

How cross-linguistic influence affects the use of duration in the production and perception of corrective and non-corrective focus types

Farhat Jabeen and Bettina Braun
Bielefeld University | University of Konstanz

Speakers use corrective focus as an explicit way to correct misunderstandings in communication. We investigate whether immersive contact with a rhythmically different language affects the production and perception of duration as a cue to corrective and non-corrective focus. We tested twenty-eight native speakers and sixty-four native listeners of Urdu, half of whom lived in Germany and used German as a second language, and half lived in Pakistan. German is a stress-timed language with head-prominence marking and makes intensive use of duration to mark corrective focus, while Urdu is a syllable-timed language with edge-prominence marking, which uses duration differently from German to mark focus types. Results showed that the majority language, German, affected focus processing in Urdu differently across modalities: In production, focus marking was not affected by country of residence, while in perception, Urdu speakers living in Germany were more sensitive to duration in the corrective focus context than Urdu speakers in Pakistan. We analyze this as cross-linguistic influence and argue that contact with a stress-timed, head-prominence majority language (here: German) affects the cue weighting in the native language Urdu in perception but not (yet) in production.

Keywords: corrective focus, non-corrective focus, duration, cross-linguistic influence, Urdu

1. Introduction

Interlocutors often tailor their messages to the prior context and to the assumed knowledge of the other participants in the conversation (Lambrecht, 1994; Vallduví & Engdahl, 1996). For instance, they mark “new information” (focus)

differently than information that is expected to be known by the interlocutor (background). This makes it easier for discourse participants to link new information to the mental representation of the discourse. Focus can be wide, encompassing an entire utterance (1a), or narrow, marking only a single word. Narrow focus can furthermore be non-corrective (1b) or corrective (1c). Methodologically, these types of foci are often elicited by questions, cf. (1a)–(c). Semantically, focus indicates “the presence of alternatives that are relevant for the interpretation of linguistic expressions” (Krifka, 2008, p.18). Depending on the language, breadth of focus is signalled by morpho-syntax, word order, and/or prosody (Jabeen, 2022; Ladd, 1980; Lambrecht, 1994; Vallduví & Engdahl, 1996).

- (1) a. Q: What happened? broad focus
 A: [Anne got a new JOB.]_F
 b. Q: Who got a new job? non-corrective narrow focus
 A: [ANNE]_F got a new job.
 c. Q: Did Mary get a new job? corrective narrow focus
 A: [ANNE]_F got a new job.

Example (1a) indicates broad focus on the entire phrase, with the accented word (shown in capital letters) as the last word in the phrase (Chomsky and Halle, 1968). Accented words in English are realized with *fo* movement and longer duration (Breen et al., 2010). Both Examples (1b) and (1c) present narrow focus on the subject. They differ information-structurally in that (1b) is a non-corrective (neutral) narrow focus, while (1c) represents a corrective narrow focus.¹ We study the phonetic differences between (1b) and (1c), in production and perception in two groups of Urdu speakers: One living in their home country (Pakistan) and the other group living abroad (Germany).

In a world in which multilingualism is the norm (Matras, 2013), it is important to understand how daily experience with different languages affects the use of prosodic cues by the users of a language, because changes in cue usage and cue weighting may lead to sound change (for first language (L1) influence cf. Dmitrieva, 2019; Schertz et al., 2020; for general sound change cf. Harrington et al., 2018; Llompart & Reinisch, 2019). The acquisition of cue weighting poses challenges for second language (L2) learners (Trouvain & Braun, 2020), both for segmental and prosodic information (Schwartz, 2022; Atterer & Ladd, 2004; Mennen et al., 2022; Zahner-Ritter et al., 2022; Maastricht et al., 2019). To under-

1. Note that some authors use the term contrastive focus to refer to corrective focus. This may be misleading as not every contrastive focus is necessarily corrective (cf. Krifka, 2008, discussion in Braun & Biezma, 2019).

stand the limits of cross-linguistic interference, it is important to investigate the effects of an L2 on an L1 (Ribbert & Kuiken, 2010).

This issue, however, has received less attention in the literature and the existing studies test mostly the production aspect. From these studies, we know that prolonged exposure to a majority language with little use of L1 may lead to attrition of the L1 at the segmental (Flege, 1987; Major, 1992) and prosodic levels (de Leeuw, 2019; de Leeuw et al., 2012; Mennen & Chousi, 2018). Schmid (2011) defines attrition as “the onset of L2 influence on L1” (p.5) and describes it as a gradual process that allows cross-linguistic influences to settle in the language use of bilingual speakers. de Leeuw et al. (2012) further qualifies attrition as a structural change in L1 following the sequential acquisition of L2. Note that the term “attrition” is not uncontroversial as it highlights the aspect of loss. We approach the question of attrition in terms of a change in cue weighting, both in production and in perception. In that sense, attrition may lead to assigning different weights to the cues present in the L2 vs. the L1. This cue re-weighting may play out differently in production and perception (Boersma & Chládková, 2011; Broselow & Kang, 2013), with changes being observed earlier in production or in perception, with conflicting evidence so far.

In this paper, we investigated if the contrasts acquired in the L2 affected the use of acoustic cues in the L1. Specifically, we compared the marking of focus type (corrective vs. non-corrective focus, cf. (1b) vs. (1c)) by native speakers of Urdu, tested in Germany and in Pakistan. The speakers tested in Germany used German regularly. As outlined in more detail in Section 2, German, the majority language in Germany, and Urdu differ in terms of rhythmic organisation. We therefore tested whether native speakers of Urdu differed based on the country of residence in their use of syllable duration as a cue to distinguish between corrective and non-corrective narrow focus in their native language, both in production (Experiment 1) and in perception (Experiment 2).

In what follows, we offer a brief overview of focus marking in Urdu (subSection 2.1). To be able to derive hypotheses on the role of country of residence, we also discuss the focus marking of Urdu’s main contact language in Germany (German). In subSection 2.2, we outline prior work on the influence of majority languages on the native language before we lay out the hypotheses in subSection 2.3. In Section 3, we present the details of the production experiment followed by a perception experiment in Section 4. The implications of our findings are discussed in Section 5 and finally, a conclusion is offered in Section 6.

2. Background

The target language for the production and perception experiments is Urdu. It is the national language of Pakistan, a *Lingua Franca* in the country (Dua, 2009), and the most frequently used medium of instruction in educational institutes (Rahman, 2019). The canonical word order in Urdu is Subject–Object–Verb (Kidwai, 2000), but the constituents can be moved to different positions to mark information structure (Jabeen, 2022; Manetta, 2012).

An important contrast between German, the majority language in Germany, and Urdu is that German is a stress–timed language, while Urdu is a syllable–timed language. Briefly, German has a phonological length contrast in vowels (accompanied by a difference in vowel quality) and it reduces unstressed syllables in terms of segmental quality, intensity, and duration (Delattre, 1969). Moreover, German is a head–prominence language, in which pitch accents (used for focus and prominence marking etc.), are associated with stressed syllables (Jun, 2012). Urdu, on the other hand, is a syllable–timed language (Hussain, 1997) with little reduction (in vowel quality and duration) in unstressed syllables. Urdu has a phonological distinction between long and short vowels as well as consonants (Saleem–ud–Din & Anjum, 2013). This leads to minimal pairs in consonantal length ['pə.t̪ɑ] ‘address’ and ['pə.t̪ɑ] ‘leaf’ and vocalic length [bi:n] ‘without’ and [bi:n] ‘flute’. Furthermore, Urdu is an edge–prominence language (Jabeen, 2019), where prominence is used to mark phrase edges instead of phrase heads. Jabeen and Delais–Roussarie (2020) have shown that Accentual Phrases in Urdu are produced with a rising *f₀* contour. Hindi, a closely related language, has been shown to have a similar prosodic structure (Féry et al., 2016; Patil et al., 2008). The majority language for our participants in Pakistan is Punjabi,² another syllable–timed language (Dhillon, 2010; Vijaykrishnan, 2003). Given the rhythmic and intonational differences between Urdu and its contact language in Germany (German), we present a brief overview of the use of duration to mark focus in these languages.

