Shafi, S., & Hellmuth, S. (Accepted/In press). Stress placement in English loanwords by speakers of Mirpur Pahari in the UK. In R. Rao (Ed.), *The phonetics and phonology of heritage languages.* Cambridge University Press.

Stress placement in English loanwords by speakers of Mirpur Pahari in the UK

We use stress placement in English loanwords into Mirpur Pahari (MP) as a test case to explore a usage-based account of loanword adaptation which incorporates internal factors (such as the phonology of the recipient language, here, MP), alongside external factors (such as amount of exposure to the donor language, here, English). Stress placement in English loanwords into MP as spoken in Pakistan versus UK show variable adaptation, from conformity to the stress rules of MP phonology, to marking stress on the same syllable of the word as in English, which is a challenge for most theories of loanword adaptation. In our approach, variable adaptation in an individual speaker's productions of English loanwords into MP are correlated with the degree of exposure of each person to English, in production data from 12 speakers of MP in the UK.

Keywords: loanword phonology; Mirpur Pahari; UK English; usage-based phonology

1. Introduction

An important issue in the literature on phonological loanword adaptation is whether or not phonological generalizations holding of loanwords can be accounted for within the phonology of the borrowing language or not (Silverman, 1992). This paper goes straight to the heart of this issue, presenting patterns of loanword adaptation which vary across speakers in exactly this regard. In our data, we will show that some speakers' patterns reflect the constraints observed in the phonology of the borrowing language, but other speakers' patterns do not; indeed, for the latter group, we find patterns of variable adaptation whereby some lexical items reflect the constraints in the phonology of the borrowing language, but other lexical items match the position of stress in the word in the source language. This paper seeks to offers a unified analysis of the internal and external factors governing

these patterns of variation, between and within speakers, in adaptation of the position of stress in English loanwords into Mirpur Pahari.

Prior work on stress patterns in loanwords, involving a range of languages, commonly shows competing effects of faithfulness to the position of stress in the word as pronounced in the source language and effects of the stress grammar of the borrowing language (Alber, 1998; Davidson & Noyer, 1997; Kager, 2000). The consistently observed role of faithfulness to the surface source word form would seem necessarily to require a loanword specific grammar, since by definition a rule or constraint requiring a match to a word in another language cannot be part of the autonomous phonology of the borrowing language (see Kang, 2010 for an overview). A number of accounts have sought to bring stress adaptation patterns under the umbrella of the phonology of the borrowing language by appeal to the role of L1 perception in the loanword adaptation process (Broselow, 2009; Hamann & Boersma, 2009). In this paper we contribute further evidence of this general pattern of stress adaptation, from novel data in an under-researched language (Mirpur Pahari) as spoken by heritage speakers in the UK. Our data reveals patterns of variable adaptation in stress position, between and within speakers, which resists any attempt to account for the patterns within a single grammar. We sketch an account in terms of externally driven change to the phonology of the borrowing language by appeal to usage-based phonology, building on ideas set out in Message Oriented Phonology (Hall, Hume, Jaeger, & Wedel, 2016), and test this hypothesis in a small-scale production study.

In section 2 we provide background information about the Mirpur Pahari language, before presenting key generalizations regarding the position of stress in English loanwords into MP as spoken in Pakistan, together with a formal analysis in Optimality Theory which demonstrates that a variable grammar is needed to account for the data. We then outline how a usage-based approach can offer a principled way of predicting patterns of variable adaptation. In section 3-4 we present the methods and

results of a small-scale empirical study designed to test these predictions, before closing the paper with a short conclusion in section 5.

2. Background to the Study

2.1 Mirpur Pahari

Mirpur Pahari (MP), also known as Mirpuri, is a non-tonal member of the Western Punjabi language family, spoken by approx. 4 million in Pakistan and 500K in the UK, and has no written form (Stow, Pert, & Khattab, 2012). There appear to be several related Pahari dialects, but limited research on the phonology of any of them, and Pahari is under-represented in research on UK minority languages (Hussain, 2015). Khan (2012) worked on English loanwords into Poonch Pahari and Tabbassam (2003) provides a basic but limited description of some aspects of Mirpur Pahari (MP). The present study of English loanwords into MP is part of a wider project to document the phonetics and phonology of MP.

A large number of regional languages are spoken in Pakistan, including Mirpur Pahari, alongside Urdu as the national/standard language. English is also widely spoken as an additional language and is used in most official correspondence. In Pakistan, then, both Urdu and English carry prestige, and are used by educated speakers in preference to their regional language (here, MP) in many contexts, and especially in urban centres. In villages, MP speakers are more likely to retain use of MP in all contexts, but there is regular contact with speakers from Mirpur and/or the UK. As a result, the MP lexicon includes many loanwords from both Urdu and English. Some of the loanwords in our Pakistan corpus (described in section 2) may have been borrowed into MP from English via Urdu, and in a few cases the indirect route is clear from e.g. the influence of Urdu orthography. Such cases are excluded from the present analysis, and we focus on tokens which are either clearly borrowed directly from English, or which show no signs of borrowing via Urdu.

A large number of Mirpuris settled in the UK from the 1960s onwards, and there is ongoing primary immigration to the UK e.g. due to marriage. British Asian speakers of Mirpuri heritage who live in cities with a large Mirpuri population (such as Bradford and Leeds) will tend to speak their heritage language, MP, at home. It is common for Mirpuris to make regular visits to Pakistan, and a constant flow of new English words thus reaches even older speakers still residing in rural areas in Mirpur. The language contact situation is thus complex, and the variable adaptation patterns which result are the primary topic of interest in the present paper.

MP has a relatively large consonant inventory, maintaining a four way laryngeal contrast in plosives: voiced, voiceless, voiceless aspirated, voiced aspirated (Tabbassam, 2003). Our Pakistan corpus MP data are consistent with the vowel inventory proposed by Khan (2012) for Poonch Pahari, with six pairs of short and long vowels: [ɪ iː e eː æ æː ə aː o oː ʊ uː]. In our corpus data, we find the following syllable types: CV, CVV, CVC, CVCC, CVVC. Complex onsets are not permitted, and the only complex codas observed are nasal-obstruent sequences.

