Title: Post-focus compression in Brahvi and Balochi

Running Title: PFC in Brahvi and Balochi

Authors:

Nasir A. Syed[a], Abdul Waheed Shah[a], Anqi Xu[b], Yi Xu[b]

[a] Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Pakistan

[b] University College London, United Kingdom

Corresponding Author:

Abdul Waheed Shah

Postal Address: Faculty of Languages and Literature

Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Postal Code 90150

Balochistan, Pakistan

00923138296025

waheedshah@luawms.edu.pk

This research was conducted in a computer laboratory in the Department of English Language and Literature, Lasbela University of agriculture, Water and Marine Sciences, (LUAWMS) Uthal, Lasbela, Balochistan, Pakistan.

Abbreviated title: Post-focus compression in Brahvi and Balochi

Abstract:

Previous research has shown that post-focus compression (PFC) — the reduction of pitch range and intensity after a focused word in an utterance, is a robust means of marking focus, but it is present only in some languages. The presence of PFC appears to follow language family lines. The present study is a further exploration of the distribution of PFC by investigating Brahvi, a Dravidian language, and Balochi, an Indo-Iranian language. Balochi is predicted to show PFC given its presence in other Iranian languages. No Dravidian language has been studied for prosodic focus before and it is not related to any languages with PFC. We recorded twenty native speakers from each language producing declarative sentences in different focus conditions. Acoustic analyses showed that, in both languages, post-focus f_0 and other correlates were significantly reduced relative to baseline neutral-focus sentences, but post-focus lowering of f_0 and intensity was greater in magnitude in Balochi than in Brahvi. The Balochi results confirm our prediction, while the Brahvi results offer the first evidence of PFC in a Dravidian language. The finding of PFC in a Dravidian language is relevant to a postulated origin of PFC, which is related to the controversial Nostratic Macrofamily hypothesis.

Keywords: Post-focus compression, PFC, Nostratic macrofamily, Dravidian languages, Indo-Iranian languages

I. INTRODUCTION

A. Post-focus compression and its cross-linguistic distributions

Focus is a communicative function for highlighting a particular component of an utterance for the purpose of directing the attention of the listener (Chen, Wang and Yang, 2014; Chen and Yang, 2015). Since the classical work of Cooper, Eady and colleagues (Cooper, Eady and Mueller, 1985; Eady and Cooper, 1986; Eady et al., 1986), it is known that the prosodic marking of focus in languages like English involves not only phonetic enhancement of focused words themselves, but also compression of post-focus words in pitch range and intensity. Such post-focus compression (PFC) is demonstrated through strictly controlled experiments in which focus is elicited through mini-dialogues (e.g., question-answer paradigm) and detailed comparison of f_0, intensity and duration between different focus conditions (Xu, 2011; Xu and Xu, 2005). With similar methods, PFC is also found in Beijing Mandarin (Chen, Wang and Xu, 2009; Xu, 1999), Korean (Lee and Xu, 2010), Japanese (Lee and Xu, 2012), Hijazi Arabic (Alzaidi and Xu, 2019), Persian (Abolhasanizadeh, Bijankhan and Gussenhoven, 2012; Taheri-Ardali and Xu, 2012), and Turkish (Ipek, 2011). Evidence of PFC is also shown for Greek (Botinis, Fourakis and Gawronska, 1999), French (Dohen and Loevenbruck, 2004), Egyptian Arabic (Hellmuth, 2006), Lebanese Arabic (Chahal, 2003), Hindi (Patil et al., 2008), German (Féry and Kugler, 2008) and Pashto (Rognoni, Bishop and Corris, 2017). It has also been demonstrated that PFC serves as an important cue for the perception of focus (Chen et al., 2009; Rump and Collier, 1996; Xu, Xu and Sun, 2004).

Despite its finding in many languages, however, PFC is not universal, as many languages do not show compression of pitch range and intensity after focus (Xu, 2011). Strikingly, whether PFC is present in a language is independent of other prosodic aspects such as lexical tone, lexical stress, etc. Languages with PFC can be either tonal, like Mandarin, or non-tonal, like English, and either have lexical stress, e.g., English and German, or no lexical stress, e.g., Korean and French (Xu, 2011). likewise, languages that lack PFC can be also either tonal, e.g., Taiwanese (Chen, Wang and Xu, 2009), Cantonese (Wu and Xu, 2010), Northern Sotho, Zulu, Chichewa, Buli, Hausa (Zerbian, Genzel and Kügler, 2010), or non-tonal, e.g., Wolof (Rialland and Robert, 2001), Malay (Azid and Xu, 2020). It is also independent of syntactic features such as word order, as languages with PFC can be SVO (English, Mandarin), SOV (Uyghur, Tibetan, Turkish), or free order (Russian) (Meyer and Mleinek, 2006), and languages without PFC can also be of different word orders (SVO: Taiwan Mandarin, Taiwanese, Hausa, VOS: Yucatec Maya). These findings have ruled out the possibility of PFC as a correlate of language specific features.

Yet another possibility is that PFC either runs in language families or is distributed with a geographical pattern. Language-family-wise, it is present in Indo-European, Indo-Iranian, Uralic and Altaic languages, but absent in Austronesian (Azid and Xu, 2020; Himmelmann, 2018; Maskikit-Essed and Gussenhoven, 2016; Wang et al., 2012), Niger-Congo (Genzel, Renans and Kügler, 2018; Zerbian et al., 2010) and Amerindian languages (Clopper and Tonhauser, 2013; Gordon, 2008; Kügler and Skopeteas, 2007; McDonough, 2002). Interestingly, the distribution pattern described above shows an intriguing parallel to the Farming/Language Dispersal Hypothesis (Diamond and Bellwood, 2003; Renfrew, 1988; Shouse, 2001), which posits that the

distribution of languages of the world today have mostly resulted from the spread of farming from major agricultural homelands, because farming supports continuous expansion of population in geographic scope. Among the homelands, one of the oldest is the Fertile Crescent, which is the birthplace of the wheat- and barley-based farming in the Near East that goes back to 11,000 BP (Diamond and Bellwood, 2003). The path of the spread of the wheat-based farming also happens to resemble the distribution of the Nostratic macrofamily (Bellwood, 2001; Pederson, 1931). The Nostratic hypothesis has been controversial, however, because it goes beyond the time-depth of traditional methods for assessing the affinity of languages through comparison of their phonology, vocabulary, or morphology (Bellwood, 2001; Campbell, 1998; François, 2015; Longobardi, 2009; Nichols, 1996). The controverses notwithstanding, the two hypotheses share some commonalities. Both postulate a Fertile Crescent origin, and both assume a spreading mechanism involving migration, language contact and language shift. There is empirical evidence, however, that PFC is a feature that does not transfer between languages through contact, so that a language can have PFC only through heritage (Chen, 2015; Chen, Xu & Guion-Anderson, 2014; Xu, 2011). The non-transferability means that any language with PFC is likely hereditarily related to other PFC languages. Thus, all PFC languages could be traced back to the Fertile Crescent, the common origin postulated by both the Farming/Language Dispersal and the Nostratic Macrofamily hypotheses.

The present study is an experimental investigation of Brahvi, a Dravidian language spoken in Pakistan, which has never been studied for focus prosody, with the goal to determine whether it shows evidence of PFC. A positive finding would provide prosodic evidence that it may be

hereditarily related to other PFC languages, as postulated by Nostratic Macrofamily hypothesis. Because a pilot study had generated results that were equivocal, we decided to include Balochi, an Indo-Iranian language spoken in the same area, as a control. Being in the same language family as Persian which has been shown to have PFC (Abolhasanizadeh et al., 2012; Taheri-Ardali and Xu, 2012), Balochi is predicted to also show PFC. There is no previous empirical work on the prosody of Balochi also. In fact, Both Brahvi and Balochi are relatively under-documented, so the next section will provide some preliminary information about the two languages.

B. Brahvi and Balochi

Brahvi, also known as Brahui, is one of the major languages of Balochistan. Brahvi people are believed to be the oldest inhabitants of Balochistan (Quintana-Murci et al., 2001; Palanichamy et al., 2015); their main habitat is in Balochistan but many Brahvi speakers also live in other provinces of Pakistan, particularly Sindh, and in other countries like Iran and Afghanistan. The total Population of Brahvis in all countries, according to Ethnolgue is about 4,220,000 people. The Brahvi language of Balochistan is divided into two major dialects, Sarawan (northern) and Jhalawan (southern). To the east of the Brahvi speaking zone, Sindhi and Eastern Balochi languages are spoken; to the west live speakers of Rakhshani Balochi. And, to the north of the Brahvi speaking zone live Pashto speaking tribes. Lasi, a dialect of Sindhi, and Makrani Balochi, are spoken in the south and southeastern parts of the Brahvi speaking areas, respectively.

Bray (1977) compared Brahvi with other Dravidian languages and established its status as a Dravidian language. According to Andronov (1980), Brahvi speakers separated from the mainstream proto-Dravidian speakers in around 4th to 3rd millennium BCE. The Dravidians

entered the Sub-continent (modern day Pakistan) before the arrival of the Aryans. Various hypotheses have been developed about the origin of Dravidian languages, but none of them is beyond debate. Overall, there are largely two main views, indigenous or foreign origin. Supporters of the indigenous origin hypothesis believe that Brahvi originated in the Sub-continent from a proto language which was spoken in ancient India before the arrival of the Aryans (Tripathy, Nirmala and Reddy, 2008). Thus, ancient India (which also included modern day Pakistan) is the homeland and birthplace of Dravidian languages. Supporters of the foreign origin hypothesis, on the other hand, hold that proto-Dravidian originated from outside of ancient India and, later on, speakers of the language migrated to the Subcontinent of India and modern-day Pakistan (Chaubey et al., 2007; Winters, 2007). In the opinion of Krishnamurti (2003), the argument that Dravidian is a language family from outside of the subcontinent is supported by the notion that Brahvi was segregated from other Dravidian languages at the time when the Dravidian tribes entered the subcontinent. Therefore, Brahvi language is an important linchpin in studying the origin and history of Dravidian languages. A study of focus prosody in Brahvi, in particular, may establish whether it has PFC, which can in turn be used as reference when studying the prosody of other Dravidian languages, which are mostly spoken in India.

