Feature Article

English Language Features Challenging for Nepali English Learners

Puskar Joshi, Zohreh R. Eslami, & Hector H. Rivera
Texas A&M University

Abstract

This paper explored and analyzed features of English pronunciation that could cause intelligibility problems for Nepali English learners (ELs), who use English as a foreign language (EFL) or English as a second language (ESL). We examined the Nepali ELs' pronunciation issues by juxtaposing them with comparable segmental and suprasegmental features of the Nepali language. We found that Nepali has far less sound inventory and voiceless consonants than English, yet it has a lot of contrastive consonants. Similarly, the Nepali language has fewer vowels, minimal diphthongs, and no triphthongs. While both English and Nepali share limited suprasegmental characteristics, they do not complement each other. Given the distinctiveness in pronunciation of the two languages, Nepali ELs are likely to face several intelligibility challenges as they encounter other non-Nepali English users. Thus, they need knowledge and appropriate training to enhance intelligibility for effective communication in workplaces and academic spaces that promote English.

Key Words *Nepali English learners; English pronunciation; intelligibility; teaching*

Introduction

This paper explores and descriptively analyses the features of English pronunciation that could cause intelligibility problems for Nepali ELs (also emergent bilinguals or multilinguals). The analyses of Nepali ELs' pronunciation problems are done by contrasting them with similar pronunciation features in Nepali. The main purpose of the current analyses is to help facilitate classroom instruction for Nepali ELs by providing teachers with insights into Nepali ELs' oral communication barriers.

Throughout this paper, *Nepali* refers to the Nepali language or people, whereas *Nepalis* refers to Nepali people or English users whose first language (L1 or home language or mother tongue) or second language (L2 or additional language) is Nepali. Finally, *English* refers to the English language spoken by native speakers of English.

Several reasons warrant this study. First, there has been a significant influx of Nepalis into English-speaking countries and educational contexts where English has been

preferred as a medium of instruction as (im)migrants and students until recently (see Mishra, 2011; Regmi, 2020), hinting that effective oral English is essential for Nepalis to survive and succeed in the new world. Thus, English medium instruction in Nepal and Nepali (im)migrant-receiving and student-welcoming countries must orient Nepali ELs toward their pronunciation issues and help them enhance intelligibility in oral English. Second, since English is a stress-timed language (Kenworthy, 1987; Roach, 2009), in contrast to the syllable-timed local languages in Nepal, Nepali ELs find English pronunciation hard to follow even if they may be aware of such differences. Finally, as English and Nepali come from different language families—English emerged from Anglo-Saxon and Nepali descended from Sanskrit—their sound systems are unique. When languages have different origins, they are likely to have distinct sound inventories. Accordingly, Nepali ELs need literacy in various English pronunciation aspects and receive appropriate training and strategies for successful communication. As the lack of extended sound inventory in L1 leads to pronunciation problems in L2 (Brekelmans, 2017), Nepali ELs must be prepared to deal with English's extended sound inventory. With all these reasons, teachers need to comprehend the phonologies of L1 and the target language to envision learners' pronunciation issues (see Burgess & Spencer, 2000) for effective classroom instruction.

Methods

This paper descriptively analyses the features of English pronunciation that are challenging for Nepali ELs by contrasting them with the features of Nepali pronunciation. The analyses are based on published literature, including the International Phonetic Alphabet (IPA).

Analyses of Contrastive Pronunciation Features of English and Nepali

Consonants

Consonants are produced in different manners and places of articulation (see Table 1). Jones (1998) describes the manner vocal cords function in producing consonant sounds as, "they may be held wide apart, they may be closed completely, or they may be held loosely together so that they vibrate when air passes between them" (p. 168). As vocal cords are held apart, the air passes between them, and voiceless sounds like [k] and [s] are heard; when we draw the vocal cords together, the air is forced to pass between them by vibrating cords that help us hear sounds like [d] and [g] (Khatiwada, 2009; Roach, 2009). Also, the vocal cords do not vibrate while making voiceless sounds, implying that vibration produces voicing.