The eastern Punjabi language has word-level tone marking lexical contrast, as well as word-level stress (Bhatia, 1993). The presence of tone is correlated with the loss of the voicing contrast in aspirated plosives. Our investigations in MP to date suggest that there are no tonal contrasts in MP (though further investigation is needed), but there is clear evidence of word-level stress (Shafi, 2019).

Primary stress is assigned in MP according to the algorithm in (1), as illustrated in (2).

(1) a) stress a final superheavy syllable,

b) else the penultimate syllable (‘penult’)

(2)	a)	pəˈsən̪d̪	choice	dərˈbaːr	shrine
	b)	ˈd̪ər.zən	seamstress	ˈtʃaː.vəl	rice

ˈχʰəs.ra	measles	ˈsoː.ti	cane
bə.ˈraɖ.rɪ	caste	tʃəpˈɽaː.si	assistant

We have not found any evidence in our data of secondary stress in MP. As a result we assume that only a single metrical foot is built per word. Stress never falls on an open/light CV syllable and the maximal monomorphemic word size appears to be trisyllabic; in our data all words of four syllables or more are compounds. In Metrical Stress Theory (Hayes, 1995), the MP stress system can be analyzed by proposing that bimoraic trochees are built right to left, with main stress applied to the rightmost foot, and that consonant extrametricality applies, as well as a complete ban on degenerate feet. We offer a formalization of the MP stress system in Optimality Theory (Prince & Smolensky, 2004) in section 2.3.

2.2 Stress Placement in English loanwords in Mirpur Pahari

A text corpus of 1200 loanwords was created by the first author, who is a first language speaker of MP, based on elicitation of grammaticality judgements from family members in Pakistan of different generations. In this text corpus of MP as spoken in Pakistan, we find four adaptation patterns, which differ in whether or not stress falls on the same syllable in the loanword as in the English realisation of the source word, as shown in (3).[1]

(3)

	gloss	*English*	*MP loanword*
Pattern A	refuse	rɪ.ˈfjuːz	rəf.ˈjuːz
	public	ˈpʌb.lɪk	ˈpəb.lək
	inspector	ɪn.ˈspɛk.tə	əns.ˈpæk.tər
Pattern B	appendix	ə.ˈpɛn.dɪks	ˈpæn.dəs

[1] [1] We do not consider segmental adaptation patterns here, such as deletion, substitution or adaptation of individual phonemes; we do discuss adaptations involving changes to syllable structure where relevant to the analysis of stress (see Shafi, 2019 for an analysis of variable adaptation of phonotactic patterns in MP loanwords).

	decision	dɪ.ˈsɪ.ʒən	də.ˈsiː.ʒən
	lettuce	ˈlɛ.tɪs	ˈlæː.təs
Pattern C	stadium	ˈsteɪ.dɪəm	sə.ˈteː.dɪəm
	ambulance	ˈæm.bju.ləns	ˈæm.bu.læns
	vaccine	ˈvæk.siːn	ˈvæk.siːn
Pattern D	shampoo	ʃæm.ˈpuː	ˈʃæm.pu
	impact (noun)	ˈɪm.pækt	əm.ˈpækt
	glucose	ˈgluː.kəʊz	gəl.ˈkoːz

In pattern A the stress in the English word falls in a position that happens already to meet the MP stress rules (as set out in section 2.1), that is, on a final superheavy syllable (as in ‘refuse’ [rɪ.ˈfjuːz]) or on a heavy penult syllable (as in ‘public’ [ˈpʌb.lɪk]$_{Eng}$ → [ˈpəb.lək]$_{MPL}$).

In pattern B, stress can be maintained on the syllable which is stressed in English, but only at the cost of some kind of structural change. For example, in ‘decision’ the vowel of the stressed penult syllable is lengthened [dɪ.ˈsɪ.ʒən]$_{Eng}$ →[də.ˈsiː.ʒən]$_{MPL}$) resulting in a suitable heavy penult syllable to bear stress. Similarly, in ‘appendix’, the word-final consonant cluster [ks] is simplified, avoiding a final superheavy syllable attracting stress away from the penult which bears stress in the English source word ([ə.ˈpɛn.dɪks]$_{Eng}$ → [ˈpæn.dəs]$_{MPL}$). The result is a word which keeps stress on the syllable that it was on in English, while still meeting the MP stress rules.

In patterns C and D, the differing demands of keeping stress on the syllable that it was on in English versus meeting the MP stress rules, are not both met. In pattern C, stress is retained in the position that matches the source word, at the expense of meeting the MP stress rules; for example, in ‘ambulance’ the final superheavy syllable is retained but does not attract stress: [ˈæm.bju.ləns]$_{Eng}$ → [ˈæm.bu.læns]$_{MPL}$. In contrast, in pattern D, stress is placed according to the MP stress rules, which

results in stress falling on a different syllable to that which is stressed in the English source word; in 'shampoo' stress falls on the heavy penult syllable rather the (non superheavy) final syllable ([ʃæm.ˈpuː]$_{Eng}$ → [ˈʃæm.pu]$_{MPL}$), and in the noun 'impact' stress is attracted to a final superheavy syllable, rather than falling on the penult as it does in English ([ˈɪm.pækt]$_{Eng}$ → [əm.ˈpækt]$_{MPL}$).

All four patterns are found in the text corpus portion which reflects the production of younger speakers of MP in Pakistan, who are educated to high school level or higher and who thus learned English as a foreign language at school. In contrast, pattern C is *not* found in the text corpus data which reflects the production of loanwords by from older MP speakers who generally have a lower level of education, and who have not learned English. Instead, for example, these older speakers are expected to produce the pattern C words as shown in (4).