Balochi is a language of Iranian family, and its history and origin are relatively clear (Birahimani, 2021). It is spoken in Pakistan and Iranian Balochistan by the people called ‘Baloch’. Almost thirty thousand Balochi speakers also live in Turkmenistan (Kokaislova, 2012). Besides, a large number of speakers of Balochi inhabit Sindh and Punjab provinces of Pakistan and some tribes of Baloch are also settled in Afghanistan. A large-scale migration of the Baloch has also occurred to Middle

Eastern countries from Pakistan so that now Balochi speakers make a considerable proportion of the population in the Arab countries. Beside Balochi, Baloch tribes also speak Brahvi, Sindhi and Saraiki languages in Pakistan.

Syntactically, both Brahvi and Balochi are SOV language (Barjasteh, 2010; Elfenbein, 2015), although Balochi is also said to have flexible word order for direct and indirect objects (Jahani and Korn, 2009). In terms of headedness, both are right headed (Jahani and Korn, 2009). Syntactically, Brahvi is an SOV language in terms of word order, and in terms of headedness, it is right headed like English, where adjective precedes noun (Elfenbein 2015; Jahani and Korn, 2009). Note, however, that factors like word order and stress pattern of language have not been shown to be determining factors of PFC according to previous research mentioned earlier. Historically, according to Dashti (2012), around 3000 years ago, some Aryan tribes moved from central Asia to Iran. In 1200 BCE, they had settled in the area around Elborz Mountains called Balashakan or Balashagan which is between Caspian Sea and Lake Van in the present-day Turkey and Azerbaijan. The name of these tribes was Balashchik, according to Dashti (2012). Due to certain circumstances, these tribes later migrated to the extreme south-eastern tip of the Iranian region and modern day Pakistani Balochistan. They have been living together with Brahvi speakers who had arrived in this area thousands of years before. Nowadays, both are called Baloch and have developed a common identity, culture and even blood relations because of inter-marriage.

There has been little experimental research on the prosody of the two languages. In the literature, only some non-experimental accounts of the stress and sentence type prosody in the two languages can be found. Based on this limited literature, neither language has contrastive lexical stress in the

narrow sense of using it to distinguish words (Elfenbein, 1997a, 1997b; Jahani and Korn, 2009; Krishnamurti, 2003). In the case of Balochi spoken in Pakistan, either a fixed stress falls on the last heavy syllable of a word (Jahani and Korn, 2009), or stress falls on a long vowel or diphthong (Elfenbein, 1997a). For Brahvi, stress falls on a long vowel in a word, but on the first syllable in words consisting of only short syllables (Elfenbein, 1997b). With regard to intonation, both languages are said to have a rise/fall contrast between declarative and interrogative sentences. For Balochi, declarative sentences and sentences with question words have a falling pitch on the sentence-final syllable, while interrogative sentences without a question word have a gradual pitch rise towards the end (Korn and Jahani 2009). For Brahvi, statements, commands and question-word questions have a falling intonation, while yes–no questions ending in an interrogative particle have a rise–fall on the particle (Karishnamurti, 2003).

There has been no prior research, however, on focus prosody in either of the two languages, although Asher (1982: 230–4) does mention that in Brahvi the role of stress is 'to express emphasis'. Here *emphasis* is a term often used to refer to focus. This suggests that there might be some audible prosodic cues for marking focus in the language, which could include PFC. But whether this is the case cannot be clear without systematic experimental investigations. The present study is the first such investigation into both Brahvi and Balochi, using a method that has been developed in the examination of focus prosody in many languages, as reviewed in Xu (2011b).

II. METHOD

The current study is a comparative investigation of focus prosody in Brahvi and Balochi, with the goal to determine if either or both of them exhibit post focus compression (PFC). The methodology follows the established method in recent studies (Chen, Wang, and Xu, 2009; Xu, 1999, 2011), which consists of a question-answer paradigm for eliciting focus at different sentence locations, direct comparison of focus and neutral conditions, examination of continuous f_0 contours, and statistical comparisons of multiple acoustic measurements at on-focus, post-focus and pre-focus locations.

A. Stimuli

For each of the languages, three short sentences were composed, all comprising of only sonorant consonants and vowels. Each sentence consists of three disyllabic words. For each sentence, four *wh-questions* were composed to elicit neutral, sentence-initial, sentence-medial and sentence-final focus, respectively. The sentences used as stimuli are listed in Tables 1 and 2.

Table 1. Brahvi Stimuli: Target sentences (right column) and wh-questions for eliciting different focus. The words to be focused are in bold face and underlined.

Sentence A:	
Q1: Ant Mas? (What happened?)	Nana mama narra. (*our uncle fled*)
Q2: Dina mama narra? (*Whose uncle fled?)*	**Nana** mama narra. (***our*** *uncle fled)*
Q3: Nana deir narra? (*Our which relative fled?)*	Nana **mama** narra. (*Our* ***uncle*** *fled)*
Q4: Nana mama ant kary? (*What did our uncle do?)*	Nana mama **narra**. (*Our uncle* ***fled****)*
Sentence B:	
Q1: Ant Mas? (What happened?)	Numa mami milla. (*your aunt was found)*
Q2: Dina mami milla? (*Whose aunt was found?)*	**Numa** mami milla. (***your*** *aunt was found)*
Q3: Nana deir milla? (*Our which relative was found?)*	Numa **mami** milla. (*your* ***aunt*** *was found*)
Q4: Nana mami amar mas? (*What happened to our aunt?)*	Numa mami **milla**. (*she was* ***found****).*
Sentence C:	
Q1: Ant Mas? (What happened?)	Ona amma manna. (*her/his mother agreed*)
Q2: Dina amma manna? (*Whose mother agreed?)*	**Ona** amma manna. (***Her/his*** *mother agreed)*
Q3: Ona dair manna? (*Which relative of his agreed*?)	Ona **amma** manna. (*His* ***mother*** *agreed)*
Q4: Ona amma ant kary? (*What did her/his mother do?)*	A: Ona amma **manna**. (*His mother* ***agreed****)*

Table 2. Stimuli Balochi: Target sentences (right column) and wh-questions for eliciting different focus. The words to be focused are in bold face and underlined.

Sentence A:	
Q1: Noori che kan? (*what does Noori do?*)	Noori Nama Wani. (*Noori reads the name*)
Q2: Kae Nama Wani? (*Who reads the name*)	**Noori** Nama Wani. (***Noori*** *reads the name*)
Q3: Noori Che Wani? (*What does Noori read?*)	Noori **Nama** Wani. (*Noori* ***reads*** *the name*)
Q4: Noori Nama Che Kan? (*What does Noori do with the name?*)	Noori Nama **Wani**. (*Noori reads the* ***name***)
Sentence B:	
Q1: Yaru che kan? (*what does Yaru do?*)	Yaru Mala Mari. (*Yaru takes care of the herd*)
Q2: Kae Mala Mari? (*Who takes care of the herd?*)	**Yaru** Mala Mari. (*Yaru takes care of the herd*)
Q3: Yaru Che Mari? (*Yaru takes care of what?*)	Yaru **Mala** Mari. (*Yaru takes care of the herd*)
Q4: Yaru Mala Che Kan? (*What does Yaru do to the herd?*)	Yaru Mala **Mari**. (*Yaru takes care of herd*)
Sentence C:	
Q1: Tai Lala Chon Bi? (*What would happen to your brother?*)	Mani Lala Milli. (*My brother will be found*)
Q2: Kai Lala Milli? (*Whose brother will be found?*)	**Mani** Lala Milli.(*My brother will be found*)
Q3: Tai Che Milli? (*Who of your relatives will be found?*)	Mani **Lala**Milli.(*My brother will be found*)
Q4: Tai Lala Milli Na Milli? (*Will your brother be found or not?*)	Mani Lala **Milli**.(*He will be found*)

B. Participants

For both Balochi and Brahvi, twenty male native speakers, aged 18-31 (mean = 23.75, standard deviation = 3.26) and 20-33 (mean = 25.25, standard deviation = 3.09), respectively, participated as subjects. The Brahvi speakers were from Khuzdar district speaking Jhalawani (southern) dialect of Brahvi. The Balochi participants were speakers of the Kechi dialect, which is part of the Southern Balochi dialect (Jahani and Korn, 2009). The participants of both groups were students of undergraduate level in Pakistan. None of them reported any speech or hearing disorders. And they were paid for their participation.

C. Procedure

Data collection was done in a computer laboratory in the Department of English Language and Literature, LUAWMS. The stimulus sentences were presented on a computer screen in Roman script by a Javascript program. The program repeated all the question-answer pairs three times in separate blocks, each with a different randomized order. The randomization order was also different for different speakers. Thus, each speaker produced 36 utterances (3 sentences * 4 focus conditions* 3 repetitions). Before the experiment, the speakers were informed in their mother tongue about the procedure, without information about the real purpose of the study. They were also shown the texts of the question-answer pairs before the start of the recording to get them familiarized with the stimuli.

Three persons conducted the recording experiment at a time. One female native speaker of Brahvi and a male native speaker of Balochi served as the interrogators for Brahvi and Balochi

respectively. Both interrogators were visiting faculty members at the Department of English Language and Literature LUAWMS (Uthal, Pakistan). The second author, who is a native speaker of Brahvi and fully competent L2 speaker of Balochi, served as the observer for both languages. For each trial, the interrogator asked the question displayed on the screen, and the participant replied by reading aloud the answer from the screen. The observer only sat and monitored to make sure everything was going on well. A Sony digital audio recording device (Sony ICD-PX 440) recorded the utterance in .wav format at a sampling rate of 44.1 kHz.