Table 1

Nepali and English Consonants

Manner of Articulation	Place of Articulation																	
	Bilabial		Labiodental		Dental		Alveolar		Postalveolar		Retroflex		Palatal		Velar		Glottal	
	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced
Plosive	p	b					t	d			**ʈ**	**ɖ**			k	g		
	pʰ	**bɦ**			**tʰ**	**dɦ**					**ʈʰ**	**ɖɦ**			**kʰ**	**gɦ**		
Affricate									tʃ	dʒ								
									ts	**dz**								
									tsʰ	**dzɦ**								
Nasal	m	**m**					n	**n**							ŋ	**ŋ**		
Tap or flap								**r**										
Fricative			f	v	θ	ð	s	z	ʃ	ʒ							h	**ɦ**
	fʰ	**vɦ**																
Latteral							l	**l**										
Approximant	w	**w**							r				j	**j**				

Adapted from Khatiwada (2009) & Roach (2009).

Notes on the table:

- English consonant sounds are printed normal.
- Nepali consonant sounds are shown in boldface.
- English consonants close to Nepali consonants are underlined.
- The consonants with superscripts represent contrastive sounds in Nepali.

Plosives

Plosives are consonant sounds that are produced by holding the air coming through the vocal cords in the mouth, compressing the air, and releasing it in the final stage. To

produce English plosives, we move two of our speech organs against each other, obstructing the passage through which the air comes from the lungs; then the sudden force of air opens the vocal organs, resulting in a sound loud enough to hear. The production of plosives can be divided into four phases—the closure phase, in which both the speech organs close the air tract; the hold phase, in which the air is held in the lungs; the release phase, in which the air is released from the lungs, and the speech organs are opened; the post-release phase in which the air is released resulting in a plosive sound. The post-release phase does not always necessarily occur. There are altogether six plosive sounds—[p], [t], [k], [b], [d], and [ɡ]—which can occur in any position in English words, e.g., [p] in *pin*, *paper*, and *pop*. (Roach, 2009)

English and Nepali plosives are similar, although Nepali uses contrastive plosives. Also, Nepalis do not adequately press the lips to produce [p], [k], and [t] sounds, therefore their English becomes unintelligible. Specifically, when they pronounce words with [p], [k], and [t] sounds at the beginning of words, like *pencil*, *tanker*, and *caravan*, those sounds may not be heard adequately. However, judging from Dalton and Seidlhofer's (1994) description of how the sound [p] is pronounced, the speaker needs to use significant force and muscular strength to make English plosives.

Affricates

Affricates start as a plosive and end as a fricative sound. They are the combination of plosive and fricatives and are produced in the same vocal organs. Affricates in English "begin with plosives and end with fricatives" (Roach, 2009, p. 39), and they have the same kind of articulation as plosives in that they have closure and hold phases. Accordingly, we have the rapid release of air with plosion and aspiration that lets us move our tongue to make the affricate sound. Put differently, affricates are produced when a stop has a fricative release, e.g., *push* in the beginning has a plosive sound followed by a vowel and a fricative release. English affricates are [tʃ] and [dʒ], which occur in any word position: [tʃ] in *church*, *catching*, and *match* and [dʒ] in *jump*, *pigeon*, and *edge*.

Nepali does not have the exact postalveolar affricates—[tʃ] and [dʒ] (see Table 1). However, Nepalis often pronounce them by replacing them with comparable postalveolar affricates—[ts] and [dz]. In the case of [tʃ], it may not differ much from [ts]. Khatiwada (2009) describes [ts] and [dz] sounds in Nepali produced through "laminal contact in the alveolar region" but heard as palatoalveolar sounds (p. 375). Interestingly, Nepalis might replace the postalveolar voiced affricate [dʒ] with voiced alveolar fricative [z], so they often pronounce *language* as [læŋgwɪz] or [læŋgwɪdz].