(4)

gloss	*English*	*MP loanword young speakers*	*MP loanword older speakers*
stadium	ˈsteɪ.dɪəm	sə.ˈteː.dɪəm	əs.tə.ˈdɪəm
ambulance	ˈæm.bju.ləns	ˈæm.bu.læns	æm.ˈboː.lə̃s
vaccine	ˈvæk.siːn	ˈvæk.siːn	vək.ˈsiːn

We see variation on a 'macro' level, therefore, between younger and older speakers, in that younger speakers' judgements permit use of 'pattern C', but older speakers' judgements do not. Within the younger speakers' patterns, we see variable adaptation of *structurally parallel words*: some are adapted so as to retain stress on the same syllable which is stressed in English (pattern C) but others are adapted so as to reflect the stress rules of MP. An example is the structural minimal pair in (3),

'vaccine'~'glucose', both of which involve a heavy penult followed by a superheavy syllable [CVC.CVVC][2], but which are adapted by younger speakers in two different ways.

In the next section we offer a formalization of MP stress in Optimality Theory (OT), to highlight the variable grammar(s) which result in the loanword adaptation patterns in (3) and (4).

2.3 Modelling variable adaptation

2.3.1 A Formal Account within Optimality Theory

In Optimality Theory surface forms are modelled as arising from the interaction of ranked constraints (McCarthy, 2008). Markedness constraints refer only to the properties of candidate surface forms, but faithfulness constraints refer also to a putative input form. Differences between grammars (and thus languages) arise in the theory from differences in constraint rankings alone, regardless of the properties of the putative input form (Prince & Smolensky, 2004). In our analysis of loanwords from English into MP, we illustrate the analysis using the surface English form as the input form, but in principle the grammar should succeed in predicting the observed surface form(s) whatever the input form.

The stress patterns observed in Mirpur Pahari (MP), as shown in (2) above, can be analysed in an OT grammar using the constraints listed in (5).

(5) a) FT-BIN: Feet are binary (under moraic analysis).

b) PARSE-σ: Syllables are parsed into feet.

c) ALLFTR (Ft, R, PWd, R): Every foot is at the right edge of the Prosodic Word.

d) STRESS-TO-WEIGHT ('SWP'): If stressed, then heavy (Crosswhite, 1998).

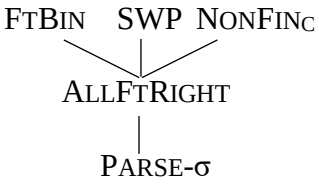
e) NON-FINALITY[C,w]: No mora-level gridmark occurs over the final consonant of a prosodic word (Hyde, 2011).

[2] As noted in section 2.1, complex onsets such as *[gl] are not permitted in MP, thus the initial syllable adjusts from [glu] to [gəl] (see Shafi, 2019 for further details).

The first three of these – FT-BIN, PARSE-σ, and ALLFTR – are standard constraints known to be indicated in the cross-linguistic typology of stress (Kager, 2007). We see the effects of the Stress-to-Weight principle ('SWP') in that all stressed syllables in MP are heavy, without exception. Finally, we use Hyde's (2011) grid-based definition of Non-Finality to capture the robustly observed pattern of consonant extrametricality in MP.[3]

We will motivate the constraint ranking in (6) through a series of worked examples in (7)-(9). Foot structure is marked with parentheses '( )', and extrametricality with angled brackets '< >'.

(6)



(7)

	/tʃaːvəl/ 'rice'	FTBIN	SWP	NONFINC	ALLFTR	PARSE-σ
☞	a. (ˈtʃaː).vəl				*	*
	b. tʃaː.(ˈvəl)			*!		*
	c. tʃaː.(ˈvə)<l>	*!	*!			*
	d. (tʃaː).(ˈvəl)			*!	*	
	e. (tʃaː).(ˈvə)<l>	*!	*!		*	

[3] Since stress is only ever realized on heavy syllables in MP, it is possible that, in order to satisfy SWP, an input short vowel may be lengthened in violation of a faithfulness constraint $\text{IDENT}_{[\text{long-V}]}$. In the interests of brevity we set this issue aside here, but an analysis along these lines is explored in detail in Shafi (2019).

(8)

	/dərbaːr/ ‘shrine’	FtBin	SWP	NonFin$_C$	AllFtR	Parse-σ
☞	a. dər.(ˈbaː)<r>					*
	b. (dər).(ˈbaː)<r>				*!	

In (7), the observed surface form is candidate a., which violates AllFtR and Parse-σ. Candidates b. and c. demonstrate the ranking of FtBin, SWP and NonFin$_C$ over AllFtR; candidates d. and e. demonstrate the ranking of FtBin, SWP and NonFin$_C$ over Parse-σ. The ungrammaticality of candidates d. and e. indicates that only one foot is formed in MP. The example in (8) shows the ranking of AllFtR over Parse-σ.

The three adaptation patterns in the loanword corpus which reflect the productions of older speakers – that is, patterns A, B and D – can be analysed within the MP grammar, using the constraints and ranking in (6). This is demonstrated by a worked example of a pattern B case, in (9).

(9)

	[ˈlɛ.tɪs]$_{Eng}$ ‘lettuce’	FtBin	SWP	NonFin$_C$	AllFtR	Parse-σ
☞	a. (ˈlæː).təs				*	*
	b. (ˈlɛ).tɪs	*!	*!		*	*
	c. læː.(ˈtə)<s>	*!	*!			*
	d. læː.(ˈtəs)			*!		*
	e. (ˈlɛ.tɪ)<s>		*!			
	f. (ˈlɛ.tɪs)	*!	*!	*!		
	g. (lɛ).(ˈtɪs)	*!		*!	*	
	h. (ˈlɛ).(tɪs)	*!	*!	*!	*	

In (9), the surface form used by older speakers is candidate a. Candidates c-f. demonstrate ranking of FTBIN, SWP and NONFIN$_C$ over ALLFTR in loanwords, as in home-grown MP words; similarly, candidates e-h. demonstrate ranking of FTBIN, SWP and NONFIN$_C$ over PARSE-σ.