D. Data analysis

The acoustic analysis was performed with the help of ProsodyPro (Xu, 2013), a Praat (Boersma, 2001) script for large-scale systematic analysis of continuous prosodic events. With ProsodyPro we took various measurements of fundamental frequency (f_0), intensity and syllable duration. Syllable segmentation was manually performed within ProsodyPro, and vocal pulse markings generated by Praat were visually and auditorily rectified to guarantee the accuracy of f_0 tracking. Measurements were taken on the basis of syllables and all measurements of intensity, f_0 and syllable duration were averaged for further analysis. All the measurements needed were automatically saved as text files by ProsodyPro. For the analysis, following the practice of previous research on focus prosody (Alzaidi and Xu, 2019; Chen et al., 2009; Xu, 1999), graphic comparisons were first made to examine if different focus conditions resulted in visible changes in f_0 contours. Based on the observation of the graphical comparisons, statistical analyses were then performed to compare measurements from pre-focus, on-focus and post-focus words with those of the same words in the neutral focus condition, with the goal to determine the similarities

and differences of the two languages in their prosodic realization of focus. Although in the overall experimental design, Balochi is treated as a control language, due to its fairly prototypical PFC patterns, as will be seen, it will be discussed before Brahvi for each set of analysis.

1. Graphical analysis

Figure 1 displays mean f_0 contours of Balochi and Brahvi sentences in four focus conditions. For Balochi, as shown in the left column, when focus is sentence-initial, the focused words have much higher f_0 peaks than the neutral focus words in the same positions. The tendency is more pronounced in sentence C than in sentences A and B. At the same time, post-focal f_0 of initial focus is much lower than that in the neutral focus condition across all sentences. When focus is sentence-medial, the on-focus increase and post-focus lowering of f_0 are less clear. Compared with neutral focus, post-focus reduction of f_0 in medial focus can be seen in sentence C but not in the other sentences. f_0 of words in the pre-focus region does not deviate from that of neutral focus. Final focus shows an on-focus raising effect in f_0 in sentences A and B compared with neutral focus. Interestingly, in contrast with medial focus, there is also a subtle pre-focus lowering in the f_0 in final focus compared with neutral focus.

The focus marking of f_0 in Brahvi, as shown in the right column of Figure 1, is overall weaker than in Balochi. For initial focus, on-focus f_0 is slightly higher than that in the same words in neutral focus, especially in sentence A, while post-focus f_0 is lower than that of neutral focus in all sentences. In medial focus, similar to Balochi, the on-focus increase of f_0 is rather limited, but the post-focus lowering of f_0 is more evident in all sentences. The f_0 of pre-focus words is higher than

that in neutral focus in sentences A and B. For final focus, the on-focus words have higher f_0 than the same words in neutral focus in sentences A and C, but the difference is smaller in sentence B.

Overall, therefore, there are visible on-focus raising and post-focus f_0 lowering in both Brahvi and Balochi, with clearer effects in Balochi. The f_0 of pre-focus words in final focus is somewhat lower than in neutral focus in Balochi but the reversed pattern is observed in Brahvi. In the following section, results of statistical analyses will be presented, which will show which of the visible differences are significant.

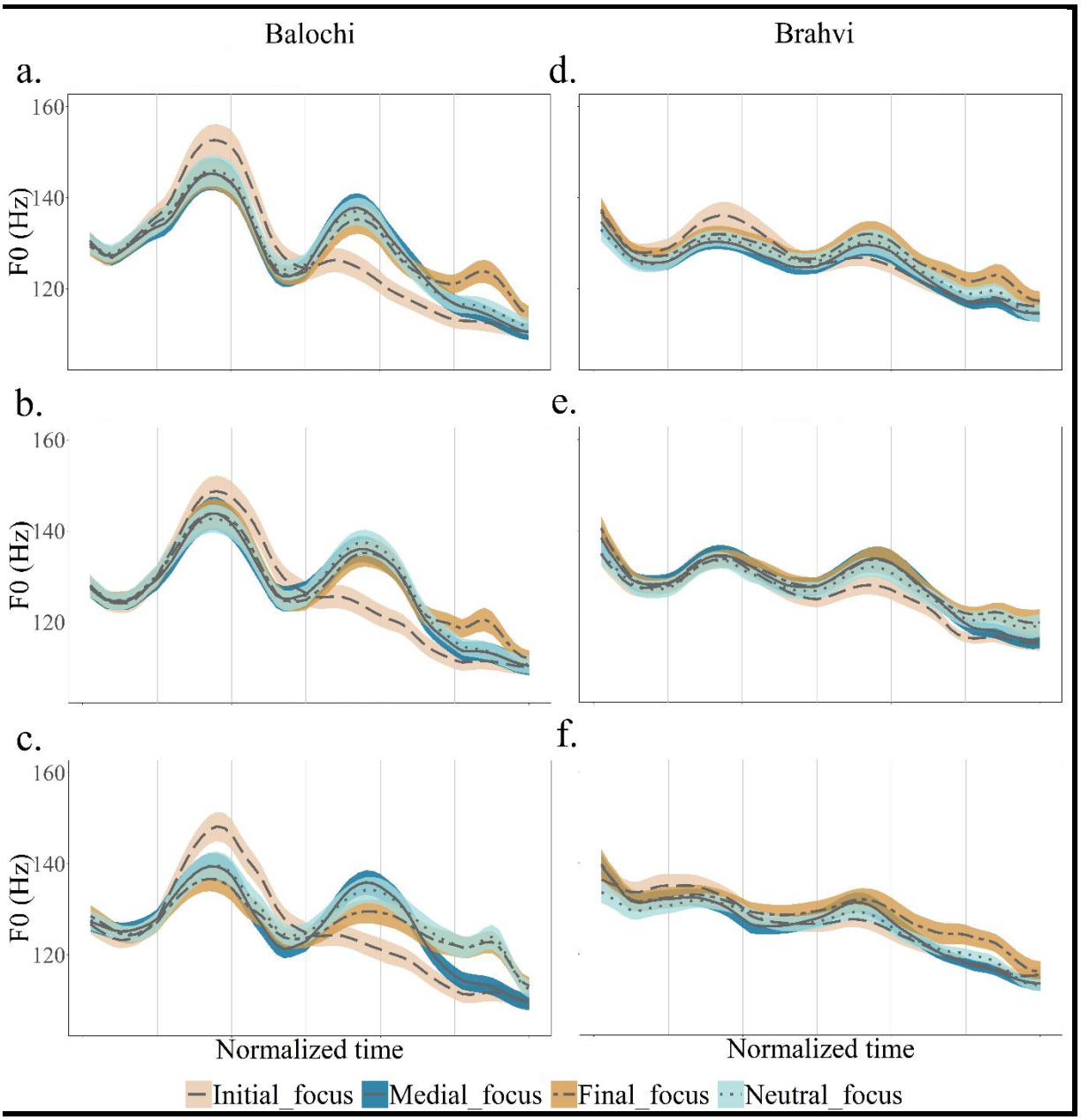

Figure 1: Time-normalized mean F_0 contours of three Balochi and Brahvi sentences in four focus conditions. Each contour is an average of three repetitions by 20 speakers in the respective language. The transparent ribbons around the mean F_0 contours represent standard errors. Vertical

lines mark syllable boundaries. a & d: Balochi and Brahvi Sentence A; b & e: Sentence B; c & f: Sentence C.

2. Statistical analysis

The data analysis was performed with linear mixed-effects models by *lme4* package (Bates et al., 2015) implemented in R (R Core Team, 2019). For each model, we included random intercepts for participant (20 participants) and for sentence (Sentence A, B and C) and by-participant random slopes for 'focus condition' (on/post/pre-focus and neutral focus), by-sentence random slopes for 'focus condition' and by-participant-sentence-interaction random slopes for 'focus condition'. *P*-values were obtained by likelihood ratio tests to check whether the inclusion of the fixed effect contributed significantly to a better model. The analysis was done in two steps. First, we tested the effect of 'focus condition' on syllable duration, intensity, maximum f_0 and mean f_0 for Balochi and Brahvi, respectively. The data were divided into small groups according to focus location (initial, medial and final). In each focus location, the dependent variables were the averaged acoustic measurements of all syllables in the pre-focus, on-focus and post-focus locations, respectively.[1] Neutral focus was treated as the baseline in all comparisons, i.e., the predictor variable 'focus condition' always consists of two levels: neutral vs. on-focus, neutral vs. post-focus and neutral vs. pre-focus. Then the effect of language on the difference between neutral and on-focus or post-

[1] As stated in the Introduction, there is no contrastive lexical stress in either of the two languages, and so we did not perform separate averaging according to stress.

focus condition in initial focus condition was analysed, based on the finding that the effect of 'focus condition' was the most significant in initial focus. The dependent variable is the difference in the acoustic measurements and the predictor variable is 'language' consisting of two levels: Balochi vs. Brahvi.

a. Balochi

Tables 3-5 show summaries of descriptive statistics and the results of linear mixed models of Balochi data for on-focus, post-focus and pre-focus locations, respectively. Table 3 shows the results of on-focus effects. For initial focus, there are significant on-focus effects on duration, mean f_0 and Maximum f_0, and a marginally significant effect on intensity. All the acoustic measurements show an increase relative to neutral focus. However, there are no significant on-focus effects for either medial or final focus.

Table 3: Mean, standard deviation and linear mixed model results of on-focus effects on syllable duration, intensity, maximum F_0 and mean F_0 in Balochi. The arrows indicate where the effects are significant and the direction of the difference relative to the neutral focus condition.