Fricatives

Fricatives are characterised as hissing sounds, which are produced as the air escapes through our mouth. While making English fricatives, the air passage of our speech organs becomes narrow at some point, and the air is expelled by pressure from the lungs "so that the escaping air produces friction" (Dalton & Seidlhofer, 1994, p. 15). Fricatives are noticeable because they can be continued and prolonged (like a hissing sound) while the air is in the lungs (Roach, 2009).

Seven English fricatives—[f], [θ], [s], [ʃ], [v], [ð], and[z]—occur in all the positions of English words, e.g., [f] in *fan*, *beautiful*, and *relief*. However, [ʒ] occurs only in the middle and final positions, such as *measure* and *rouge*, and [h] occurs only in the initial and middle positions, as in *honey* and *ahead*.

In making English labiodental fricatives, the lower lip is in contact with the upper teeth (Roach, 2009); the tongue is usually placed inside the teeth, with the tip of the tongue touching the inside of the lower front teeth. Indeed, the labiodental fricatives are "articulated by the inside of the lower lip making light contact with the upper teeth" (Dalton & Seidlhofer, 1994, p. 15). Since Nepali does not have labio-dental fricatives (see Table 1), Nepalis tend to replace these sounds with contrastive bilabial fricatives [f^{h}] and [$b^{ɦ}$]—respectively. The former is a voiceless bilabial fricative, and the latter is a voiced bilabial fricative, thus *fan* is pronounced as [f^{h}æn], and *van* is pronounced as [$v^{ɦ}$æn].

While making English dental fricatives, the tip of the tongue meets the front teeth as if the tongue was placed between the teeth (Roach, 2009). The dental fricatives of English sometimes cause intelligibility problems for Nepalis because Nepali has dental plosive sounds [t], [d], [t^{h}], and [$d^{ɦ}$] but not dental fricatives. Usually, Nepalis replace English dental fricatives with contrastive Nepali dental plosives.

In making postalveolar fricatives as [ʃ] and [ʒ] in English, the tongue is in contact with an area slightly further back than that for [s] and [z]. In pronouncing [s] and [z], the tip of the tongue meets the alveolar, and the centre of it raises. Nepalis often replace this sound with the voiceless alveolar fricative [s]. Consequently, when Nepalis pronounce *shoes* [ʃu:z], they say [su:z].

Nepali does not have the voiced postalveolar fricative [ʒ] sound. Therefore, Nepalis cannot intelligently pronounce English words with this sound without adequate training. Whenever this sound appears, they tend to replace it with the voiced alveolar fricative [z] sound, e.g., they pronounce *measure* as [mezə]. Sometimes, they replace [z] with the contrastive voiced alveolar affricate [dz] sound, as shown in Khatiwada (2009).

Finally, Nepalis can pronounce the English voiceless alveolar fricative [s] sound without difficulty, yet they have a problem with word-initial consonant clusters. Thus, if [s] is followed by [p], [t], [k], [m], or [n], Nepalis tend to pronounce the sound by inserting [ɪ]

before [s]; e.g., *smart* is pronounced as [ɪsma:t]. Presumably, this problem is caused by mother tongue interference because Nepalis tend to insert [ɪ] sound before them. However, if [s] is followed by an approximant [w] or lateral [l], Nepalis pronounce the sound without any obstruction.

Other Consonants

Other consonants namely nasal, tap or flap, lateral, and approximant do not cause major intelligibility issues for Nepali ELs. To pronounce nasal sounds like [m n], the air coming from the lungs passes through the nasal cavity. However, to pronounce, tap or flap sound like Nepali [r], one articulatory organ (tongue) taps or flaps against another organ once. On the other hand, to pronounce lateral sounds like [l], the air passes through the sides of the tongue. Finally, to pronounce approximants like [w], vocal organs like lips come together, but without touching each other. Specifically, Nepalis pronounce English dental fricatives (fricatives pronounced by touching the back of the upper teeth by the blade of the tongue), nasals, lateral, tap or flap, velar plosives (plosives pronounced by raising the body of the tongue to touch the palate), and the approximant sounds (see Table 1) without intelligibility issues although they tend to replace these English consonants with contrastive Nepali consonants. In specific, they may replace English dental fricatives (fricatives pronounced by forcing the tip of the tongue against the back of the teeth) with Nepali dental plosives (plosives pronounced by placing the tip of the tongue behind the teeth and forcing the airflow out), English velar plosives (plosives pronounced by raising the body of the tongue against the palate and forcing the airflow out) with Nepali velar plosives, and English nasals with Nepali nasals. However, the bilabial approximant (approximant sounds made by partially stopping the airflow coming from the lungs and releasing it) [w] often causes intelligibility problems if it appears in the beginning of an English word starting with *wa* letters as in *water*, *Walter*, or *walk.* Usually, Nepalis pronounce *wa* as if it has [ɑ:] sound. (Khatiwada, 2009; Roach, 2009)