A further worked example in (10) confirms that pattern D cases can also be analysed within the autonomous MP grammar (setting aside segmental changes to resolve the complex onset).

(10)

	[ˈgluː.kəʊz]$_{Eng}$ ‘glucose’	FTBIN	SWP	NONFIN$_C$	ALLFTR	PARSE-σ
☞	a. gəl.(ˈkoː)<z>					*!
	b. (ˈgəl).koːz				*!	*!
	c. gəl.(ˈkoːz)			*!		*
	d. (gəl).(ˈkoː)<z>				*!	

Pattern C is observed in tokens in the text corpus portion which reflects the productions of younger speakers and is modelled by introducing a loan-phonology specific constraint, MATCHSTRESS (Davidson & Noyer, 1997), shown in (11).

(11) MATCHSTRESS: Stress falls on the same vowel in the source word as in the loan word.

(12)

	[ˈvæk.siːn]$_{Eng}$ ‘vaccine’	MATCH	FTBIN	SWP	NONFIN$_C$	ALLFTR	PARSE-σ
☞	a. (ˈvæk).siːn					*	*
	b. væk.(ˈsiː)<n>	*!					*

(13)

	[ˈgluː.kəʊz]$_{Eng}$ ‘glucose’	FTBIN	SWP	NONFIN$_C$	ALLFTR	PARSE-σ	MATCH
☞	a. gəl.(ˈkoː)<z>					*!	*
	b. (ˈgəl).koːz				*!	*!	

The Pattern C example in (12) demonstrates that, in this adaptation pattern, MATCHSTRESS is highly ranked. It is not possible to derive patterns C and D within the same grammar, however, because, for younger speakers, structurally parallel cases yield different outputs: [ˈvæk.siːn] ~ [gəl.ˈkoːz]. Pattern C requires a grammar in which MATCHSTRESS outranks ALLFTR, as shown in (12), but pattern D requires a grammar in which ALLFTR outranks MATCHSTRESS, as shown in (13).

Patterns A and B both involve stress falling on the vowel that would be stressed in the English source word, so the relative ranking of MATCHSTRESS will not impact on the analysis. Patterns A and B are thus also compatible with a grammar containing the loan-specific constraint MATCHSTRESS., as shown in (14). This is the same pattern B case shown in (9) for older speakers, without MATCHSTRESS.

(14)

	[ˈlɛ.tɪs]$_{Eng}$ ‘lettuce’	MATCH	FTBIN	SWP	NONFIN$_C$	ALLFTR	PARSE-σ
☞	a. (ˈlæː).təs					*	*
	b. (ˈlɛ).tɪs		*!	*!		*	*
	c. læː.(ˈtə)<s>	*!	*!	*!			*
	d. læː.(ˈtəs)	*!			*!		*
	e. (ˈlɛ.tɪ)<s>			*!			
	f. (ˈlɛ.tɪs)		*!	*!	*!		
	g. (lɛ).(ˈtɪs)	*!	*!		*!		
	h. (ˈlɛ).(tɪs)		*!	*!	*!	*	

We can therefore model the macro-variation between younger and older speakers as due to a change in the grammar, brought about by introduction of the loan-specific MATCHSTRESS constraint. However, it is not possible to model the observed intra-speaker variable adaptation between patterns C and D in the text corpus data reflecting younger MP speakers' productions, within a single grammar; external factors must be involved. Given the precise definition of MATCHSTRESS adopted here, a potential unifying feature in the younger speakers' pattern D cases might be frequency. Specifically, if speakers have not heard the source word in English they have nothing to 'match' to and MATCHSTRESS is vacuously satisfied, so only see the effects of the autonomous MP rules in these words.

The role of external factors could, in principle, be formalised within a model such as Stochastic OT (Boersma & Hayes, 2001), which assigns variable weightings to constraints so as to generate predictions about which grammar will obtain in a particular case. Rather than pursue this analysis further here, instead, in the next section, we explore how a usage-based theory might accommodate degree of exposure to English as part of a variable grammar, based on the insights of Message Oriented Phonology (Hall et al., 2016). The predictions of this model are then tested in a small-scale empirical study in sections 3-4.

2.4 The Present Study

As shown above, it is not possible to explain the variable adaptation of English loanwords into MP both within and between speakers, without accepting a change in the grammar which is triggered by one or more external factors. Specifically, in an OT account of the type illustrated here, some external factor is needed to trigger re-ranking of MATCHSTRESS in the grammar of younger speakers.

In this section we sketch an alternative approach, with a view to generating testable predictions. We propose modelling the variable adaptation in terms of degree of 'mere exposure' – in this case – to

English. In Message Oriented Phonology (MOP), the contents of each speaker's lexicon form the primary data over which statistical learning applies, thus the lexicon contributes directly to the emergence of phonological categories (Hall et al., 2016). We apply this model to the present language contact situation in which lexical items are borrowed from one language into another. We assume a shared lexical space for each speaker, between source and target languages: as the proportion of source language lexical items in the lexicon grows (here, English words), so the influence of the stress patterns in those words on the stress grammar increases.

This approach predicts inter-speaker variation in loanword adaptation as the norm, since individual speakers do not share an identical lexicon, but instead may add new items to their lexicon in a different order or at a different speed. The model also predicts intra-speaker variation since the changes to the grammar take place gradually over time, as the balance of word types in the lexicon changes. This situation, with both inter- and intra- speaker variation, is exactly the pattern we see in the MP loanword corpus, with overall macro-variation between younger and older speakers, which can be modelled in the form of different grammars, and also variable adaptation patterns within the productions of younger speakers, requiring a variable grammar.

The specific prediction of a usage-based approach in this context is that the amount and type of variable adaptation observed in an individual speaker's productions of English loanwords into MP will be correlated with the degree of exposure of each person to English. In the remainder of the chapter we test this hypothesis using English vocabulary size as a proxy for degree of exposure to English.