On focus					
		Duration	Intensity	Mean F_0	Maximum F_0
Initial focus	Initial	M = 171; SD = 31	M = 65; SD = 5	M = 137; SD = 18	M = 145; SD = 21
	Neutral	M = 159; SD = 27	M = 64; SD = 5	M = 134; SD = 18	M = 140; SD = 20
	Difference	$\chi2$ = 5.209, df = 1, p = 0.022 ↑	$\chi2$ = 2.894, df = 1, p = 0.089	$\chi2$ = 6.575, df = 1, p = 0.010 ↑	$\chi2$ = 9.9319, df = 1, p = 0.002 ↑
Medial focus	Medial	M = 189; SD = 18	M = 67; SD = 5	M = 131; SD = 18	M = 141; SD = 23
	Neutral	M = 183; SD = 16	M = 67; SD = 5	M = 132; SD = 18	M = 140; SD = 21
	Difference	$\chi2$ = 2.568, df = 1, p = 0.109	$\chi2$ = 0.200, df = 1, p = 0.655	$\chi2$ = 0.224, df = 1, p = 0.636	$\chi2$ = 0.318, df = 1, p = 0.573
Final focus	Final	M = 196; SD = 31	M = 61; SD = 5	M = 122; SD = 15	M = 130; SD = 18
	Neutral	M = 186; SD = 26	M = 61; SD = 5	M = 120; SD = 14	M = 128; SD = 17
	Difference	$\chi2$ = 1.488, df = 1, p = 0.223	$\chi2$ = 0, df = 1, p = 1	$\chi2$ = 1.313, df = 1, p = 0.252	$\chi2$ = 0.980, df = 1, p = 0.322

Table 4 shows statistic results of post-focus effects. For initial focus, there are significant effects on all acoustic measurements of post-focus words. All of them have lower values relative to the neutral focus condition. For medial focus, there is only significant post-focus effect on intensity, which is lower than in the neutral focus condition.

Table 4: Mean, standard deviation and linear mixed model results of post-focus effects on syllable duration, intensity, maximum F_0 and mean F_0 in Balochi. The arrows indicate where the effects are significant and the direction of the difference relative to the neutral focus condition.

Post focus					
		Duration	Intensity	Mean F_0	Maximum F_0
Initial focus	Initial	M = 170; SD = 16	M = 62; SD = 5	M = 121; SD = 16	M = 127; SD = 17
	Neutral	M = 184; SD = 20	M = 64; SD = 5	M = 126; SD = 16	M = 134; SD = 19
	Difference	$\chi2$ = 6.378, df = 1, p = 0.012 ↓	$\chi2$ = 11.303, df = 1, p < 0.001 ↓	$\chi2$ = 8.224, df = 1, p = 0.004 ↓	$\chi2$ = 8.976, df = 1, p = 0.003 ↓
Medial focus	Medial	M = 176; SD = 23	M = 60; SD = 5	M = 118; SD = 15	M = 125; SD = 17
	Neutral	M = 186; SD = 26	M = 61; SD = 5	M = 120; SD = 14	M = 128; SD = 17
	Difference	$\chi2$ = 2.255, df = 1, p = 0.133	$\chi2$ = 4.327, df = 1, p = 0.038 ↓	$\chi2$ = 1.518, df = 1, p = 0.218	$\chi2$ = 1.633, df = 1, p = 0.201

Table 5 shows statistic results of pre-focus effects. The only significant effect is on intensity of pre-focus words for final focus: It is lower in pre-focus words than in neutral focus words. Moreover, the main effect is marginally significant on duration, mean f_0 and maximum f_0. The duration of final focus is shorter than neutral focus in pre-focus region. A similar tendency is found in mean f_0 and maximum f_0, with final focus showing lower f_0 than neutral focus in pre-focus region.

Table 5: Mean, standard deviation and linear mixed model results of pre-focus effects on syllable duration, intensity, maximum F_0 and mean F_0 in Balochi. The arrow indicates where the effect is significant and the direction of the difference relative to the neutral focus condition.

Pre-focus					
		Duration	Intensity	Mean F_0	Maximum F_0
Medial	Medial	M = 163; SD = 33	M = 64; SD = 6	M = 134; SD = 19	M = 140; SD = 21
	Neutral	M = 159; SD = 27	M = 64; SD = 5	M = 134; SD = 18	M = 140; SD = 20
	Difference	$\chi2$ = 1.2373, df = 1, p = 0.266	$\chi2$ = 1.485, df = 1, p = 0.223	$\chi2$ = 0.079, df = 1, p = 0.778	$\chi2$ = 0.069, df = 1, p = 0.792
Final	Final	M = 166; SD = 22	M = 65; SD = 6	M = 132; SD = 18	M = 139; SD = 21
	Neutral	M = 171; SD = 20	M = 66; SD = 5	M = 133; SD = 18	M = 140; SD = 21
	Difference	$\chi2$ = 3.476, df = 1, p = 0.062	$\chi2$ = 5.123, df = 1, p = 0.024 ↓	$\chi2$ = 3.148, df = 1, p = 0.076	$\chi2$ = 2.842, df = 1, p = 0.092

b. Brahvi

Tables 6-8 show summaries of descriptive statistics and results of linear mixed models of Brahvi data for on-focus, post-focus and pre-focus effects, respectively. Table 6 shows that, similar to Balochi, for initial focus, there are significant on-focus effects on duration, mean f_0 and maximum f_0, but no significant effect on intensity. The duration of initial focus is longer and mean f_0 and maximum f_0 are higher than in neutral focus. There are no on-focus effects on any of the acoustic measurements in the medial focus condition. For final focus, there is a significant effect on intensity, with higher intensity in final focus than in neutral focus. The on-focus effects for final focus on duration and mean f_0 are only marginal. Both measurements have slightly greater values in final focus than in neutral focus.

Table 6: Mean, standard deviation and linear mixed model results of on-focus effects on syllable duration, intensity, maximum F_0 and mean F_0 in Brahvi. The arrows indicate where the effects are significant and the direction of the difference relative to the neutral focus condition.

On focus					
		Duration	Intensity	Mean F_0	Maximum F_0
Initial focus	Initial	M = 163; SD = 22	M = 69; SD = 7	M = 133; SD = 17	M = 139; SD = 19
	Neutral	M = 157; SD = 23	M = 68; SD = 7	M = 131; SD = 14	M = 137; SD = 16
	Difference	$\chi2 = 4.384$, df = 1, $p = 0.036$ ↑	$\chi2 = 2.060$, df = 1, $p = 0.151$	$\chi2 = 3.895$, df = 1, $p = 0.048$ ↑	$\chi2 = 4.851$, df = 1, $p = 0.028$ ↑
Medial focus	Medial	M = 166; SD = 21	M = 68; SD = 7	M = 129; SD = 16	M = 133; SD = 18
	Neutral	M = 166; SD = 21	M = 69; SD = 7	M = 130; SD = 16	M = 133; SD = 16
	Difference	$\chi2 = 0.079$, df = 1, $p = 0.778$	$\chi2 = 2.393$, df = 1, $p = 0.122$	$\chi2 = 0.813$, df = 1, $p = 0.367$	$\chi2 = 0.069$, df = 1, $p = 0.793$
Final focus	Final	M = 193; SD = 22	M = 66; SD = 8	M = 123; SD = 16	M = 129; SD = 20
	Neutral	M = 188; SD = 24	M = 65; SD = 7	M = 121; SD = 12	M = 127; SD = 14
	Difference	$\chi2 = 2.932$, df = 1, $p = 0.087$	$\chi2 = 3.937$, df = 1, $p = 0.047$ ↑	$\chi2 = 3.243$, df = 1, $p = 0.072$	$\chi2 = 2.405$, df = 1, $p = 0.121$

Table 7 shows statistic results of post-focus effects. For initial focus, the post-focus effects are significant on intensity, mean f_0 and maximum f_0. Compared with neutral focus, these measurements in post-focus words are all lowered in value. Duration is also shorter than neutral focus in the post-focus region but it is not significant. For medial focus, there are significant post-focus effects on duration, intensity and mean f_0, which all have smaller values than in the neutral focus condition. But the post-focus effect on maximum f_0 is only marginally significant.

Table 7: Mean, standard deviation and linear mixed model results of post-focus effects on syllable duration, intensity, maximum F_0 and mean F_0 in Brahvi. The arrows indicate where the effects are significant and the direction of the difference relative to the neutral focus condition.

Post focus					
		Duration	Intensity	Mean F_0	Maximum F_0
Initial	Initial	M = 171; SD = 21	M = 66; SD = 7	M = 123; SD = 12	M = 127; SD = 14
	Neutral	M = 177; SD = 22	M = 67; SD = 7	M = 125; SD = 14	M = 130; SD = 15
	Difference	$\chi2 = 2.140$, df = 1, $p = 0.144$	$\chi2 = 9.638$, df = 1, $p = 0.002$ ↓	$\chi2 = 7.332$, df = 1, $p = 0.007$ ↓	$\chi2 = 6.837$, df = 1, $p = 0.009$ ↓
Medial	Medial	M = 183; SD = 21	M = 65; SD = 8	M = 120; SD = 12	M = 125; SD = 14
	Neutral	M = 188; SD = 24	M = 65; SD = 7	M = 121; SD = 12	M = 127; SD = 14
	Difference	$\chi2 = 4.264$, df = 1, $p = 0.039$ ↓	$\chi2 = 6.566$, df = 1, $p = 0.010$ ↓	$\chi2 = 4.336$, df = 1, $p = 0.037$ ↓	$\chi2 = 3.157$, df = 1, $p = 0.076$

Table 8 shows statistic results of pre-focus effects. The only significant pre-focus effect is on intensity in final focus, which is lower than in the neutral focus condition. The pre-focus effect on maximum f_0 is only marginal in final focus, because f_0 is actually slightly higher than in the neutral focus condition. There are no significant pre-focus effects in medial focus.

Table 8: Mean, standard deviation and linear mixed model results of mean duration, intensity, maximum F_0 and mean F_0 of pre-focus words in Brahvi. The arrow indicates where the effect is significant and the direction of the difference relative to the neutral focus condition.