Vowels

When we pronounce vowels, there is no sound obstruction as the air passages through our mouth (Jones, 1998; Roach, 2009). Vowels are produced at different places and manners, which determines the length and quality. As vowels are produced as long and short, opening the mouth and tongue positions are crucial in pronouncing them (Roach, 2009). Table 2 illustrates the manner and place English vowels, and the *schwa* are produced.

Table 2

Description of English Vowels and Schwa

Sound	*Type*	*Position & Manner of Articulation*	*Examples*
i:	*Long*	Close front; slightly spread lip position.	Sheep, beach

ɜ:		Mid-central; neutral lip position.	Bird, fern
ɑ:		Open but not back vowel; neutral lip position.	Card, half
ɔ:		Fully back, strong lip rounding.	Board, torn
u:		Less back, moderately rounded lips.	Food, soon
ɪ	*Short*	Close front; slightly spread lip.	Bit, fish
e		Front; slightly spread lips.	Bet, men
æ		Not quite open; slightly spread lips.	Bat, man
ʌ		Central and more open; neutral lip position.	Cut, come
ɒ		Not fully back, but open mid and open in tongue height; slightly rounded lip position.	Pot, cross
ʊ		More open and closer to central; rounded lips.	Book, full
ə	*Schwa*	Mid and central; not pronounced with force	About, character

Source: Roach (2009, pp. 8-17)

English has 11 vowels (long and short vowels) and the *schwa* [ə], which is found in weak and unstressed syllables (Roach, 2009). Nepali vowels are characterised as nasal (Khatiwada, 2009). Koffi (2019) argues that Nepali does not have five English vowel counterparts: [ɪ], [ɛ], [æ], [ɔ], and [ʊ]. However, this statement is not entirely true because of the presence of two vowels—[i] and [u]—which are often used as [ɪ] and [ʊ]. For instance, when Nepalis pronounce English words like *beach*, it is heard as *bitch*—an embarrassment in many situations. Similarly, when they pronounce the word *sheep*, it is heard as *ship.* Also, they tend to favour [ʊ] over [u:]; so, when they pronounce the word *pool,* it is heard as *pull.* The other vowel problem Nepalis often encounter is related to *schwa*. Because Nepali lacks *schwa*, it causes a communication challenge to Nepalis because they stress the syllable with *schwa*, as *about* [**ə**baʊt] instead of [ə**baʊt**]. Khatiwada (2009) points to another issue in that Nepalis may replace the *schwa* with the vowel [ʌ], which is described as "slightly, rounded, and […] acoustically lower and more back" (p. 377).

Diphthongs and Triphthongs

Diphthongs are the fusion of two vowels in a single syllable, e.g., *day* [deɪ], which combines [e] and [ɪ]. Likewise, triphthongs combine three vowels within a single syllable, for instance, *mower* [məʊə]. English has eight diphthongs, [ɪə], [eə], [ʊə], [eɪ], [aɪ], [ɔɪ], [əʊ], and [aʊ], and five triphthongs, [eɪə], [aɪə], [ɔɪə], [əʊə], and [aʊə]. (Roach, 2009)

The most crucial aspect of English diphthongs is that it must occur in a syllable (Catford, 1988, p. 116). Indeed, there is a movement of sound from one (first) vowel to another (second) in a diphthong, meaning there are two sounds and one movement. As the sound moves, the second part remains short (Roach, 2009). Nevertheless, in a triphthong, one

more sound with one extra movement, from the first to the third vowel. The challenge in using triphthongs is that they must be pronounced quickly. However, foreigners usually must notice whether triphthongs have two or more combined sounds (Roach, 2009).