2. Pakistan has many regional languages and Urdu as the *Lingua Franca*. Depending on the region, Urdu is in a contact situation with many languages. Our respondents were recorded in Punjab and the majority language there is Punjabi (Eberhard et al., 2020). English is the official language of Pakistan, and our respondents would have been exposed to Pakistani English. This could arguably influence their production of focus in Urdu. However, Jabeen (2022)’s analysis of the prosodic marking of wide and non–corrective narrow focus is based on speakers living in Pakistan. Her data shows that the use of acoustic cues to mark focus in Urdu differs from that in English. In any case, our analysis is concerned with differences between Urdu speakers living in Germany and Pakistan, notwithstanding the influence of regional languages or English on the variety of Urdu spoken in Pakistan.

2.1 Duration as a cue to focus marking in Urdu–Hindi and German

The acoustic realization of focus breadth and types of narrow focus in Urdu is under-studied. Generally, the research on Hindi prosody is considered to be applicable to Urdu as well (Butt et al., 2016, 2020), so some findings for Hindi are included here. It must be noted that the speakers in most of the studies reported here were recorded outside Pakistan/India,³ therefore an effect of community language (German or English) on their findings cannot be excluded. Genzel and Kügler (2010) analyzed the prosodic marking of wide focus vs. corrective narrow focus – cf. (1a) vs. (1c) – by Hindi speakers living in Germany. The authors reported that although the stressed syllable contributed the most to focus induced lengthening, all the syllables in correctively focused words were produced with longer duration compared to their counterparts in wide focus. Furthermore, Puri (2018) analyzed the prosodic realization of wide and corrective narrow focus by late learners of English and simultaneous bilingual speakers of Hindi and Indian English, all tested in the US. She found that both groups of speakers used longer duration of stressed vowels in Hindi to mark corrective focus compared with wide focus.

Jabeen (2022) investigated the use of intonational and durational cues to mark the difference between wide focus and non-corrective narrow focus by Urdu speakers living in Pakistan – cf. (1a) vs. (1b). In terms of duration, speakers lengthened the noun in non-corrective narrow focus compared to wide focus. Regarding the marking of corrective vs. non-corrective narrow focus – cf. (1b) vs. (1c) – Jabeen and Braun (2018) tested a small sample of Urdu speakers residing in Germany. They reported that the correctively focused object nouns were produced with significantly longer syllable durations than in non-corrective focus.

The existing analyses of focus marking show that Urdu, similar to Hindi, uses duration to distinguish between focus breadth i.e., wide and narrow focus on the one hand and between corrective and non-corrective narrow focus on the other. Both Urdu and Hindi speakers living in Germany elongate all the syllables in the target words produced in the corrective focus context (Genzel and Kügler (2010) compared to wide focus; Jabeen and Braun (2018) compared to non-corrective focus). Note that only Jabeen and Braun tested for differences between focus types i.e., corrective vs. non-corrective focus. Moreover, studies on the use of prosodic cues to investigate the perception of focus types are missing.

3. Jabeen (2022) is an exception as her data was collected in Pakistan. But she did not include corrective focus in her analysis.

For German, there have been many studies on the acoustic correlates of focus breadth as well as different types of narrow focus. German speakers are claimed to use a three-way distinction for duration: Kügler (2008) found that the stressed syllables of correctively focused words were longer than their counterparts in non-corrective narrow focus, which in turn were longer than in wide focus (Cangemi et al., 2015; Kügler & Gollrad, 2015).

At first glance, German and Urdu appear to be similar in their use of duration to signal focus breadth and type. However, the data for Urdu and Hindi was collected in different countries (Germany, US) with different majority languages. This may cast a new light on the findings of these studies as these participants were exposed to languages rhythmically and intonationally different from their native language. Therefore, it cannot be ruled out that these speakers' use of duration is affected by their exposure to the majority languages. In the next subsection, we briefly summarize the existing literature to illustrate the cross-linguistic influence on prosody.

2.2 Cross-linguistic influence on prosody from L2 to L1

Multilingual speakers need to acquire and adequately use the phonological inventory and the phonetic realization of the components for each of their languages. Since this paper is concerned with speakers and listeners who are surrounded by a majority language that is different from their L1 in terms of its rhythmic structure and intonation, we focus our literature review on this population. The existence of two or more language systems may lead to cross-linguistic influence from one language, typically the more frequently used one, to the less frequently used language (Beddor, 2009, 2012; Flege & Bohn, 2021; Piccinini & Arvaniti, 2021). Language users' productions are also affected by the patterns they are exposed to as L2 listeners (Chen et al., 2021; Tremblay et al., 2021). Research lags behind when it comes to the influence of the L2 on the L1 prosody. Most of the research on bidirectional influences between the L1 and the L2 is concerned with intonation (e.g., Mennen, 2004; Mennen et al., 2022; Maastricht et al., 2016) and focuses on production.

Among the few studies investigating the influence of the L2 on the L1 rhythm, Henriksen (2016) reported the existence of bidirectional influence between L1 (Spanish, syllable-timed) and the later acquired L2 (English, stress-timed). He showed that these speakers' use of consonant and vowel durations in L2 English was similar to that of native American English speakers, thus indicating a close approximation with native-like (stress-timed) rhythm. However, their production of rhythm in L1 Spanish differed from that of monolingual Spanish L1 speakers. Henriksen argued that this was indicative of L2 influence on L1 rhythmic

patterns. Notably, these bilingual speakers were recorded in Michigan, US, where English is the most frequently used community language. Hence the rhythm patterns exhibited by these bilingual speakers may also reflect the influence of the majority language on the speakers' L1.

The discussion above shows that L1 prosody is flexible and that transfer between languages may be bidirectional. Moreover, prosodic transfer is influenced by the extent of contact between languages and their phonetic and phonological similarities/differences (Elordieta & Romera, 2021). These factors may lead to L1 attrition i.e., L2 usage leading to a change in L1 structures (de Leeuw et al., 2012), or transfer i.e., L1 structures affect the use of cues in L2, or both. While the change in structures may occur eventually, the re-weighting of cues may happen ahead of an absolute change. Notably, most of the research in this regard is concerned with the production of rhythmic patterns and intonation. There is a lack of studies investigating the influence of L2 on speakers' L1 in the perception of durational cues of focus marking or the re-weighting of this cue. Our study aims to address this by investigating the influence of majority language (German) on speakers' use of duration to mark corrective and non-corrective focus types in Urdu (L1) from both production and perception perspectives. If such an influence exists, we further discuss if it has led to attrition in terms of change in L1 structures or variation in cue-weighting. Our hypotheses are presented in the next subsection.

2.3 Hypotheses

Previous research by Jabeen and Braun (2018) showed that Urdu speakers living in Germany produce longer syllable durations in correctively focused nouns compared to non-corrective narrow focus. However, their sample size was small (only four speakers) and did not include native speakers in their home country, Pakistan. It is possible that the use of syllable duration differs among speakers living in different countries. We build on Jabeen and Braun's preliminary results and test 28 native speakers and 64 native listeners of Urdu to investigate whether country of residence (and hence contact with a rhythmically different language) affects

- the durational marking of corrective vs. non-corrective focus by native speakers of Urdu.
- Urdu-speaking listeners' sensitivity to durational cues to corrective vs. non-corrective focus.