Pre-focus					
		Duration	Intensity	Mean F_0	Maximum F_0
Medial	Medial	M = 157; SD = 20	M = 68; SD = 8	M = 131; SD = 16	M = 138; SD = 18
	Neutral	M = 157; SD = 23	M = 68; SD = 7	M = 131; SD = 14	M = 137; SD = 16
	Difference	$\chi2 = 0.111$, df = 1, $p = 0.740$	$\chi2 = 2.651$, df = 1, $p = 0.104$	$\chi2 = 0.03$, df = 1, $p = 0.862$	$\chi2 = 2.243$, df = 1, $p = 0.134$
Final	Final	M = 159; SD = 20	M = 68.1; SD = 7	M = 131; SD = 16	M = 137; SD = 18
	Neutral	M = 162; SD = 21	M = 68.4; SD = 7	M = 130; SD = 15	M = 135; SD = 16
	Difference	$\chi2 = 1.977$, df = 1, $p = 0.160$	$\chi2 = 5.092$, df = 1, $p = 0.024$ ↓	$\chi2 = 0$, df = 1, $p = 1$	$\chi2 = 3.489$, df = 1, $p = 0.062$

c. Comparison of Brahvi and Balochi

Although the two languages show similar patterns in their focus prosody, the magnitude of the change in the acoustic measurements appeared different in our initial inspection. For further analysis, we calculated the differences in duration, intensity, maximum f_0 and mean f_0 between on-focus and neutral focus words, and between post-focus and neutral focus words for both languages in the initial focus condition where the effect of focus is the most significant. Repetitions were collapsed across sentences for each speaker before calculating the differences. The distribution of acoustic measurements of the difference in on-focus region and post-focus region is displayed in Figures 2 and 3, respectively.

Figure 2 shows that there is higher variability in the difference of average tokens produced by Balochi speakers than by Brahvi speakers in all acoustic dimensions. More values were above zero in Balochi than in Brahvi, which suggests a stronger on-focus raising effect. Notably, there are a few tokens by Brahvi speakers that have outstanding on-focus raising effects on mean f_0 and maximum f_0 (Figure 2b & 2d). A similar trend is shown in Figure 3. In the post-focus region, Balochi speakers vary greatly in their marking of focus compared to the Brahvi speakers. Balochi speakers produced more tokens with negative difference values than Brahvi speakers. Again, a few sentences by Brahvi speakers do show a strong post-focus f_0 decrease (see Figure 3b & 3d). But Balochi speakers show stronger post-focus compression in all the acoustic measurements than Brahvi speakers.

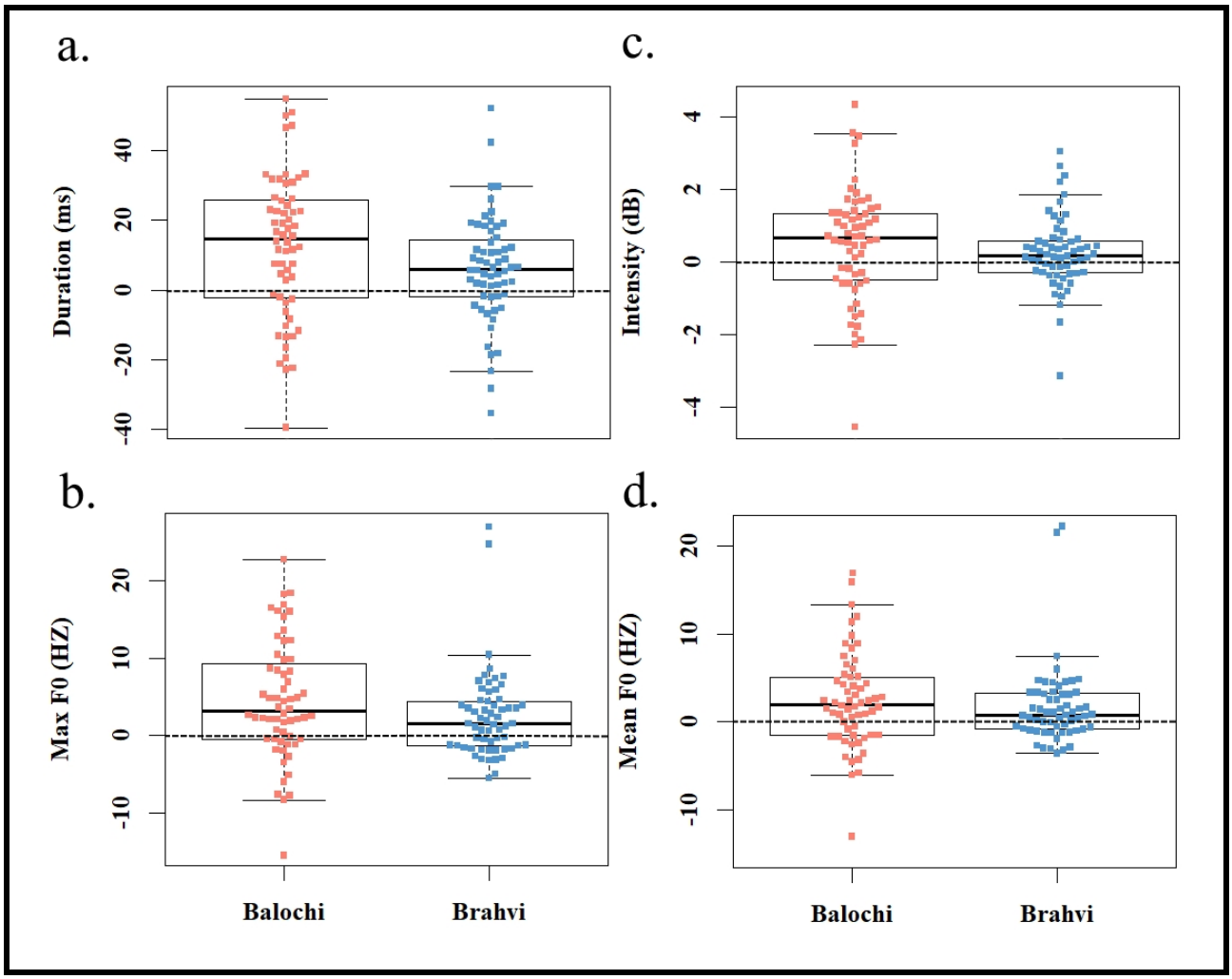

Figure 2. Difference in duration, intensity, maximum F_0 and mean F_0 between initial focus and neutral focus in the on-focus region in Balochi and Brahvi. Each point represents the average of each sentence produced by one participant.

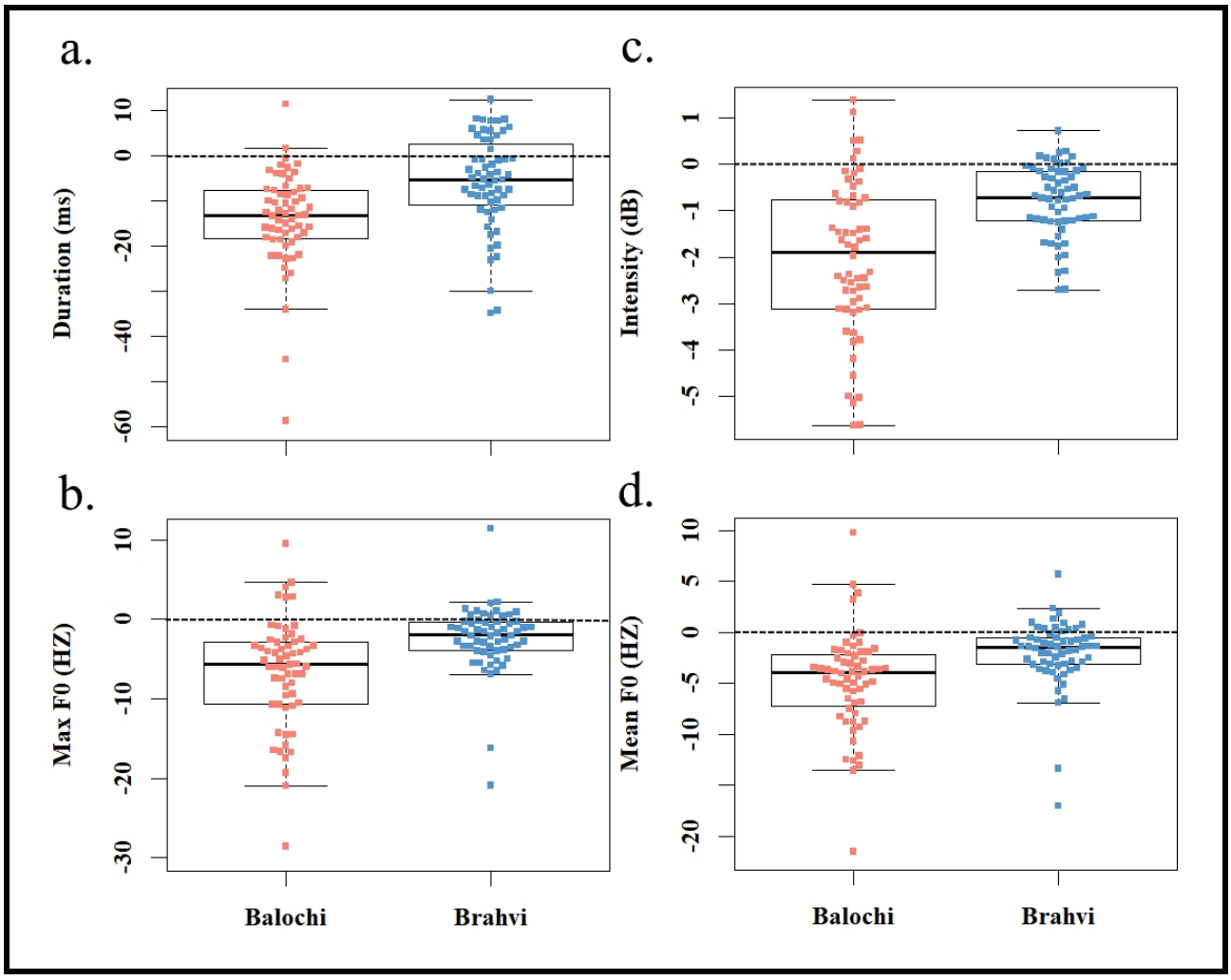

Figure 3. Difference in duration, intensity, maximum F_0 and mean F_0 between initial focus and neutral focus in the post-focus region in Balochi and Brahvi. Each point represents the average of each sentence produced by one participant.