In Nepali, there are as many as ten diphthongs [ui], [ei], [oi], [ʌi], [ai], [iu], [eu], [ou], [ʌu], and [au] (Bandhu, 1989 cited in Khatiwada, 2009) but no triphthongs. Among the diphthongs, Nepali seems to share only [eɪ] and [aɪ] with English because it lacks distinction between [i:] and [ɪ] (see Khatiwada, 2009). That said, while Nepalis may recognise diphthongs as they hear them in English, it is not necessarily easy for them to distinguish between words containing single vowels and diphthongs. Eventually, they fail to internalise and use diphthongs correctly. Also, since Nepali does not have triphthongs, Nepalis could reduce triphthongs to diphthongs. Consequently, when they say *dose*, it could be heard as [dəs], or [dʌs]; not [dəʊs]. And when they say *lower*, it could be heard as [ləwə], [lʌwʌ], [lɔwə], or [lɔwʌ]; not [ləʊə].

Word-Stress

English words (except for unstressed ones) are pronounced with distinctive force on a stressed syllable. In English, stress occurs at regular intervals between the stressed syllables even if there are many intervening unstressed ones. That said, only one syllable of each English polysyllabic word is pronounced with primary stress (Kenworthy, 1987). For example, *photographic* is pronounced as [fəʊtə**gr<u>æ</u>**fɪk], where the primary and secondary stresses fall on the third (bold and underlined) and first (bold) syllables. Here, the third syllable of *photographic* is pronounced with the speaker's muscular effort, suggesting that the speaker is talking about *photographic*. Nevertheless, if the word is pronounced with a stress on other syllables, it may mean something else or nothing. Kenworthy (1987) warns that if the speaker stresses the wrong syllable in English, listeners will have a problem identifying the word.

Essentially, length, loudness, vowel quality, and pitch help recognise English stress. However, Vanderplank (1993) argues that stressed syllables are not necessarily longer and louder. Roach (2009) also considers some of the four factors equally crucial and claims that sometimes a syllable can be made noticeable either by one or two factors. For Catford (1988), loudness is more associated with sound rather than stress.

Also, the pitch of the voice is higher on the stressed syllables than on other syllables in English, and the former has a different vowel quality from vowels in the neighbouring syllables. While stressed syllables can have any vowel in them (long or short), syllabic consonants, like in *listened* [lɪsnd], and, *schwa*, like in *manor* [mænə], appear only in unstressed syllables. (Roach, 2009)

In languages like Nepali, stress does not play a crucial role in speakers' pronunciation. Indeed, Nepalis do not use primary stress frequently. Even if they do, they stress any

syllable, particularly the first or second syllable of a word (Acharya, 1991), based on what they want to emphasise and how they want to do so. Sometimes, they may pronounce words with a level stress.

Also, English word-stress rules cannot capture all words, creating intelligibility problems for foreigners. While English stress does not have a fixed place, i.e., any syllable of a polysyllabic word can carry the primary stress (Kenworthy, 1987), it is "fixed in the sense that every word has [unique] stress pattern" (Dalton & Seidlhofer, 1994, p. 39). Interestingly, the changing nature of word-stress also applies to Nepali: When two words with primary stress on the initial syllable form a compound word, the primary stress falls on the first syllable (Acharya, 1991). Nevertheless, stress does not drastically alter the meaning of a Nepali word as in English.

Finally, English stress is mobile, so it moves from one syllable to another within semantic families of words. For example, the stress in the word *photograph* is on the first syllable, whereas in *photography*, it is on the second syllable. Thus, learners can only learn the English word-stress in a communicative context, meaning they must get chances to practise the correct stress patterns in real contexts. As Dalton and Seidlhofer (1994) warn that "incorrect word-stress decreases intelligibility" (p. 39), learners need to use word-stress approximating English norms in different communicative situations. Even though stress shifts are largely consistent on the word category change in English, teachers can decide on the classroom instructional strategies based on their expertise and curriculum requirement.