Based on studies of cross-linguistic influence and the prosodic marking of focus in the two countries of residence, we test the following hypotheses:

- H1: In production, a rhythmically dissimilar majority language, which strongly uses duration to mark the differences between corrective and non-corrective focus (German), leads to a stronger use of the durational cue in Urdu spoken in Germany compared with Urdu spoken in Pakistan.
- H2: In perception, this rhythmically dissimilar majority language, leads to a higher sensitivity for duration in judging the type of focus in Urdu used by speakers in Germany than Pakistan.

3. Production experiment

In this experiment, we manipulate focus context (corrective vs. non-corrective, within-subjects) and country of residence (Germany vs. Pakistan, between-subjects). As the noun and the corresponding case marker form a prosodic phrase together (Jabeen & Delais-Roussarie, 2020), we analyzed the duration of the first and second syllables of the target noun as well as the following case marker.

3.1 Methods

3.1.1 *Participants*

Twenty-eight female Urdu speakers (12 living in Pakistan; 16 living in Germany) were recorded. Two German residents were excluded because most of their target sentences contained hesitations, which would have made duration measurements invalid. All the participants filled a questionnaire providing information about their age, gender, mother tongue, education, and the extent of their Urdu usage for personal and official communication. The German residents further indicated their duration of stay and contexts (personal, official, academic) of German usage. The speakers were aged between nineteen and thirty-five years (mean Pakistan: 21 years; mean Germany: 30 years) and held a university degree. The Pakistani residents used Urdu extensively for everyday and official communication. The Urdu speakers living in Germany had grown up in Pakistan using Urdu as a medium of formal instruction and in their daily lives. They had moved to Germany as young adults for academic purposes and had been living there for at least nine months and at most seven years at the time of recording. While living in Germany, they were still using Urdu for communication with friends and family in Pakistan. Importantly, they were also using German for official communication and in their everyday life.

3.1.2 Materials

The stimuli for the production experiment comprised twelve Urdu SOV sentences with the object as the target noun placed at the immediately preverbal position (See Appendix A for the full dataset). This is the canonical word order in Urdu and the focused constituents are preferably placed at the preverbal position (Jabeen, 2022). All the target nouns were bisyllabic trochaic words ('CV:CV), followed by a monosyllabic accusative case marker ("ko"). For each target sentence, we constructed two contexts, one eliciting a corrective and one a non-corrective reading for the target sentence. (2a)–(b) show the English translation of one pair of contexts for one of the target sentences, presented in (2c).

- (2) a. **Non-corrective focus:** Your uncle bought a new computer but sold it after a few days. Your brother asks if your uncle had shown that computer to anyone before selling it. You say:
- b. **Corrective focus:** Your uncle bought a new computer but sold it after a few days. Your brother thinks that the uncle had shown that computer to Yaseen before selling it. You correct him and say:
- c. 'ma.mu=ne ['lɛ.la=ko]_F 'd̪i.kʰa 'd̪i.ja t̪ʰa
 uncle Laila shown given was
 "(The) uncle had shown (it) to Laila."

We further constructed twenty-eight fillers, consisting of thirteen statements and 15 numeric sequences, such as "2 + 10 = ?". All the fillers represent wide focus in the experiment (elicited without context sentences), thus counterbalancing the number of focus contexts in the experimental trials.

3.1.3 Procedure

Focus context was manipulated within-subjects and within-items. We created one list containing the non-corrective focus context and half of the fillers in the first block and the corrective focus context and the other half of the fillers in the second block. As corrective focus is claimed to be more marked than non-corrective focus (Zerbian, 2015), we hypothesized that once the speakers are exposed to the marked focus type, they adopt its marking across the board. We decided to use the current set up to avoid this potentially confounding factor. Within each block of focus context, the target items were presented in a random order. For all the Pakistani and three of the German residents, the data was recorded using a head-mounted Shure microphone and a ZoomH6 recording device at the sampling frequency of 44,100Hz (16Bit) in a quiet room. The remaining German residents were recorded with their computers' microphones using Zoom as it was not possible to meet them in person due to the Covid Pandemic related restrictions. The speakers were asked to read the context and the stimulus sentences from MS.

PowerPoint slides presented on a screen. The participants controlled the pace of the experiment themselves. The average duration of the recording sessions was eleven minutes.

3.1.4 *Data analysis*

3.1.4.1 *Acoustic analysis*

The syllable boundaries in the target sentences were labelled manually using Praat (Boersma & Weenink, 2022). Segmentation was performed by the first author, a native speaker of Urdu, using the broadband spectrogram and standard segmentation criteria (Turk et al., 2006). Thirty-four target sentences (2%) produced by German residents were removed from the analysis as the speakers either did not pause between the context and the target sentence ($n=30$; 16 corrective, 14 non-corrective) or hesitated before producing the target noun ($n=4$; 2 corrective, 2 non-corrective). The former category was excluded as variability in pausing led to differences in prosodic phrasing and, consequently, phonetic realization. The remaining 254 items were used for the analysis explained below.

3.1.4.2 *Statistical analysis*

Prior to analysis, we normalized the syllable durations by dividing them by the overall sentence duration. This accounted for differences in speaking rate and yielded roughly normally distributed data. Finally, the relative durations were multiplied by 100 as percentages are easier to comprehend for the readers. We first fitted an omnibus Linear Mixed-Effects Regression model (Baayen et al., 2008) with focus context (corrective vs. non-corrective), country of residence (Germany vs. Pakistan), and syllable number (1st or 2nd syllable of the noun, case marker) in R (R Core Team, 2023) using “lme4” (Bates et al., 2015) and “lmerTest” (Kuznetsova et al., 2017). Trial number was added as a control predictor, both as main effect and in interaction with focus context as participants may lose interest during the experiment and this lack of attention could lead to differential use of acoustic cues for the two focus contexts over the course of the experiment. The initial models contained participants and items as crossed random effects (random intercepts only). The random slopes for focus context were added if this more complex model converged and if this improved the fit of the model determined on the basis of the log-Likelihood comparison of models using the `anova()`-function. In the next step, non-significant interactions ($p > 0.05$) and non-significant main effects that did not occur in significant interactions were removed. Finally, data points with residuals larger than 2.5 standard deviations were eliminated from the dataset and the model was refitted. This last step excluded less than 2% of the data. In models with interactions, we used the F-statistics of the `anova()`-output (Satterthwaite’s method) to determine whether main

effects or interactions were significant. In case of significant interactions, we then split the models by one of the factors and tested for effects of the other factor. In this case, we report the Estimates and slopes from the summary-output to give an indication of the direction and strength of potential effects.

3.2 Results

The three-way-interaction between syllable number, country of residence, and focus context was not significant ($p=0.7$). Neither was the interaction between residence and focus context ($p=0.7$), the interaction between focus context and trial number ($p=0.7$), nor the control predictor trial number ($p=0.8$). These terms were removed. The final model (see Appendix B) did not show significant main effects of focus context ($F(1,1711)=1.07, p=0.2$), but there was a significant effect of syllable number ($F(2,1705)=1168, p<0.0001$), country of residence ($F(1,24)=5.2, p=0.03$) and significant interactions between syllable number and residence ($F(2,1705)=10.9, p<0.0001$) as well as syllable number and focus context ($F(2,1705)=6.1, p=0.002$). These interactions are visualized in Figure 1. To investigate these two-way-interactions further, we fitted separate models for each syllable.⁴

For the FIRST SYLLABLE of the target noun, there was no interaction between focus context and country of residence ($F(1,545)=0.03, p=0.8$), but main effects of country of residence ($\beta=1.05, SE=0.44, t=2.3, p=0.02$) and focus context ($\beta: 0.6, SE=0.13, t=4.7, p<0.0001$): The first syllable was significantly longer for Pakistani residents (Est. 13.35) compared with the German ones (Est. 12.3) and the first syllable was longer in corrective focus (Est. 12.3) compared to non-corrective focus context (Est. 11.7).