3. Methodology

3.1 Materials

A picture description task was designed in order to elicit 52 target loanwords, taken from the corpus described in section 2, and selected so as to elicit tokens of all potential adaptation types, but also

choosing words more amenable to visualisation. Picture description was used because there is no agreed written form for MP; although it is possible to represent MP in Urdu script we expected that some or all of our participants would not be confident reading MP written in this way. The full list of target loanwords is provided in the Appendix. Each word was elicited in isolation (picture naming) and in context (picture description). We also used picture description to elicit 19 target words in MP, to establish whether the stress patterns used in MP by the Bradford speakers matched the patterns described above for speakers in Pakistan.

To obtain a measure of vocabulary size across participants, we used the Vocabulary Size Test (VST, Nation & Beglar, 2007). This is a measure of written receptive vocabulary size in English and has undergone validation (Beglar, 2010; Gyllstad, Vilkaite, & Schmitt, 2015). For practical reasons, we wanted a test that could be administered on paper rather than online, since we would not always have internet access in testing sites; this ruled out use of an alternative checklist-based vocabulary size measure, the V_YesNo test (Meara & Miralpeix, 2016). The VST allows us to use the same test for participants across a wide range of levels of proficiency in English, including early bilinguals, in contrast to e.g. the earlier Vocabulary Levels Test (Laufer & Nation, 1999). We used the original 140 word version of the VST, which tests the first 14000 words of English, at a 1 in 100 sampling rate. We favoured this over the 20000 words versions of the test, which, although shorter, uses a lower sampling rate (1 in 200), since it has been found that, even at a 1 in 100 sampling rate the VST test slightly overestimates actual vocabulary size (Gyllstad et al., 2015). The VST is nevertheless easy to administer, and to interpret, and is a well-established test.

3.2 Participants

First language speakers of MP living in Bradford were recruited as participants, through personal contacts of the first author, and with the help of the community research team of the *Born in Bradford*

study (https://borninbradford.nhs.uk/). Criteria for inclusion were being a first language speaker of MP resident in Bradford, aged 18 or over. A language background questionnaire was used to build up a picture of the language use of each participant, but information provided on that questionnaire, such as knowledge of other languages, was not used as criteria for exclusion. We report here on the responses of 12 participants (f1-f6 female/m1-m6 male); details of the language background of each, including age at onset of acquisition of English (AOA) and length of residence (LOR) in the UK, are in Table 1.

Table 1. Demographic and language profile of the Bradford participants.

	age	*place of birth*	*AOA*	*LOR*	*highest education*	*language use at work*
f1	25	Mirpur	22	3	high school in Pakistan	MP, English
f2	47	Mirpur	6	41	high school in UK	English
f3	45	Mirpur	21	24	primary school in Pakistan	English, Urdu
f4	24	Bradford	n/a	24	undergraduate degree in UK	MP, English
f5	38	Mirpur	26	12	high school in Pakistan	MP, English
f6	36	Mirpur	4	32	high school in UK	MP, English
m1	34	Mirpur	25	9	undergraduate degree in Pakistan	English
m2	44	Bradford[4]	5	39	high school in UK	MP, English
m3	38	Bradford	n/a	38	high school in UK	MP, English
m4	31	Bradford	n/a	31	high school in UK	MP, English
m5	45	Mirpur	28	17	undergraduate degree in Pakistan	MP, English
m6	41	Mirpur	1	40	high school in UK	MP, English

[4] Although born in in the UK, participant m2 lived in Mirpur as a child between the ages of 5-10 years.

This participant sample is by design heterogenous in terms of relevant factors such as length of time living in UK, or language used at work, and so has the potential to reveal whether differences in amount of exposure to English impact on loanword adaptation. All 12 speakers produced MP words elicited using picture description with the expected MP stress patterns as described in section 2.1.

3.4 Procedure

Data collection sessions were conducted in MP by the first author, who is a first language speaker of MP, and took place in participants' homes or in suitable public places such as an office or community centre. Audio recordings were made in .wav format at 44.1kHz 16 bit using a Marantz PMD620 solid state recorder, via a head-mounted Shure SM10A microphone. After provision of informed consent, each participant provided answers to the language background questionnaire, before starting recordings. The loanword picture description task was one of several elicitation tasks undertaken in MP, for a wider study. At the end of the recording session each participant was asked to complete the Vocabulary Size Test in English on paper; the test comprises 140 multiple choice questions in which the participant must choose the correct meaning for a word in English.

3.4 Analysis

The production of each target loanword was phonetically transcribed by the first author. These transcriptions were checked and then coded for adaptation strategy by the second author. Each token was thus classified as an example of one of the adaptation strategies A, B, C or D, described in sections 2. The VST answer key was used by the first author to generate a score for each participant. The adaptation strategy category used in each token serves as a dependent variable, with VST scores and other relevant participant-related variables as potential predictors. Tests for correlation in the data were performed by the second author in R (Core Team, 2014) using Spearman's Rank and Kendall's Tau.

4. Results

Figure 1 shows a count of the number of observed tokens of each adaptation pattern type (A/B/C/D) in the 52 elicited loanwords for each speaker; the number of missing items is also reported, as in some cases the picture description task failed to elicit the desired target word.