Linear mixed models were constructed to test whether there are significant language effects on focus marking. The results as well as the mean and standard deviation are shown in Table 9. The two languages do not differ in the magnitude of on-focus raising, but significant differences are seen in post-focus effects. As shown in Table 9, in post-focus region, the main effect of language

is significant on the difference in intensity and maximum f_0 and marginally significant on duration and mean f_0. This indicates that post-focus compression is more pronounced in Balochi than Brahvi.

Table 9. Mean, standard deviation and results of linear mixed models of the difference between Balochi and Brahvi, with the difference between initial focus and neutral focus in syllable duration, intensity, maximum F_0 and mean F_0 as dependent variables. The asterisk represents significant main effect.

		Duration	Intensity	Maximum F_0	Mean F_0
On focus	Balochi	12 (21)	4 (8)	2 (5)	0 (2)
	Brahvi	6 (15)	2 (6)	2 (5)	0 (1)
		$\chi2 = 2.378$, df = 1, $p = 0.123$	$\chi2 = 0.686$, df = 1, $p = 0.407$	$\chi2 = 1.176$, df = 1, $p = 0.278$	$\chi2 = 0.318$, df = 1, $p = 0.573$
Post focus	Balochi	-14 (11)	-7 (7)	-5 (5)	-2 (2)
	Brahvi	-6 (10)	-2 (4)	-2 (3)	-1 (1)
		$\chi2 = 2.814$, df = 1, $p = 0.093$	$\chi2 = 6.543$, df = 1, $p = 0.011$ *	$\chi2 = 4.578$, df = 1, $p = 0.032$ *	$\chi2 = 3.667$, df = 1, $p = 0.056$

III. DISCUSSION

Overall, the results of both graphic comparisons and statistical analyses have provided evidence that focus is prosodically marked in both Brahvi and Balochi, and that for both languages, there is more consistent post-focus compression than on-focus enhancement of prosodic properties. For Balochi, significant on-focus enhancement occurs only in initial focus, and only in duration, mean f_0 and maximum f_0, but not in intensity (Table 3). Significant post-focus compression for initial focus occurs in all the four acoustic measurements, although for medial focus only in intensity (Table 4). For Brahvi, on-focus enhancement for initial focus is significant in terms of duration,

mean f_0 and maximum f_0, just like in Balochi, but there is also on-focus enhancement of intensity in final focus (Table 6). There is significant post-focus compression in terms of intensity, mean f_0, and maximum f_0 in initial focus, and also significant post-focus compression in terms of duration, intensity and mean f_0 in medial focus (Table 7). For both languages, there are pre-focus effects only for final focus and only in terms of intensity lowering (Tables 5 and 8).

As shown in Figures 1-3 and Table 9, the magnitude of PFC is smaller in Brahvi than in Balochi. A potential source of the difference is that the neutral focus condition in Balochi is not fully neutral, because the wh-question used in the neutral focus condition asks what the person did, rather than simply asking what happened as in the Brahvi experiment. This may have inadvertently introduced a broad late-focus in that condition, making it somewhat similar to final focus. This may explain why there is no on-focus effect for any of the measurements for final focus in Balochi (Table 3), contrary to the significant on-focus effect on intensity and marginally significant effects on duration and mean f_0 for final focus in Brahvi (Table 6). A close look at Figure 1, however, shows that in Balochi, post-focus f_0 contours are largely flattened in initial focus, whereas in Brahvi a prominent f_0 peak remains in the post-focus region in initial focus. As a result, f_0 contours in the post-focal region are well separated in Balochi for final focus (Figure 1a-c), while in Brahvi the standard-error ribbons are partially overlapped (Figure 1d-f). This, together with the significant differences in intensity and maximum f_0 shown in Table 9, suggests that PFC is weaker in Brahvi than in Balochi.

The finding of PFC in Balochi is not surprising, given similar findings in Persian, another Indo-Iranian language (Abolhasanizadeh et al., 2012; Taheri-Ardali and Xu, 2012). But it does add to

the accumulating evidence that the distribution of PFC falls largely along language family lines (Chen et al., 2009; Xu, 2011). The finding of PFC in Brahvi is more significant. First of all, to the best of our knowledge, this is the first clear evidence of PFC in any Dravidian language. It therefore raises the question whether PFC is also present in other Dravidian languages, which are all spoken thousands of miles away in the southern parts of India from the western part of Pakistan where Brahvi is spoken. Secondly, the presence of PFC in Brahvi may provide new clues for the dispute between the indigenous and foreign origin accounts of Dravidians in the Indian subcontinent. One of the most extreme versions of the indigenous origin hypothesis claims that the Dravidians were probably hunter-gatherers rather than farmers (Fuller, 2003), and so proto-Dravidians existed in the Subcontinent during the pre-agriculture era. Given the proposed link between the spread of PFC and the spread of farming based on the farming/language dispersal hypothesis discussed in the Introduction, the presence of PFC in Brahvi might counter the idea of a hunter-gatherer origin of Dravidian. Given the pattern of distribution of PFC languages, it may be unlikely that PFC had emerged in a Dravidian language independently of other PFC languages, unless there is evidence (none so far) to show that such independent emergence is possible.

The finding of PFC in Brahvi, however, may not be as nearly as helpful for settling the disagreements among different versions of the foreign origin of the Dravidian hypothesis. But it could at least be a useful guide against any version with an assumption of a foreign origin from a non-farming population. For example, one account believes that proto-Dravidians originated from Africa and moved to the western part of Iran and the Near East where they also developed Elamo-Dravidian language (McAlpin, 1980, 1981; Southworth and McAlpin, 2013), and from there, they

spread to the Indian subcontinent via Balochistan and the Indus Valley, giving birth to great civilizations (Winters, 2008, 2012). A slightly different view is that Iran is the birthplace and original homeland of Dravidian languages (Palanichamy et al., 2015; Quintana-Murci et al., 2001). The presence of PFC in Brahvi as found in the present study would be evidence disfavoring any version of the foreign origin hypothesis that assumes a non-farming origin either in Africa or elsewhere.

The weaker form PFC in Brahvi as compared to that in Balochi may not be easy to explain. Past research has seen various factors that may lead to the interaction of focus marking with the acoustic marking of other functions. Examples of such interaction are found in Turkish and Uygur where the SOV structure of those languages may have made it hard for both final focus and penultimate focus to be effectively marked prosodically, despite robust PFC marking of earlier focus in an utterance (Ipek, 2011; Wang, Qadir and Xu, 2013). In Japanese, focus seems to interact with lexical pitch accent, so that PFC is present in sentences consisting of accented words (although significant f_0 rises are still present in post-focus regions, (Sugahara, 2002), but it does not occur in sentences consisting of unaccented words (Ishihara, 2011, 2015; Lee and Xu, 2012, 2018). But neither of these seem a likely factor here, given that Brahvi and Balochi are both SOV languages (Barjasteh, 2010; Elfenbein, 2015), and both lack contrastive lexical stress, as reviewed in the Introduction. Interestingly, Rahmani, Rietveld and Gussenhoven (2015) have found that Persian has a word accent that uses f_0 to contrast morphologically simplex versus complex words, and that this accentual contrast is fully neutralized in post-focal locations. Being also an Iranian language like Persian, Balochi is said to have "a fairly strong stress accent" that usually falls on a long vowel or

diphthong, which is usually on the first of several long syllables (Elfenbein (1997: 774). The fairly strong post-focal flattening of F_0 contours in Balochi as seen in Figure 1 suggests that stress accent is likely neutralized in post-focal positions, just like in Persian. The less flattened post-focal Brahvi may suggest a lack of accent neutralization in the language. Further studies are needed to ascertain whether this is indeed the case, given currently limited knowledge about both languages.

The finding of PFC in Brahvi, a Dravidian language, may be relevant to the controversy over the Nostratic Macrofamily hypothesis (Bellwood, 2001; Pederson, 1931). A major problem with the hypothesis is that the construction of such a macrofamily has exceeded the time-depth of traditional methods for assessing the affinity of languages through comparison of their phonology, vocabulary, or morphology (François, 2015; Longobardi, 2009; Nichols, 1996). Based on various estimates, after about 10,000-14,000 years, even the basic vocabularies (i.e., those that change the slowest) would have been replaced, making it virtually impossible to recognize any cross-language similarities that are not due to chance (Gray, 2005; Longobardi, 2009; Nichols, 1996; Ringe, 1992). Focus prosody explored in the present study, in contrast, is an aspect of speech that is likely to change slower than many other linguistic facets. This is because the non-transferability of PFC (Chen, 2015; Chen et al., 2014; Swerts and Zerbian, 2010; Wu and Chung, 2011) means that PFC is unlikely to be involved in accidental transfers frequently seen in cross-linguistic borrowing of other kinds of language properties. PFC may therefore potentially serve as a marker of linguistic affinity with greater time depth than traditional markers. In the case of the present study, for example, the presence of PFC in Brahvi suggests that Dravidian languages may be remotely related to the Indo-Aryan languages spoken in the Indo-Pak subcontinent, despite the dissimilarities in

both phenotype and genetic characteristics between the speakers of Dravidian and Indo-Aryan languages (Ali et al., 2014; Diamond & Bellwood, 2003). Also different from traditional markers, the examination of focus prosody requires strictly controlled experiments and precise acoustic comparisons. The application of this method, as done in the present study and many previous ones, could introduce experimental phonetics into historical linguistics. This could help to increase the scope as well as depth of investigation of language affinity and lineage.