As for the relative duration of the SECOND SYLLABLE of the focused noun, there was a significant interaction between focus context and country of residence ($F(1,544)=4.9, p=0.02$). Figure 1 illustrates that the German residents produced the second syllable with slightly longer duration in corrective focus (Est. 9.0) than in non-corrective focus context (Est. 8.8). Urdu speakers living in Pakistan, however, showed an opposite trend (corrective Est. 9.1; non-corrective Est. 9.28). Note, however, that separate models for each country of residence showed that these differences merely approached significance ($p=0.09$ for Pakistani residents, $p=0.1$ for German residents).

4. For speaker-based differences in the use of duration for focus marking across syllables, see Figure 3 in Appendix C.

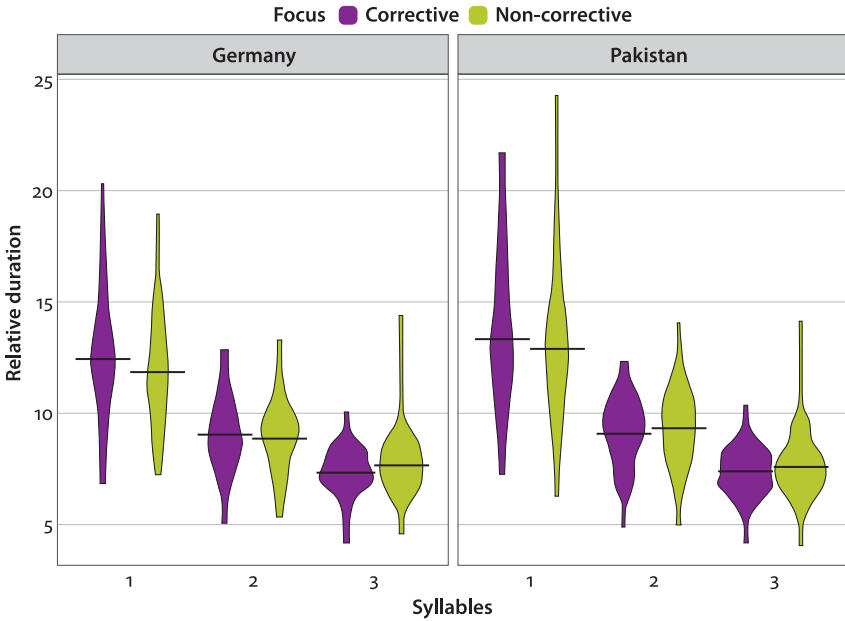


Figure 1. Relative duration (syllable duration/sentence duration \times 100) of syllables in the focused noun (1, 2) and the case marker (3). The horizontal crossbars indicate the mean

The CASE MARKER showed no interaction between focus context and residence ($F(1, 545), p=0.6$), no effect of country of residence ($F(1, 24), p=0.9$), but a significant effect of focus context, which was opposite to the first syllable of the noun. It was shorter for corrective focus (Est. 7.3) than for non-corrective focus context (Est. 7.4). The statistical analysis showed that this difference was significant ($\beta: 0.16, SE=0.001, t=2.2, p=0.02$).

3.3 Discussion

Urdu speakers in both countries produced the first syllable of the focused noun longer in corrective than in non-corrective focus and the case marker shorter. There was weak evidence that the second syllable of the noun was produced differently depending on the country of residence: The German residents slightly lengthened the second syllable of the noun in corrective compared to the non-corrective focus context, while Pakistani residents slightly shortened it. This opposite effect of focus context dependent on country of residence was significant for the second syllable, but not the effect of focus context for this syllable. Furthermore, the three-way-interaction between country of residence, focus context,

and syllable number was not significant either, weakening the effects of country of residence on focus marking.

One may question whether the statistical power was high enough to detect a three-way-interaction between country of residence, focus context, and syllable number at all. A post-hoc power analysis, using the *simr* package (Green & MacLeod, 2016) in R, indicated that the power for detecting the significant three-way-interaction was 84%. This is the typical power used in Linguistics, so it is unlikely that the lacking three-way-interaction can be attributed to a lack of statistical power. Nevertheless, it may be interesting to replicate the experiment with a larger sample size. The findings reported here show that there is no evidence in favour of H₁, which predicted larger duration differences between corrective and non-corrective focus for German than Pakistani residents.

It is conceivable that participants with a longer stay in Germany have a higher usage of German and that they may differ more from Urdu speakers tested in Pakistan than those with less usage. The current dataset does not support this, at least not for the variable duration of stay. A post-hoc analysis for the subset of German residents showed no interaction between duration of stay, focus context, and syllable number ($p > 0.3$). Therefore, we conclude that at the group level, Urdu speakers in Germany do not differ significantly from Urdu speakers in Pakistan regarding the use of duration for focus marking.

In the current study, focus context was manipulated between-blocks, i.e., participants produced the non-corrective item in the first block and the corrective item in the second block. Since each focus indicates the presence of alternatives (Krifka, 2008), it is easy for speakers to construe alternatives once they are made aware of their existence. This is what the corrective focus contexts do: They explicitly state an alternative. Presenting the non-corrective focus contexts reduces the likelihood that participants construe such alternatives, thus minimizing the effects of focus context. Alternatively, this fixed order of focus contexts (non-corrective before corrective) may have reduced participants' awareness of the contrast, hence weakening the effects of focus context. Potential block order effects are an empirical question in the pragmatics-phonetics interface that need to be investigated further (Einfeldt et al., 2023).

The current data suggests that German resident Urdu speakers did not differ systematically from Urdu speakers living in Pakistan, although the German residents were exposed to a contact language (German) which uses duration to distinguish between corrective and non-corrective focus. Our findings show that the previously mentioned studies on Urdu and Hindi focus marking (Genzel & Kügler, 2010; Jabeen & Braun, 2018) testing speakers outside of Pakistan seem partly valid. Genzel and Kügler as well as Jabeen and Braun had reported the lengthening of all the syllables in target words produced in corrective focus.

However, the lengthening of only the first syllable in our data alludes to an influence of the majority language (German) on the findings reported by Genzel and Kügler and Jabeen and Braun. It is possible that cross-linguistic influence in production takes time to establish (Schmid, 2011) and that Urdu listeners in the two countries of residence already show differential sensitivity to the duration cue. In the next section, we investigate the use of syllable duration in the perception of focus contexts.

4. Perception experiment

In this experiment, we manipulate presented focus context (corrective vs. non-corrective context, manipulated within-subjects), recorded focus type (original recorded in corrective or non-corrective context, used as basis for manipulation, manipulated within-subjects), duration condition (lengthened or shortened focused noun, within-subjects), and country of residence (Germany vs. Pakistan, manipulated between-subjects). To test participants' use of duration for the perception of focus, we analyze the perceived fit between context and auditory stimulus on a 5-pt-Likert scale.