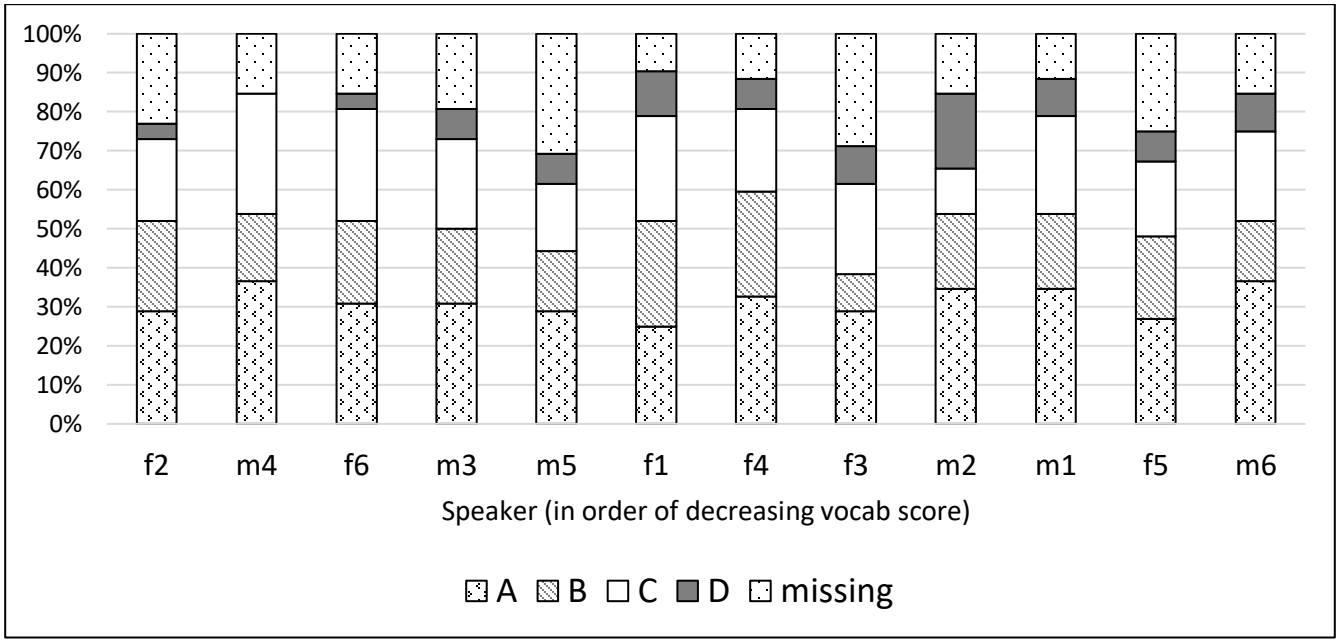


Figure 1. Count of each loanword adaptation pattern type (A/B/C/D/missing) by speaker.

All but one of the Bradford speakers show all four adaptation patterns, broadly matching the patterns observed in the loanword text corpus for the speech of younger speakers in Pakistan. This suggests that language use/exposure is indeed the factor which triggers use of the pattern C adaptation strategy, rather than age itself. The younger speakers in Pakistan whose speech patterns are reflected in the loanword text corpus have greater exposure to English than the older speakers in Pakistan represented there, whose data patterns shows no instances of pattern C.

In Figure 1 the speakers are ordered from left to right according to their score in the vocabulary size test, from highest to lowest score. Speaker m6 did not complete the full vocabulary test and is thus

excluded from further analysis. Speakers' individual vocabulary test scores are shown also in Figure 2 below. All speakers show a score in the test which is consistent with being a second language speaker of English. Nation (2012) notes that a score of 50-60 in the test (indicating receptive knowledge of the 5000-6000 most frequent word families) is similar to that shown by an undergraduate student for whom English is an additional language; an L2 English speaking PhD student would typically have a score of 90, approaching the typical vocabulary size of a 13 year old L1 English speaker (10,000 words). In our sample (excluding participant m6), scores ranged from 49-94 (mean 69.2, SD 22.5).

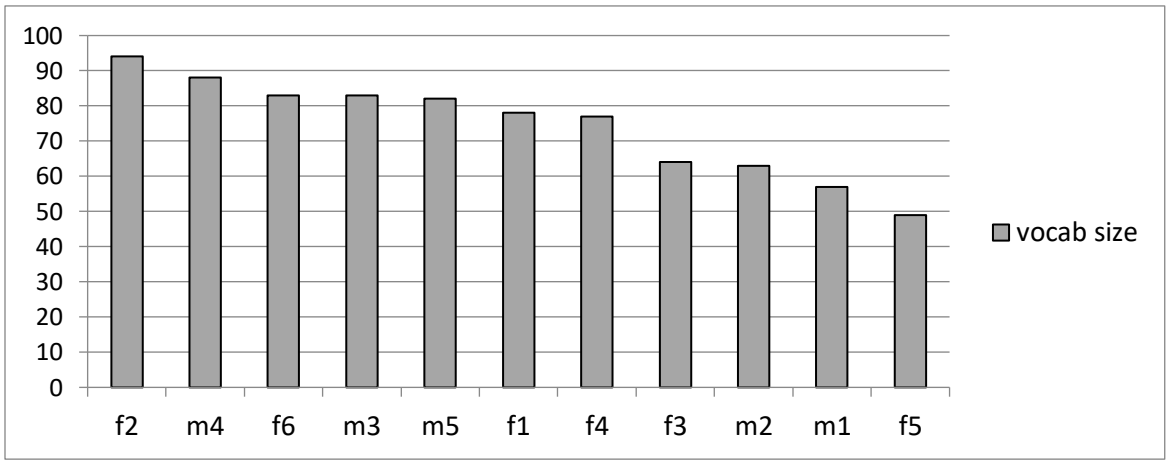


Figure 2. Vocabulary size test score, by speaker.

