tonal split and laryngeal contrast of onset consonant in Lili Wu Chinese. *The Journal of the Acoustical Society of America*, 147, 2901–16. DOI: <https://doi.org/10.1121/10.0001000>
- Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K.** (2011). VoiceSauce: A program for voice analysis. *Proceedings of the International Congress of Phonetic Science XVII* (pp. 1846–49). <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Shue/Shue.pdf>
- Shultz, A. A., Francis, A. L., & Llanos, F.** (2012). Differential cue weighting in perception and production of consonant voicing. *The Journal of the Acoustical Society of America*, 132, EL95–EL101. DOI: <https://doi.org/10.1121/1.4736711>
- Sidwell.** (2009). Classifying the Austroasiatic languages: history and state of the art: Lincom Europa.
- Sidwell, P., & Rau, F.** (2015). Austroasiatic comparative-historical reconstruction: An overview. *The Handbook of Austroasiatic Languages*, 1, 221–363. DOI: [https://doi.org/10.1163/9789004283572\\_005](https://doi.org/10.1163/9789004283572_005)
- Solé, M.-J.** (2018). Articulatory adjustments in initial voiced stops in Spanish, French and English. *Journal of Phonetics*, 66, 217–41. DOI: <https://doi.org/10.1016/j.wocn.2017.10.002>
- Stevens, K. N., & House, A. S.** (1956). Studies of formant transitions using a vocal tract analog. *The Journal of the Acoustical Society of America*, 28, 578–85. DOI: <https://doi.org/10.1121/1.1908403>
- Stevens, K. N., & Klatt, D. H.** (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *The Journal of the Acoustical Society of America*, 55, 653–59. DOI: <https://doi.org/10.1121/1.1914578>
- Styler, W.** (2017). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America*, 142, 2469–82. DOI: <https://doi.org/10.1121/1.5008854>
- Svantesson, J.-O., & House, D.** (2006). Tone production, tone perception and Kammu tonogenesis. *Phonology*, 23, 309–33. [www.jstor.org/stable/4420277](http://www.jstor.org/stable/4420277). DOI: <https://doi.org/10.1017/S0952675706000923>
- Tạ, T. T., Brunelle, M., & Nguyễn, T. Q.** (2019). Chrau register and the transphonologization of voicing. S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 2094–98). [https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS\\_2143.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS_2143.pdf)

- Thongkum, T. L.** (1987). Another Look at the Register Distinction in Mon. *UCLA Working Papers in Phonetics*, 67, 29–48. <https://escholarship.org/uc/item/6t1916dq>
- Thongkum, T. L.** (1989). An acoustic study of the register complex in Kui (Suai). *Mon-Khmer Studies*, 15, 1–19. <http://sealang.net/sala/archives/pdf8/theraphan1986acoustic.pdf>
- Thongkum, T. L.** (1991). An Instrumental Study of Chong Register. J.H.C.S. Davidson (Ed.), *Austroasiatic Languages: Essays in honour of H.L. Shorto* (pp. 141–60). London: School of Oriental and African Studies, University of London. <https://digital.soas.ac.uk/AA00000478/00001>
- Thurgood, E.** (2004). Phonation Types in Javanese. *Oceanic Linguistics*, 43, 277–95. <https://www.jstor.org/stable/3623359>. DOI: <https://doi.org/10.1353/ol.2005.0013>
- Thurgood, G.** (1999). *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change*. Honolulu: University of Hawai'i Press.
- Thurgood, G.** (2002). Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica*, 19, 333–63. DOI: <https://doi.org/10.1075/dia.19.2.04thu>
- Tiede, M. K.** (1996). An MRI-based study of pharyngeal volume contrasts in Akan and English. *Journal of Phonetics*, 24, 399–421. DOI: <https://doi.org/10.1006/jpho.1996.0022>
- Watkins, J.** (2002). *The Phonetics of Wa: Experimental Phonetics, Phonology, Orthography and Sociolinguistics*. Canberra: Australian National University. <https://doi.org/10.15144/PL-531>
- Watters, S. A.** (2002). The sounds and tones of five Tibetan languages of the Himalayan region. *Linguistics of the Tibeto-Burman Area*, 25, 1–65. <http://sealang.net/sala/archives/pdf8/watters2002sounds.pdf>
- Wayland, R.** (1997). *Acoustic and Perceptual Investigation of Breathy and Clear Phonation in Chanthaburi Khmer: Implications for the History of Khmer Phonology*. Doctoral dissertation, Cornell University.
- Wayland, R., & Jongman, A.** (2002). Registrogenesis in Khmer: A phonetic account. *Mon-Khmer Studies*, 32, 101–15. <https://lc.mahidol.ac.th/documents/PublicationMultimedia/MonKhmer/Vol32/wayland2002registronesis.pdf>
- Wayland, R., & Jongman, A.** (2003). Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics*, 31, 181–201. DOI: [https://doi.org/10.1016/S0095-4470\(02\)00086-4](https://doi.org/10.1016/S0095-4470(02)00086-4)

**How to cite this article:** Brunelle, M., Tà, T. T., Kirby, J., & Đinh, L. G. 2020 Transphonologization of voicing in Chru: Studies in production and perception. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11(1): 15, pp. 1–33. DOI: <https://doi.org/10.5334/labphon.278>

**Submitted:** 04 June 2020    **Accepted:** 22 August 2020    **Published:** 13 October 2020

**Copyright:** © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 