IV. CONCLUSION

Through a carefully designed experiment and systematic acoustic analysis, we have found evidence that both Brahvi and Balochi use post-focus compression (PFC) in the prosodic marking of focus. The finding of PFC in Balochi is consistent with the prediction that the distribution of this prosodic feature is along language family lines. The finding of PFC in Brahvi adds one more language family, Dravidian, as having been tested for the presence of PFC. This addition further increases the resemblance of the cross-linguistic distribution of PFC to the hypothetical Nostratic Macrofamily. Because Brahvi is a satellite Dravidian language isolated in the northern part of the Indo-Pak subcontinent, however, there is a need to further investigate focus prosody in the southern Dravidian languages which form the majority of the language family.

The present study has also demonstrated, as similar studies in recent years, the feasibility and benefit of a new means of testing language affinity and lineage for both contemporary and historical linguistic research, based on an experimental paradigm that focuses on speech prosody.

The experimental aspect makes sure that the findings are replicable, and the focus on prosody allows the testing of linguistic affinity with greater time depth than traditional methods.

V. STATEMENT OF ETHICS

This study observed ethical requirements related to human subjects. All participants of this study were adults (age range: 18-31 years, mean: 23.75 years). Therefore, permission for recording was obtained from participants themselves instead of their parents. First of all, we explained the whole scheme of study without pointing the specific objectives to the participants and then requested them to participate in the study. Only those students who were willing to participate at their own will were selected for participation. Written consent for recording their voices was obtained from those participants who were willing to participate in the experiment. We assured the participants that these recordings will only be heard by the researchers and transcribed for research purpose. They were also informed that only statistical information obtained from acoustic analysis of their productions in summarized form (mean values) will be quoted in the paper. The anonymity of the participants was also assured to them. Later on, they signed a consent form to give us permission to record their productions. Each participant was called for recording at the time of his convenience. They were explicitly told that they were at liberty to leave the experiment at any stage. They were also compensated for their time in terms of money. Since no risk or danger was involved in this study, therefore, the paper was exempt from ethical committee approval. In consultation with the Department of English language and linguistics LUAWMS the team of researchers decided not to approach any Committee for ethical approval of the project.

VI. CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

VII. FUNDING SOURCES

This research project was completed with funding provided by Higher Education Commission of Pakistan under the head of Thematic Research Grants Program.

VIII. AUTHOR CONTRIBUTIONS

Nasir Abbas Syed and Professor Yi Xu conceived and designed the study, Nasir Abbas Syed and Abdul Waheed Shah conducted the experiments; Anqi Xu performed the statistical analysis. All participated in the writing of the manuscript.

References

Abolhasanizadeh, V., Bijankhan, M., and Gussenhoven, C. (**2012**). The Persian pitch accent and its retention after the focus, *Lingua* **122**, 1380-1394.

Alzaidi, M., Xu, Y. and Xu, A. (**2019**). Prosodic Encoding of Focus in Hijazi Arabic, *Speech Communication* **106**, 127-149.

Andronov, M. S. (**1980**). *The Brahvi Language*. (Nauka Publishing House Central, Msocow).

Asher, R. E. (**1982**). *Tamil (Lingua Descriptive Studies)* (North-Holland, Amsterdam).

Azid, M.S.; Xu, Y.: Prosodic focus in Malay without post-focus compression. Pertanika Journal of Social Sciences & Humanities **28**: 91-108 (2020).

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., Eigen, C., and Rcpp, L. (**2015**). Package 'lme4', *Convergence* 12.

Bellwood, P. (**2001**). Early agriculturalist population diasporas? Farming, languages, and genes, *Annual review of anthropology* **30**, 181-207.

Birahimani, A. H. (2021). Reviewing the history and development of aspiration in Eastern Balochi. *Journal of Historical Linguistics*, ***11***(3), 457-498.

Boersma, P. (**2001**). Praat, a system for doing phonetics by computer, *Glot International* **5:9/10**, 341-345.

Botinis, A., Fourakis, M., and Gawronska, B. (**1999**). Focus identification in English, Greek and Swedish, *The 14th International Congress of Phonetic Sciences* (San Francisco), pp. 1557-1560.

Bray, D. (**1977**). *The Brahvi Language: Introduction and Grammar Part 1*. (Brahvi Academy, Quetta).

Campbell, L., and Poser, W. J. (**2008**). *Language Classification: History and Method* (Cambridge University Press, Cambridge).

Chahal, D. (**2003**). Phonetic Cues to Prominence in Lebanese Arabic, in *The 15th International Congress of Phonetic Sciences* (Barcelona), pp. 2067-2070.

Chaubey, G., Metspalu, M., Villems, R., & Kivisild, T. (**2007**). Reply to Winters, *BioEssays **29***, 499-499.

Chen, L., and Yang, Y. (**2015**). Emphasizing the only character: Emphasis, attention and contrast, *Cognition* **136**, 222-227.

Chen, L., Wang, L., and Yang, Y. (**2014**). Distinguish between focus and newness: An ERP study, *Journal of Neurolinguistics* 31, 28-41.

Chen, S.-w., Wang, B., and Xu, Y. (**2009**). Closely related languages, different ways of realizing focus, *Interspeech 2009* (Brighton, UK), pp. 1007-1010.

Chen, Y. (**2015**). Post-focus compression in English by Mandarin learners, *The 18th International Congress of Phonetic Sciences* (Glasgow, UK).

Chen, Y., Guion-Anderson, S., and Xu, Y. (**2012**). Post-Focus Compression in Second Language Mandarin, *Speech Prosody 2012* (Shanghai), pp. 410-413.

Chen, Y., Xu, Y., and Guion-Anderson, S. (**2014**). Prosodic realization of focus in bilingual production of Southern Min and Mandarin, *Phonetica* **71**, 249-270.

Clopper, C. G. and Tonhauser, J. (2013). The Prosody of Focus in Paraguayan Guaraní. *International Journal of American Linguistics* **79**(2): 219-251.

Cooper, W. E., Eady, S. J., and Mueller, P. R. (**1985**). Acoustical aspects of contrastive stress in question-answer contexts, *Journal of the Acoustical Society of America* **77**, 2142-2156.

Dashti, N. (**2012**). *The Baloch and Balochistan: A historical account from the Beginning to the fall of the Baloch State*. (Trafford Publishing, Bloomington).

Delforooz, B. B. (**2010**). *Discourse Features in Balochi of Sistan*: *(oral narratives*). Doctoral Dissertation University of Uppsala.

Delforooz, B. B. (**2019)**. Two Brahui Texts with Glossary and Grammatical Analysis. *Iranian Journal of Applied Language Studies*, Vol 12, No 1, 2020, pp. 89-122.

Diamond, J., and Bellwood, P. (**2003**). Farmers and their languages: the first expansions, *Science* **300**, 597-603.

Dohen, M., and Lœvenbruck, H. (**2004**). Pre-focal rephrasing, focal enhancement and post-focal deaccentuation in French, *The 8th International Conference on Spoken Language Processing* (Jeju, Korea), pp. 1313-1316.

Duan, W., and Jia, Y. (**2015**). Contrastive study of focus phonetic realization between Jinan dialect and Taiyuan dialect, *2015 International Conference Oriental COCOSDA held jointly with 2015 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)*, pp. 47-52.

Eady, S. J., and Cooper, W. E. (**1986**). Speech intonation and focus location in matched statements and questions, *Journal of the Acoustical Society of America* 80, 402-416.

Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R., and Lotts, D. W. (**1986**). Acoustic characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments, *Language and Speech* **29**, 233-251.

Elfbein, J. (**1997a**). Balochi Phonology, In *Phonologies of Asia and Africa*, edited by Alan S. Kaye and Peter T. Daniels (Eisenbrauns, Winona Lake, Indiana). Volume 1, pp. 761-777.

Elfbein, J. (**1997b**). Brahui Phonology, In *Phonologies of Asia and Africa*, edited by Alan S. Kaye and Peter T. Daniels (Eisenbrauns, Winona Lake, Indiana). Volume 1, pp. 797-811.

Elfenbein, J. (1998). Brahvi. In *The Dravidian Languages*, Edited by Sanford B. Steever. London: Routlege. Pp 388-414.

Féry, C. and Fanselow, G., (2020). Prosody of discontinuous nominal phrases in Indian languages. *Journal of South Asian Linguistics*, *10*, pp.60-87.

Féry, C., and Kügler, F. (**2008**). Pitch accent scaling on given, new and focused constituents in German, *Journal of Phonetics* 36, 680-703.

Flege, J. E. (1995). Second language Speech learning: Findings, and Problems. *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues.* W. Strange. Timonium, MD: York Press pp.

François, A. (**2015**). Trees, waves and linkages: models of language diversification, *The Routledge handbook of historical linguistics* (Routledge), pp. 179-207.

Fuller, D. Q. (**2003**). An agricultural perspective on Dravidian historical linguistics: archaeological crop packages, livestock and Dravidian crop vocabulary, in *Assessing the Languaging/Farming Dispersal Hypothesis*, edited by P. Bellwood, and C. Renfrew (McDonald Institute for Archaeological Research, Cambridge), pp. 191-213.

Genzel, S., Renans, A., and Kügler, F. (**2018**). Focus and its prosody in Akan and Ga, *Speech Prosody 2018*, pp. 724-728.

Gordon, M. (2008). The intonational realization of contrastive focus in Chickasaw. *Topic and Focus*: Springer. pp. 69-82.

Gray, R. (**2005**). Pushing the Time Barrier in the Quest for Language Roots, *Science* **309**, 2007-2008.

Hartmann, K. (2008). Focus and tone. *Acta Linguistica Hungarica* **55**(3): 415-426.