4.1 Methods

4.1.1 *Participants*

Sixty-four⁵ speakers of Urdu took part in the perception experiment, twenty-nine living in Germany (5 females, 22 males) and thirty-five in Pakistan (18 females, 17 males). Participants' age ranged between twenty and fifty years (mean age Germany: 25 years, mean age Pakistan: 27 years). All participants reported to hold a university degree and used Urdu in their personal life. The German residents also used German for everyday and official communication. Due to logistic reasons, the participants were different from those in the production experiment.

4.1.2 *Stimuli*

As stimuli, we used the same set of twelve SOV sentences, which were produced anew by one of the participants of the production experiment. The speaker was

5. Originally, we had collected data from seventy-six participants. However, due to a calculation error by the experimental set up, the response times for 12 Pakistani residents were not calculated.

chosen because she had clear articulation and an agreeable voice with little glotalization. She had lived in Germany for three years at the time of recording. She was recorded in a sound-attenuated booth with a head-mounted microphone. The recording conditions and the equipment were identical to the lab-based procedure reported in the production experiment. Table 1 presents the mean absolute and relative durations of the syllables in the target noun and the following case marker produced by the model speaker (henceforth “recorded focus type”).

Table 1. Average absolute syllable duration (in seconds) of the focused nouns and case markers in the two recorded focus types. The relative duration (syllable duration / sentence duration \times 100) is given in the parentheses

Context	1st syllable	2nd syllable	Case marker
Corrective	0.238 (13.1%)	0.158 (8.7%)	0.132 (7.3%)
Non-corrective	0.195 (11.5%)	0.132 (7.8%)	0.136 (8.0%)

These original recordings are expected to contain the full set of cues for the respective focus context. To explicitly investigate the impact of duration differences in the recorded focus types, we created two manipulated conditions, generated by PSOLA resynthesis (Moulines & Charpentier, 1990):

- Long noun duration condition: The first and second syllables of nouns produced in corrective focus were shortened (short noun duration condition in recorded corrective focus type), while the following case markers were lengthened.
- Short noun duration condition: The first and second syllables of the nouns produced in non-corrective focus were lengthened (long noun duration condition in recorded non-corrective focus type), whereas their following case markers were shortened.

The shortening and lengthening of syllables were based on the average ratio of the durations for each syllable produced in non-corrective and corrective focus, respectively. For instance, the first syllable had an average duration of 0.238s in the corrective recording and 0.195s in the non-corrective recording. The ratio between the two ($0.238s/0.195s = 1.22$) was used to lengthen all first syllables in the non-corrective recordings (see first row of Table 2) and the inverse ratio to adapt all first syllables of the corrective recordings.

For the perception experiment, we used the two contexts (corrective and non-corrective, henceforth “presented focus context”) and the target sentences recorded therein. In the experiment, there were four conditions: Two conditions with their original duration (“long noun duration”, originally produced in a correc-

Table 2. Average ratio used to manipulate the duration of syllables in the recorded focus types

Effect of manipulation	1st syllable	2nd syllable	Case marker
from non-corrective focus utterance to long noun duration condition	1.22	1.19	0.97
from corrective focus utterance to short noun duration condition	0.81	0.83	1.02

tive focus context and “short noun duration”, originally produced in non-corrective focus context) and two conditions with manipulated syllable duration (“short noun duration”, in otherwise corrective focus type and “long noun duration” in otherwise non-corrective focus type). We further created fourteen *wh*-questions and respective contexts to be used as fillers. This ensured that the target sentences as well as the fillers were always presented with context sentences. To have all the stimuli synthetically modified (and avoid confounds), the overall *f₀* value of utterances with original duration was raised by 10Hz. This shift in *f₀* does not affect the duration of syllables. It was an arbitrary number that was not too disruptive. Others have raised *f₀* by a factor (cf. Braun et al., 2011).

4.1.3 Apparatus and procedure

Presented focus context and recorded focus type were manipulated within-subjects and within-items, i.e., each target sentence was presented in a non-corrective as well as in a corrective focus context and had been originally recorded in a corrective and non-corrective context. Duration condition was manipulated within-subjects, but between-items such that each participant received the long condition for half of the items and the short condition for the other half. This resulted in eight experimental conditions: 2 presented focus contexts × 2 recorded focus types × 2 duration conditions (long vs. short noun duration). We set up two lists, manipulated between-subjects, see Table 3. Each list comprised forty-eight target sentences and fourteen fillers (*wh*-questions) presented in a random order.

The experiment was set up using a web-based interface created on the Xojo platform. The participants were informed in Urdu that the experiment aimed to investigate focus realization in different contexts in Urdu. They were also made aware that the target sentences were synthetically modified. The task instructed them to silently read the context presented on the screen, play the sentence, and listen to the recording carefully. Then they were required to rate the naturalness of the sentence in the given context on a five-point Likert scale. A target sentence could not be played more than three times in total. To make sure that the par-

Table 3. Contexts and items in two experimental lists

Presented focus context	Recorded focus type	Duration condition	Items	
			List 1	List 2
Non-Corrective	Non-Corrective	Long	1-6	7-12
Non-Corrective	Non-Corrective	Short	7-12	1-6
Non-Corrective	Corrective	Long	1-6	7-12
Non-Corrective	Corrective	Short	7-12	1-6
Corrective	Non-Corrective	Long	1-6	7-12
Corrective	Non-Corrective	Short	7-12	1-6
Corrective	Corrective	Long	1-6	7-12
Corrective	Corrective	Short	7-12	1-6

ticipants listened to a target sentence before rating its naturalness, the web interface did not allow the rating to take place before the sentence was played at least once. The perception experiment was followed by a short questionnaire. They were asked to indicate their gender, age, mother tongue, and education. We also measured response times from the onset of the visual display of the context until the rating of a sentence was provided. Participants were not instructed to respond quickly or placed under a time constraint. The average duration of the experiment was ten minutes.

4.1.4 Data analysis

Prior to analysis, we removed data points with response times smaller than four seconds, because it is impossible to read the context in less than two seconds and provide naturalness ratings to sound files that had an average duration of two seconds. This removed 8% of the data from the analysis (125 data points from German and 103 data points from Pakistani residents). For the statistical analysis, we fitted Linear Mixed-Effects Regression⁶ models in R. We used participants' ratings as dependent variable and presented focus context (corrective vs. non-corrective context), recorded focus type (original corrective or non-corrective utterance), duration condition (long vs. short noun duration), country of residence, as well as their interactions as fixed factors. As random effects, we included participants (nested under list) and items as crossed random effects. The statistical modelling was similar to that reported in Experiment 1.

6. Likert-scales are not strictly continuous variables. However, simulation studies have shown that Linear Mixed-Effects Regression models provide similar outcomes (Brauer & Curtin, 2018).

4.2 Results

The results showed a significant three-way-interaction between presented focus contexts, noun duration and country of residence ($F(1,2304)=4.4, p=0.03$), an interaction between presented focus context and recorded focus type ($F(1,2299)=13.7, p=0.0002$) and a main effect of recorded focus type ($F(1,59)=20, p<0.0001$), see Figure 2. The complete model output is provided in Appendix D. The main effect of recorded focus type (upper row vs. lower row in Figure 2) indicates that corrective focus utterances resulted in overall higher rating than their non-corrective counterparts, which is further qualified by the interactions.