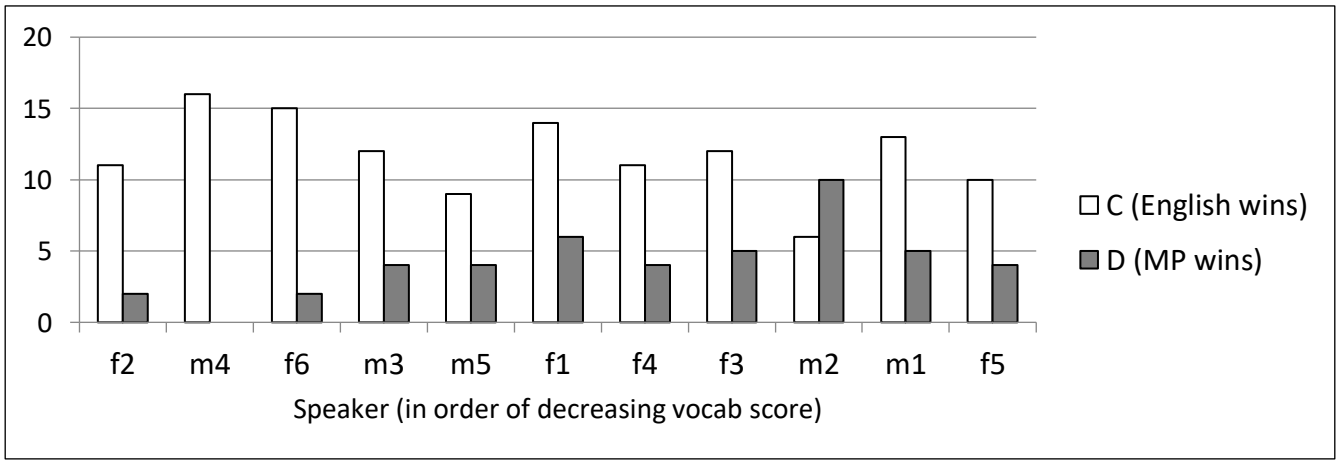


Figure 3. Count of pattern C and D tokens, by speaker.

Figure 3 shows the count of pattern C and D tokens, by speaker. Visual comparison of Figures 2 and 3 suggests an inverse relationship between the number of D tokens and a speaker's vocabulary size test score. Correlation tests confirm an inverse monotonic relationship between vocabulary size and the number of D cases (Spearman's rho = -0.7028; p = .0159; Kendall's tau = -0.5557; p = .0238), as illustrated in Figure 4, but no such relationship between vocabulary size and number of C cases (Spearman's rho = 0.3867; p = .24; Kendall's tau = 0.299; p = .0238).

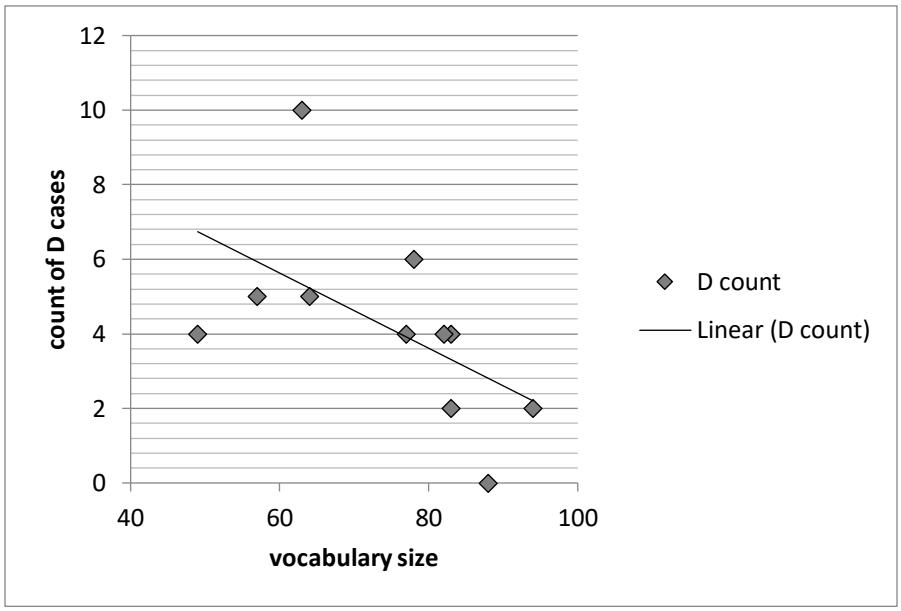


Figure 4. Scatter plot of vocabulary size and count of pattern D tokens by speaker.

Although the sample size is small, this result is consistent with the hypothesis that exposure to English is inversely related to the number of D cases (which follow MP rules), rather than to the number of C cases (which show stress on the syllable that is tressed in the English source word). We interpret this as an indication that all of the Bradford speakers have a sufficient number of English words in their lexicon, as would be expected, to generate an effect of faithfulness to the position of

stress in the English source word. In OT terms, all the speakers have had sufficient exposure to English to have 'switched' to a grammar which includes MATCHSTRESS.

The amount of observed intra-speaker variation in residual use of MP-like pattern D co-varies with vocabulary size to a significant extent, with increased exposure to English reflecting a reduced tendency to 'ignore' the English source word stress pattern in favour of adhering to the rules of MP.

5. Discussion and Conclusions

We have presented original data from a text corpus of English loanwords into the under-researched language Mirpur Pahari (Shafi, 2019), which show variation between younger and older speakers in Pakistan in the placement of stress in a subset of words (pattern C~D), as well as variable adaptation by younger speakers of otherwise structurally parallel words.

The patterns of variation between and within speakers were formalised in an OT grammar, which shows that the patterns of adaptation favoured by older speakers can be modelled with the same constraints and ranking that accounts for stress assignment in the MP language itself. A variable grammar, with re-ranking of an additional MATCHSTRESS constraint, is needed to account for the adaptation patterns of younger speakers in Pakistan, but the OT analysis does not offer a way to explain what triggers the re-ranking of constraints. An account taking degree of exposure to English as the external trigger for change was outlined within Message Oriented Phonology, and tested in a small-scale study with 12 MP speakers in the UK, in which adaptation patterns are correlated with an objective measure of receptive vocabulary size in English. The results are consistent with an account of variation both between and within speakers in terms of exposure to English.

Acknowledgements

Thanks to the Born in Bradford study community engagement team for help with recruitment of participants.