Hartmann, K. and Zimmermann, M. (2007). In place-out of place? Focus in Hausa. In *On information structure: meaning and form*. K. Schwabe and S. Winkler. Amsterdam: Benjamins pp. 365-403.

Hartmann, K. and Zimmermann, M. (2007). Focus Strategies in Chadic: The Case of Tangale Revisited. *Studia Linguistica* **61**: 95-129.

Hellmuth, S. (**2006**). Focus-related pitch range manipulation (and peak alignment effects) in Egyptian Arabic, *Speech Prosody 2006* (Dresden, Germany), pp. PS4-12-164.

Himmelmann, N. P. (**2018**). Some preliminary observations on prosody and information structure in Austronesian languages of Indonesia and East Timor, *Perspectives on information structure in Austronesian languages*, pp. 347-374.

Ipek, C. (**2011**). Phonetic realization of focus with no on-focus pitch range expansion in Turkish, *The 17th International Congress of Phonetic Sciences* (Hong Kong), pp. 140-143.

Ishihara, S. (**2011**). Japanese focus prosody revisited: Freeing focus from prosodic phrasing, *Lingua* **121**, 1870-1889.

Ishihara, S. (**2015**). Syntax-phonology interface, *Handbook of Japanese Phonetics and Phonology*, edited by H.Kubozono (Mouton de Gruyter, Berlin).

Jahani, K. (2019). *A Grammar of Modern Standard Balochi*. Uppsala: Uppsala University Press.

Jahani, C., & Korn, A. (2009). Balochi. In Gernot Windfuhr (ed.) *The Iranian Languages* (pp. 634-692). London: Routledge.

Kaiser, M., and Shevoroshkin, V. (**1988**). Nostratic, *Annual Review of Anthropology* **17**, 309-329.

Kokaislova, P. (**2012**). Ethnic Identity of the Baloch People, *Central Asian and the Caucasus Journal of Social and Political Studies **(3)**, 45-55.*

Kolipakam, V., Jordan, F. M., Dunn, M., Greenhill, S. J., Bouckaert, R., Gray, R. D., &Verkerk, A. (**2018**). A Bayesian phylogenetic study of the Dravidian language family, *Royal Society open science*, ***5***, 171504.

Krishnamurti, B. (**2003**). *The Dravidian Languages*. (Cambridge, University Press Cambridge).

Kügler, F. and Skopeteas, S. (2007). On the universality of prosodic reflexes of contrast: The case of Yucatec Maya. *The 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany

Lee, A., and Xu, Y. (**2012**). Revisiting focus prosody in Japanese, *Speech Prosody 2012* (Shanghai), pp. 274-277.

Lee, A., and Xu, Y. (**2018**). Conditional realisation of post-focus compression in Japanese, *Speech Prosody 2018* (Poznań, Poland), pp. 216-219.

Lee, Y.-c., and Xu, Y. (**2010**). Phonetic Realization of Contrastive Focus in Korean, *Speech Prosody 2010* (Chicago), pp. 100033:100031-100034.

Longobardi, G., and Guardiano, C. (**2009**). Evidence for syntax as a signal of historical relatedness, *Lingua* **119**, 1679-1706.

Maskikit-Essed, R., and Gussenhoven, C. (**2016**). No stress, no pitch accent, no prosodic focus: the case of Ambonese Malay, *Phonology* **33**, 353-389.

McAlpin, D. W. (1980). Is Brahui Really Dravidian?. *Annual Meeting of the Berkeley Linguistics Society* (Vol. 6, pp. 66-72).

McAlpin, D. W. (1981). Proto-Elamo-Dravidian: The evidence and its implications. *Transactions of the American Philosophical Society*, *71*(3), 1-155.

McAlpin, D. W. (2015). Brahui and the Zagrosian hypothesis. *Journal of American Oriental Society, 135*(3), 551-586.

McDonough, J. (2002). The prosody of interrogative and focus constructions in Navajo. In *Formal Approaches to Functional Phenomena*. A. Carnie and H. Harley. Amsterdam: John Benjamins pp. 191-206.

Meyer, R., & Mlinek, I. (**2006**). How prosody signals force and focus- A study of pitch accents in Russian yes-no questions. *Journal of Pragmatics*, 38 (10), 1615-1635.

Nichols, J. (**1996**). The comparative method as heuristic, in *The Comparative Method Revised*, edited by M. Durie, and M. Ross (Oxford University Press, Oxford), pp. 39–71.

Palanichamy, M. G., Mitra, B., Zhang, C. L., Debnath, M., Li, G. M., Wang, H. W., Agrawal, S., Chaudary, T. K., & Zhang, Y. P. (**2015**). West Eurasian mtDNA lineages in India: an insight into the spread of the Dravidian language and the origins of the caste system, *Human genetics **134***, 637-647.

Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., and Vasishth, S. (**2008**). Focus, word order and intonation in Hindi, *Journal of South Asian Linguistics* **1**, 55-72.

Pedersen, H. (**1931**). *The Discovery of Language: Linguistic Science in the Nineteenth Century. English translation by John Webster Spargo* (Indiana University Press, Bloomington, IN).

Quintana-Murci, L., Krausz, C., Zerjal, T., Sayar, S. H., Hammer, M. F., Mehdi, S. Q., & Jobling, M. A. (**2001**). Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. *The American Journal of Human Genetics*, ***68***, 537-542.

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

Rahmani H., Rietveld T., & Gussenhoven C. (**2015**) Stress Deafness Reveals Absence of Lexical Marking of Stress or Tone in the Adult Grammar. *PLoS ONE* 10(12): e0143968. https://doi.org/10.1371/journal.pone.0143968

Renfrew, C. (**1988**). *Archaeology and Language* (Cambridge University Press, New York).

Rialland, A., and Robert, S. (**2001**). The intonational system of Wolof, *Linguistics* **39**, 893–939.

Ringe, D. A. (**1992**). On calculating the factor of chance in language comparison, *Transactions of the American Philosophical Society* **82**, 1-110.

Rognoni, L., Bishop, J., and Corris, M. (**2017**). Pashto Intonation Patterns, *Interspeech 2017*, pp. 1228-1232.

Rump, H. H., and Collier, R. (**1996**). Focus conditions and the prominence of pitch-accented syllables, *Language and Speech* **39**, 1-17.

Shen, C., and Xu, Y. (**2016**). Prosodic Focus with Post-focus Compression in Lan-Yin Mandarin, *Speech Prosody 2016* (Boston, USA), pp. 340-344.

Shouse, B. (**2001**). Spreading the Word, Scattering the Seeds, *Science* **294**, 988-989.

Soohani, B., Ahangar, A. A., & Oostendorp, M. van (**2011**). Stress Pattern System in Central Sarawani Balochi. *Iranian Journal of Applied Language Studies*, Vol 3, (1) 151-194.

Southworth, F. C., & McAlpin, D. W. (2013). 30 South Asia: Dravidian linguistic history. *The encyclopedia of global human migration.*

Sugahara, M. (**2002**). Conditions on post-focus dephrasing in Tokyo Japanese., *Speech Prosody 2002*.

Swerts, M., and Zerbian, S. (**2010**). Prosodic transfer in Black South African English, *Speech Prosody 2010* (Chicago).

Taheri-Ardali, M., and Xu, Y. (**2012**). Phonetic Realization of Prosodic Focus in Persian, *Speech Prosody 2012* (Shanghai), pp. 326-329.

Tripathy, V., Nirmala A., Reddy, B. M. (**2008**). Trends in Molecular Anthropological Studies in India, *International Journal of Human Genetics* **8**,1-20.

Wang, B., Wang, L., and Qadir, T. (**2011**). Prosodic encoding of focus in six languages in China, *The 17th International Congress of Phonetic Sciences* (Hong Kong), pp. 144-147.

Wang, B., Zhang, Y., Xu, Y., and Ding, H. (**2017**). Prosodic focus in three northern Wu dialects: Wuxi, Suzhou and Ningbo, *ExLing 2017* (Heraklion, Greece), pp. 117-120.

Winters, C. A. (**2007**). Did the Dravidian speakers originate in Africa? *BioEssays* **27**, 497-498.

Winters, C. (**2008**). Origin and Spread of Dravidian Speakers, *International Journal of Human Genetics* **8**, 325-329.

Winters, C. (**2012**). Origin of the Niger-Congo Speakers. Web med Central *GENETICS 1*-18;3(3): WMC003149. doi: 10.9754/journal.wmc.2012.003149.

Wu, W. L., and Chung, L. (2011). Post-focus compression in English-Cantonese bilingual speakers, *The 17th International Congress of Phonetic Sciences* (Hong Kong), pp. 148-151.

Wu, W. L. and Xu, Y. (**2010**). Prosodic focus in Hong Kong Cantonese without post-focus compression. *Proceedings of Speech Prosody Chicago*, 10040.

Xu, Y. (**1999**). Effects of tone and focus on the formation and alignment of f_0 contours, *J. Phonetics* **27**, 55-105.

Xu, Y. (**2011**). Post-focus compression: Cross-linguistic distribution and historical origin, *The 17th International Congress of Phonetic Sciences* (Hong Kong), pp. 152-155.

Xu, Y. (2013). ProsodyPro — A tool for large-scale systematic prosody analysis, Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), (Aix-en-Provence, France 2013), pp. 7-10.

Xu, Y.; Chen, S.-w.; Wang, B.: Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family? The Linguistic Review **29**: 131-147 (2012).

Xu, Y., and Xu, C. (**2005**). Phonetic realization of focus in English declarative intonation, *J. Phonetics* **33**, 159-197.

Xu, Y., Xu, C. X., and Sun, X. (**2004**). On the Temporal Domain of Focus, *Speech Prosody 2004* (Nara, Japan), pp. 81-84.

Zerbian, S., Genzel, S., and Kügler, F. (**2010**). Experimental work on prosodically-marked information structure in selected African languages (Afroasiatic and Niger-Congo), *Speech Prosody 2010* (Chicago), pp. 100976:100971-100974.