To investigate the three-way-interaction between presented focus contexts, noun duration, and country of residence further, we calculated separate models for the two countries of residence. For German residents, there was an interaction between presented focus context and duration ($F(1,1122)=6.6, p=0.01$). Separate models for each presented focus context showed that German residents rated the long noun durations as more natural than short noun durations in the corrective focus context ($\beta=0.21, SE=0.06, t=3.4, p=0.0005$), while there was no effect of duration in non-corrective focus contexts ($p>0.8$). Note that this interaction was independent of the recorded focus type, i.e., not influenced by other prosodic information in the utterances. For Pakistani residents, there was no interaction between presented focus context and duration ($F(1,1370)=0.06, p=0.7$), but a significant interaction between presented focus context and recorded focus utterance ($F(1,1370)=5.5, p=0.01$) as well as main effects of presented focus context ($F(1,1370)=23, p<0.0001$) and recorded focus type ($F(1,1370)=41.2, p<0.0001$). Separate models for each presented focus context showed that there was a preference for the corrective focus type in both focus contexts, but this preference was stronger in corrective focus type ($\beta=0.32, SE=0.08, t=4.0, p=0.0003$) than in non-corrective focus type ($\beta=0.14, SE=0.04, t=3.3, p=0.0009$). Crucially, Pakistani residents did not show an effect of duration ($F(1,1477)=0.04, p=0.8$), nor an interaction between presented focus context and duration ($F(1,1371)=0.06, p=0.7$).

4.3 Discussion

The results of the perception experiment lend support to H2. It is intriguing that the Pakistani residents paid more attention to the original focus recording and not to the duration manipulation, while German residents showed the opposite pattern. This differential sensitivity indicates that the Pakistani residents were sensitive to other prosodic aspects of the utterance (cf. Jabeen & Braun (2018) for tonal

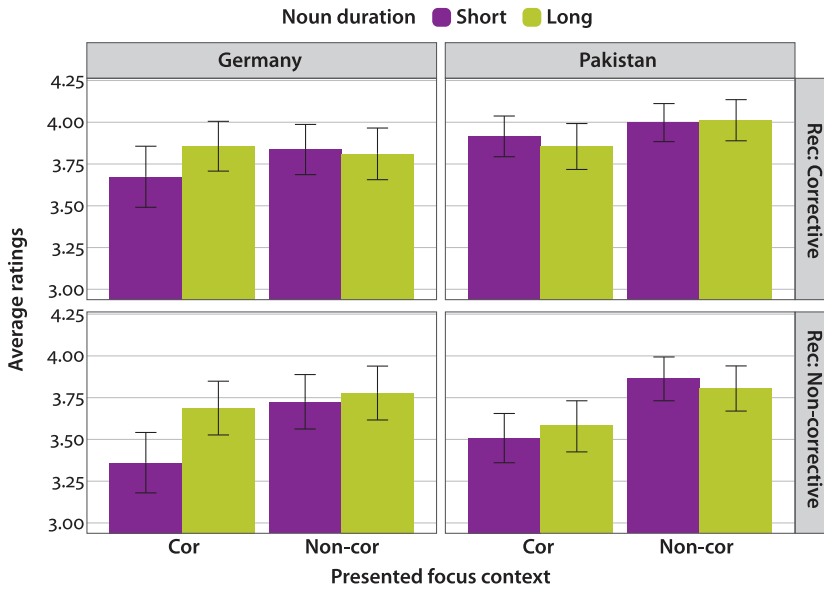


Figure 2. Ratings by Urdu speakers split for presented focus context (cor(rective) and non-cor(rective) on x-axis), duration condition, and recorded focus type. The whiskers represent the 95% Confidence Interval

differences between corrective and non-corrective focus), but not to the duration of the syllables in the focused noun phrase. The German residents, in contrast, downgraded the stimuli with short noun duration (and a long case marker) in the presented corrective context. This indicates that long noun duration (with a short case marker) is a necessary condition for marking corrective focus for German residents, while they did not pay attention to other prosodic cues.

One may wonder why the participants' sensitivity to prosody (duration for German residents, original focus type for Pakistani residents) is limited to the corrective focus and does not hold also for non-corrective focus contexts. However, such an asymmetric pattern is not unheard of (Schwartz, 2022). In fact, it ties in with other findings which show that listeners react negatively to reduced prosodic cues (e.g., lower ratings for short noun duration in corrective context for German residents) but do not downgrade prosodic cues that are too strong for a given condition (e.g., no difference in ratings for long and short noun duration in the non-corrective focus context for German residents). For instance, Braun (2004) has shown that in contrastive topic contexts, German listeners clearly preferred topics with long durations and high and late pitch peaks. In non-contrastive contexts, however, there was no effect of prosody. One possible explanation for this is that participants accommodate the context to create a fit with the stronger

prosodic cues. Alternatively, it is also possible that the corrective focus context is considered more marked (Zerbian, 2015) and the speakers pay more attention to the cues in that context compared to the less marked non-corrective focus context.

5. General discussion

Our study focused on the impact of stress-timed, head-prominence marking German as an L2 on the use of duration in syllable-timed, edge-prominence marking Urdu as an L1. Our production experiment shows that syllable duration is used to distinguish between different types of narrow focus by both groups of Urdu speakers: The first syllable of the noun was lengthened, and the case marker was shortened in corrective compared to non-corrective focus contexts. The duration of the second syllable of the noun showed weak evidence of cross-linguistic influence: While German residents slightly lengthened this syllable in corrective contexts, the Pakistani residents slightly shortened it. Statistically, this differential use of duration in the second syllable of the noun was significant, but generalization was limited due to the lacking three-way-interaction between focus context, country of residence, and syllable number. This led us to err on the conservative side and refute H1. In perception, the effect of country of residence on the use of duration as a cue to presented focus context was significant: Only the German residents employed duration as a cue to focus context, while Pakistani residents did not. This finding is in line with H2. This modality-specific influence (production vs. perception) of the majority language may be interpreted as a stronger (and/or earlier) influence in perception than in production. Pakistani residents were sensitive to the original recording setting, which suggests that they used other cues for the task. There is very little room for comparison with other studies, as only a few studies have tested both modalities (cf. Altenberg, 2005; Beddor et al., 2018; Cangemi et al., 2015; Nagle, 2018).

In addition to gaining insights into cross-linguistic influence from L2 to L1, our findings also provide new data on the prosodic marking of focus types in a lesser-studied language. Jabeen (2022) has already shown that Urdu speakers use longer duration in the production of narrow focus as compared with wide focus. Together with findings by Jabeen (2022), we can conclude that speakers of Urdu use duration in a gradient manner (corrective focus > non-corrective focus > wide focus) and they manipulate it to distinguish between types of narrow focus i.e., non-corrective and corrective focus.

The majority language in the multilingual context can affect the use of phonetic cues in contact languages (Beddor, 2012), thus leading to language attrition,

understood as re-weighting of cues in L1 under the influence of L2. Our data on the perception of duration in corrective and non-corrective focus contexts provides an indication of a re-weighting of the duration cue by Urdu participants tested in Germany. It is a matter for future research to investigate if Urdu speakers living in Germany use longer duration in the corrective focus context in their L2 German too.

One caveat is that the evidence regarding the production and perception differences comes from a between-subjects comparison. In future studies, one could test the same group of participants on both tasks (cf. Beddor et al., 2018), and increase the sample size of participants to be able to investigate more fine-grained differences on the basis of duration of stay (Elordieta & Romera, 2021; Mennen & Chousi, 2018; Mennen et al., 2022), percentage of usage of the majority language (Braun et al., 2014; Levy et al., 2019), as well as the attitude towards the majority language (Elordieta & Romera, 2021). Moreover, further research is planned to test how long the sensitivity to the durational cue lasts upon Urdu speakers' return to Pakistan (cf. Chamorro et al. (2016) for Spanish).