Appendix

Table A: Target words elicited from UK participants, with expected realisation in MP

source word	**realisation in MP**	**source word**	**realisation in MP**
perfume	pər.ˈfjuːm	camera	ˈkæmb.rə
alarm	ˈlaː.rəm	baby	ˈbeː.bi
balloon	bæ.ˈluːn	burger	ˈbər.gər
shampoo	ˈʃæm.pu	lettuce	ˈlæː.təs
giraffe	dʒə.ˈraː.fa	printer	pə.ˈrĩn.tər
cartoon	kar.ˈtuːn	paper	ˈpeː.pər
introduce	ɪn.tər.ˈduːs	mushroom	ˈmaʃ.rum
stadium	sə.teː.ˈdɪəm	curtain	ˈkər.tən
inspector	əns.ˈpæk.tər	blender	bə.ˈlæn.dər
injection	ən.ˈdʒæk.ʃən	teacher	ˈtiː.tʃər
library	læb.ˈreː.ri	student	sə.ˈtuː.də̃t
mobile	mə.ˈbaɪl	computer	kəmˈpuː.tər
basket	ˈbas.kɪt	carpet	ˈkar.pət
banana	bəˈnaː.nə	pilot	ˈpaɪ.lət
bracelet	bə.ˈræs.lət	decision	də.ˈsiː.ʒən
almond	ˈal.mə̃d	doctor	ˈdak.tər
pepper	ˈpeː.pər	patient	ˈpeː.ʃə̃t
potato	pə.ˈtæː.to	appendix	ˈpæn.dəs

monkey	ˈmaŋ.ki	cylinder	sə.ˈlæn.dər
zebra	ˈzeb.ra	hospital	həs.pəˈʈaːl
salad	sə.ˈlaːɖ	elephant	æ.lɪ.ˈfænt
tomato	tə.ˈmaː.to	cucumber	kə.ˈkəm.bər
onion	ˈon.jən	ambulance	æm.ˈboː.lə̃s
engagement	əŋ.ˈgeːdʒ.mɪt	broccoli	bə.ˈrok.li
umbrella	əm.ˈreː.la	terrorist	tæ.ra.ˈrɪst
mechanic	məˈkæː.nək	helicopter	hæ.lɪ.ˈkap.tər

References

Alber, B. (1998). Stress preservation in German loan-words. In W. Kehrein & R. Wiese (Eds.), *Phonology and morphology of the Germanic languages* (pp. 113-114). Tübingen: Niemeyer.

Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing, 27*(1), 101-118.

Bhatia, T. K. (1993). *Punjabi: A Cognitive Descriptive Grammar*. London: Routledge.

Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry, 32*(1), 45-86.

Broselow, E. (2009). Stress adaptation in loanword phonology: Perception and learnability. In P. Boersma & S. Hamann (Eds.), (pp. 191-234). Berlin: Mouton de Gruyter.

Core Team, R. (2014). *R: A language and environment for statistical computing*. Retrieved from Vienna, Austria: http://www.R-project.org/

Crosswhite, K. (1998). Segmental vs. prosodic correspondence in Chamorro. *Phonology, 15*(03), 281-316.

Davidson, L., & Noyer, R. (1997). Loan phonology in Huave: nativization and the ranking of faithfulness constraints. *Proceedings of the West Coast Conference on Formal Linguistics, 15*, 65-79.

Gyllstad, H., Vilkaite, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *International Journal of Applied Linguistics, 166*(2), 278-306.

Hall, K. C., Hume, E., Jaeger, T. F., & Wedel, A. (2016). *The message shapes phonology*. Ms. UBC, University of Canterbury, University of Rochester, University of Arizona.

Hamann, S., & Boersma, P. (2009). Loanword adaptation as first-language phonological perception. In A. Calabrese & W. L. Wetzels (Eds.), *Loan phonology* (pp. 11-58). Amsterdam: John Benjamins.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Hussain, S. (2015). Missing from the 'minority mainstream': Pahari-speaking diaspora in Britain. *Journal of Multilingual and Multicultural Development, 36*(5), 483-497.

Hyde, B. (2011). Extrametricality and non-finality. In M. Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 1027-1051). Oxford: Blackwell.

Kager, R. (2000). Stem stress and peak correspondence in Dutch. In J. Dekkers, F. Reinoud, H. van der Leeuw, & J. M. van de Weijer (Eds.), *Optimality Theory: phonology, syntax, and acquisition* (pp. 121-150). Oxford: Oxford University Press.

Kager, R. (2007). Feet and metrical stress. In P. De Lacy (Ed.), *The Cambridge Handbook of Phonology* (pp. 195-227).

Kang, Y. (2010). Tutorial overview: Suprasegmental adaptation in loanwords. *Lingua: International Review of General Linguistics, 120*(9), 2295-2310.

Khan, A. Q. (2012). *Phonology of Pahari: A Study of Segmental and Suprasegmental Features of Poonch Dialect.* University of Azad Jammu & Kashmir, Muzaffarabad.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing, 16*(1), 33-51.

McCarthy, J. (2008). *Doing Optimality Theory: Applying Theory to Data*. Oxford: Blackwell.

Meara, P., & Miralpeix, I. (2016). *Tools for Researching Vocabulary* (Vol. 105): Multilingual Matters.

Nation, P. (2012). Vocabulary Size in a Second Language. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Oxford: Wiley.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The language teacher, 31*(7), 9-13.

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar. [Revision of Prince & Smolensky (1993).]*. Oxford Blackwell.

Shafi, S. (2019). *A Phonological Analysis of English Loanwords in Mirpur Pahari: Exploring Variable Adaptation in Optimality Theory.* University of York, York.

Silverman, D. (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology, 9*(2), 289-328.

Stow, C., Pert, S., & Khattab, G. (2012). Translation to practice: Sociolinguistic and cultural considerations when working with the Pakistani heritage community in England, UK. In S. McLeod & B. A. Goldstein (Eds.), *Multilingual Aspects of Speech Sound Disorders in Children* (pp. 24-27). Bristol: Multilingual Matters.

Tabbassam, N. (2003). Phonological survey of Pahari Language. *Saroosh - Journal of Mirpur College, 11*(2), p380-397.