6. Conclusion

The current study documents change in the weighting of the duration cue for the perception of non-corrective and corrective focus. Our findings show that a cue learnt in an L2 can influence the weighting of this cue in the L1. However, at the group level, the influence of the L2 did not show generalizable effects in production. It is possible that the use of the duration cue is a process in the making for Urdu speakers living in Germany. Our results provide an insight into the extent of cross-linguistic influence of L2 on L1 and highlight the role of L1 attrition in the use of the duration cue in corrective focus marking.

Disclosure statement

The authors report no conflict of interest.

Consent












All the participants gave informed consent to participate in the experiments.

Acknowledgements

We thank Habiba for producing the stimuli for the perception experiment, Joachim Kleinmann for setting up the perception experiment, Christoph Schwarze, three anonymous reviewers, and the editor for valuable comments on the earlier versions of the manuscript.

References

- Altenberg, E. (2005). The judgment, perception, and production of consonant clusters in a second language. *International Review of Applied Linguistics in Language Teaching*, 43(1), 53–80.
- Atterer, M., & Ladd, R. (2004). On the phonetics and phonology of “segmental anchoring” of F0: Evidence from German. *Journal of Phonetics*, 32, 177–197.
- Baayen, H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Beddor, P.S. (2009). A coarticulatory path to sound change. *Language*, 85(4), 785–821.
- Beddor, P.S. (2012). Perception grammars and sound change. In M.J. Solé & D. Recasens (Eds.), *The initiation of sound change. Perception, production, and social factors* (pp. 37–55). John Benjamins.
- Beddor, P.S., Coetzee, A.W., Styler, W., McGowan, K.B., & Boland, J.E. (2018). The time course of individuals’ perception of coarticulatory information is linked to their production: Implications for sound change. *Language*, 94(4), 931–968.
- Boersma, P., & Chládková, K. (2011). Asymmetries between speech perception and production reveal phonological structure. In W.S. Lee & E. Zee (Eds.), *Proceedings of the ICPHS 2011* (pp. 328–331). Hong Kong.
- Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer* (v. 6.3) [computer program]. Available at <http://www.praat.org/> [retrieved 15.11.2022]
- Brauer, M., & Curtin, J.J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3), 389–411.
- Braun, B. (2004). Answers to the perception of thematic contrast and questions regarding the perception of thematic “non-contrast”. In B. Bel & I. Marlien (Eds.), *Proceedings of Speech Prosody 2004* (pp. 685–688), Nara, Japan.
- Braun, B., & Biezma, M. (2019). Prenuclear L*+H activates alternatives for the accented word. *Frontiers in Psychology*, 10.
- Braun, B., Dainora, A., & Ernestus, M. (2011). An unfamiliar intonation contour slows down on-line speech comprehension. *Language and Cognitive Processes*, 26(3), 350–375.
- Braun, B., Galts, T., & Kabak, B. (2014). Lexical encoding of L2 tones: The role of L1 stress, pitch accent and intonation. *Second Language Research*, 30(3), 323–350.
- Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7), 1044–1098.

- Broselow, E., & Kang, Y. (2013). Phonology and speech. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge handbook of second language acquisition* (pp. 529–554). Cambridge: Cambridge University Press.
- Butt, M., Jabeen, F., & Bögel, T. (2016). Verb cluster internal wh-phrases in Urdu: Prosody, syntax and semantics/pragmatics. *Linguistic Analysis*, 40(3–4), 445–487.
-  Butt, M., Jabeen, F., & Bögel, T. (2020). Ambiguity resolution via the syntax-prosody interface: The case of kya ‘what’ in Urdu/Hindi. In G. Kentner & J. Kremer (Eds.), *Prosody in syntactic encoding* (pp. 85–117). De Gruyter.
- Cangemi, F., Krüger, M., & Grice, M. (2015). Listener-specific perception of speaker-specific productions in intonation. In S. Fuchs, D. Pape, C. Petrone, & P. Perrier (Eds.), *Individual differences in speech production and perception* (pp. 123–145). Peter Lang.
-  Chamorro, G., Sorace, A., & Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Bilingualism: Language and Cognition*, 19(3), 520–532.
-  Chen, A., Çetinçelik, M., Roncaglia-Denissen, M. P., & Sadakata, M. (2021). Native language, L2 experience, and pitch processing in music. *Linguistic Approaches to Bilingualism*, 13(2), 218–237.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Harper and Row.
-  Delattre, P. (1969). An Acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics in Language Teaching*, 7(4), 295–326.
-  de Leeuw, E. (2019). Native speech plasticity in the German-English late bilingual Stefanie Graf: A longitudinal study over four decades. *Journal of Phonetics*, 73, 24–39.
-  de Leeuw, E., Mennen, I., & Kirby, J. M. (2012). Singing a different tune in your native language: First language attrition of prosody. *International Journal of Bilingualism*, 16, 101–116.
- Dhillon, R. K. (2010). Stress and tone in Indo-Aryan languages (Publication No. 3415170) [Doctoral dissertation, Yale University]. ProQuest.
-  Dmitrieva, O. (2019). Transferring perceptual cue-weighting from second language into first language: Cues to voicing in Russian speakers of English. *Journal of Phonetics*, 73, 128–143.
- Dua, H. (2009). Urdu. In K. Brown & S. Ogilvie (Eds.), *Concise encyclopaedia of languages of the world* (pp. 1133–1138). Elsevier.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2020). *Ethnologue: Languages of the world*. Dallas, Texas: SIL International. (Twenty-third edition, Online version: <http://www.ethnologue.com> (accessed on 15.09.2020))
-  Einfeldt, M., Sevastjanova, R., Zahner-Ritter, K., Kazak, E., & Braun, B. (2023). The use of active learning systems for stimulus selection and response modelling perception experiments. *Computer Speech & Language*, 83, 101537.
-  Elordieta, G., & Romera, M. (2021). The influence of social factors on the prosody of Spanish in contact with Basque. *International Journal of Bilingualism*, 25(1), 286–317.
-  Féry, C., Pandey, P., & Kentner, G. (2016). The prosody of focus and givenness in Hindi and Indian English. *Studies in Language*, 40(2), 302–339.
-  Flege, J. E. (1987). The production of ‘new’ and ‘similar’ phones in a foreign language: Evidence for the effect of Equivalence Classification. *Journal of Phonetics*, 15, 47–65.

- doi** Flege, J.E., & Bohn, O.-S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press.
- Genzel, S., & Kügler, F. (2010). The prosodic expression of contrast in Hindi. In M. Hasegawa–Johnson, A. Bradlow, J. Cole, K. Livescu, J. Pierrehumbert, & C. Shih (Eds.), *Proceedings of Speech Prosody 2010*, 143:1–4, Chicago, IL, USA. Retrieved from https://www.isca-speech.org/archive_vo/sp2010/sp10_143.html
- doi** Green, P., & MacLeod, C.J. (2016). SIMR: an R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498.
- doi** Harrington, J., Kleber, F., Reubold, U., Schiel, F., & Stevens, M. (2018). Linking cognitive and social aspects of sound change using agent-based modeling. *Topics in Cognitive Science*, 10, 707–728.
- doi** Henriksen, N. (2016). Convergence effects in Spanish–English bilingual rhythm. In J. Barnes, N. Veilleux, S. Shattuck–Hufnagel, & A. Brugos (Eds.), *Proceedings of Speech Prosody 2016* (pp. 721–725). Boston, USA.
- Hussain, S. (1997). Phonetic correlates of stress in Urdu (Publication No. 9814230) [Doctoral dissertation, Northwestern University, USA]. ProQuest.
- Jabeen, F. (2019). Prosody and word order: Prominence marking in declaratives and wh-questions in Urdu/Hindi [Doctoral dissertation, University of Konstanz, Germany]. KOPS. <http://nbn-resolving.de/urn:nbn:de:bsz:352-2-19mbzsf16gr5x7>
- doi** Jabeen, F. (2022). Word order, intonation, and prosodic phrasing: Individual differences in the production and identification of narrow and wide focus in Urdu. *Languages*, 7(2), 103.
- doi** Jabeen, F., & Braun, B. (2018). Production and perception of prosodic cues in narrow and corrective focus in Urdu/Hindi. In K. Klessa, J. Bachan, A. Wagner, M. Karpiński, & D. Śledziński (Eds.), *Proceedings of Speech Prosody 2018* (pp. 30–34). Poznań, Poland.
- doi** Jabeen, F., & Delais–Roussarie, E. (2020). The Accentual Phrase in Urdu/Hindi: A prosodic unit at the interplay between rhythm and intonation. In N. Minematsu, M. Kondo, T. Arai, & R. Hayashi (Eds.), *Proceedings of Speech Prosody 2020* (pp. 542–546). Tokyo, Japan.
- Jun, S.A. (2012). Prosodic typology. In S.A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430–453). Oxford University Press.
- Kidwai, A. (2000). *XP–adjunction in Universal Grammar: Scrambling and binding in Hindi–Urdu*. Oxford University Press.
- doi** Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3–4), 243–276.
- Kügler, F. (2008). The role of duration as a phonetic correlate of focus. In P. Barbosa, S. Madureira, & C. Reis (Eds.), *Proceedings of Speech Prosody 2008* (pp. 591–594). Campinas, Brazil.
- doi** Kügler, F., & Gollrad, A. (2015). Production and perception of contrast: The case of the rise–fall contour in German. *Frontiers in Psychology*, 6.
- doi** Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Ladd, D.R. (1980). *The structure of intonational meaning: Evidence from English*. Indiana University Press.
- doi** Lambrecht, K. (1994). *Information structure and sentence form*. Cambridge University Press.

- doi Levy, H., Konieczny, L., & Hanulíková, A. (2019). Processing of unfamiliar accents in monolingual and bilingual children: Effects of type and amount of accent experience. *Journal of Child Language*, 46(2), 368–392.
- doi Llompart, M., & Reinisch, E. (2019). Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech*, 62(3), 594–622.
- doi Maastricht, L., Krahmer, E., & Swerts, M. (2016). Prominence patterns in a second language: Intonational transfer from Dutch to Spanish and vice versa. *Language Learning*, 66(1), 124–158.
- doi Maastricht, L., Krahmer, E., Swerts, M., & Prieto, P. (2019). Learning direction matters: A study on L2 rhythm acquisition by Dutch learners of Spanish and Spanish learners of Dutch. *Studies in Second Language Acquisition*, 41, 87–121.
- doi Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, 76, 190–208.
- doi Manetta, E. (2012). Reconsidering rightward scrambling: Postverbal constituents in Hindi–Urdu. *Linguistic Inquiry*, 43(1), 43–74.
- doi Matras, Y. (2013). Languages in contact in a world marked by change and mobility. *Revue Française de Linguistique Appliquée*, 18(2), 7–13.
- doi Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, 32, 543–563.
- doi Mennen, I., & Chousi, D. (2018). Prosody in first-generation adult immigrants and second-generation heritage-language users: The timing of prenuclear rising accents. In K. Klessa, J. Bachan, A. Wagner, M. Karpiński, & D. Śledziński (Eds.), *Proceedings of Speech Prosody 2018* (pp. 828–832). Poznań, Poland.
- Mennen, I., Reubold, U., Endes, K., & Mayr, R. (2022, April 20–22). *L2-induced changes in native intonation: An investigation of four dimensions of intonation in English migrants to Austria*. [Conference presentation]. 10th International Symposium on the Acquisition of Second Language Speech (New Sounds 2022), Barcelona, Spain.
- doi Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, 9, 453–467.
- doi Nagle, C. (2018). Examining the temporal structure of the perception–production link in second language acquisition: A longitudinal study. *Language Learning*, 68(1), 234–270.
- Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., & Vasishth, S. (2008). Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics*, 1(1), 55–72.
- doi Piccinini, P. & Arvaniti, A. (2021). Dominance, mode, and individual variation in bilingual speech production and perception. *Linguistic Approaches to Bilingualism*, 9(4–5), 628–658.
- doi Puri, V. (2018). Focus in Indian English and Hindi late and simultaneous bilinguals. *Linguistic Approaches to Bilingualism*, 8(3), 343–371.
- R Core Team. (2023). *R: A language and environment for statistical computing* (v. 4.3.1) [Computer software manual]. Vienna, Austria.
- doi Rahman, T. (2019). Mother tongue education policy in Pakistan. In A. Kirkpatrick & A. J. Liddicoat (Eds.), *The Routledge international handbook of language education policy in Asia* (pp. 364–381). Taylor and Francis Group.

- doi** Ribbert, A., & Kuiken, F. (2010). L2-induced changes in the L1 of Germans living in the Netherlands. *Bilingualism: Language and Cognition*, 13(1), 41–48.
- Saleem-ud-Din, S. M., & Anjum, S. (Eds.). (2013). *Oxford Urdu–English dictionary*. Oxford University Press.
- doi** Schertz, J., Carbonell, K., & Lotto, A. J. (2020). Language specificity in phonetic cue weighting: Monolingual and bilingual perception of the stop voicing contrast in English and Spanish. *Phonetica*, 77(3), 186–208.
- doi** Schmid, M. (2011). *Language Attrition*. Cambridge: Cambridge University Press.
- doi** Schwartz, G. (2022). Asymmetrical cross-language phonetic interaction: Phonological implications. *Linguistic Approaches to Bilingualism*, 12(2), 103–132.
- doi** Tremblay, A., Kim, S., Shin, S., & Cho, T. (2021). Re-examining the effect of phonological similarity between the native- and second-language intonational systems in second-language speech segmentation. *Bilingualism: Language and Cognition*, 24(2), 401–413.
- Trouvain, J., & Braun, B. (2020). Sentence prosody in a second language. In C. Gussenhoven & A. Chen (Eds.), *The Oxford handbook of language prosody* (pp. 605–618). Oxford University Press.
- doi** Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schließer (Eds.), *Methods in empirical prosody research* (pp. 1–28). De Gruyter.
- doi** Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, 34, 459–519.
- Vijaykrishnan, K. G. (2003). Stress and tone in Punjabi. *Language in India*, 3(11), 32.
- doi** Zahner-Ritter, K., Einfeldt, M., Wochner, D., James, A., Dehé, N., & Braun, B. (2022). Three kinds of rising–falling contours in German wh–questions: Evidence from form and function. *Frontiers in Communication*, 7.
- doi** Zerbian, S. (2015). Markedness considerations in L2 prosodic focus and givenness marking. In E. Delais–Roussarie, M. Avanzi, & S. Herment (Eds.), *Prosody and language in contact. Prosody, phonology and phonetics* (pp. 7–27). Springer.

Data availability statement

The data for this paper, including appendices, is available at the following link in the OSF repository: https://osf.io/knfcd/?view_only=85065bd5eocd4beba24foofd10719ffd

Address for correspondence

Farhat Jabeen
CITEC, Inspiration 1
Bielefeld University
33615 Bielefeld
Germany

farhat.jabeen@uni-bielefeld.de



<https://orcid.org/0000-0002-4681-1437>

Co-author information

Bettina Braun
University of Konstanz

bettina.braun@uni-konstanz.de



<https://orcid.org/0000-0002-7696-3255>

Publication history

Date received: 22 December 2022

Date accepted: 20 August 2023

Published online: 20 October 2023