Rhythm

Rhythm in English is useful to indicate what is coming next (see Martin, 1972, 1975; Vanderplank, 1993). Rhythm is often compared to our heartbeat; it foregrounds important syllables like images on a background (Dalton & Seidlhofer, 1994). In English, rhythm occurs at regular intervals on the stressed syllables, so listeners expect to hear it from speakers (Kenworthy, 1987). However, whether speakers speak rhythmically could vary: "[Sometimes people] speak very rhythmically […] while at other times [they] speak arhythmically […] if [they] are hesitant or nervous" (Roach, 2009, p. 109).

Rhythm in Nepali occurs at regular intervals on stressed or unstressed syllables. However, as each syllable is articulated for approximately the same duration in Nepali (Acharya, 1991), Nepalis tend to use syllable-timed rhythm, affecting their intelligibility in oral English. This is the direct influence of the syllable-timed nature of the Nepali language.

Connected Speech

When people speak English, they pronounce phonemes of some words differently because of the neighbouring words, such as *in the* [innə] wherever the pair of words is

used together. Indeed, assimilation changes our tongue position during conversation and prevents the tongue from reaching the ideal position, where the individual sounds are made (Dalton & Seidlhofer, 1994). Also, as we speak, some sounds may disappear, e.g., *looked back* [lʊk bæk], a situation called the zero realisation of sound or sound deletion or elision (Roach, 2009) where sounds like consonant clusters and the *schwa* may become weak due to elision. Sometimes people may insert extra sounds to link words, as in [mi:dɪər ɪvent], which links *media* and *event*. Here, the extra sound [r] enhances pronunciation accuracy and speech fluency by allowing speakers to fill the gaps between words (Roach, 2009).

As Nepalis tend to pronounce every word and syllable clearly, it is natural for them to anticipate English speakers doing the same. That said, elision and assimilation are not uncommon in Nepali. Elision like *gako thiye* [gʌko t^{ɽ}i̲ẽ̲] for *gayeko thiye* (or has gone) [gʌeko t^{ɽ}i̲ẽ̲] and *bhako thiyo* [b^{h}ʌko t^{ɽ}ijo] for *bhayeko thyo* (or has happened) [b^{h}ʌeko: t^{ɽ}ijo] are commonly used in Nepali. Interestingly, people may hear [gako] and [b^{h}ako] rather than [gʌko] and [b^{h}ʌko] as examples of assimilation. Yet elision and assimilation of Nepali and English are unrelated, meaning the knowledge of one does not help one understand the meaning of another. As the elision and assimilation of Nepalis get influenced by their mother tongue, there are always chances of misunderstanding when they converse with other English speakers.

Intonation

Intonation is the rise and fall of voice when people speak. In natural settings, people's voices may go up, down, and stay level as they talk.

In English, speakers "can change the meaning of what they say through [intonation]" (Hammer, 1996, p. 12). This view is supported by Bradford (1988): "[People] can mean different things by using the same group of words, arranged in the same order, but saying them in different ways" (p. 1). Kenworthy (1987), too, agrees: "[Listeners] get information using pitch of the voice and speakers send information using pitch variation" (p. 19). In other words, unintelligible intonation leads to breaking up the communication because the correct information is unlikely to get conveyed to the listener.

When people interact, they show their feeling, attitude, and mood towards what they are talking about and the person they are speaking to. Such attributes can be positive, negative, or neutral and can be expressed through intonation. Specific to English, people use intonation for various purposes. They use falling [↘] intonation for finality and definiteness, lack of interest in a topic or for a closed communicative situation. Kenworthy (1987) asserts that the pitch usually drops when the last item in a list has been reached, e.g., *Bill has a daughter, two sons, and a granddaughter*. They also use falling intonation to expect a supportive response or other people to have

the same information as *You saw my computer, didn't you?* In contrast, they use rising [↗] intonation for general questions, like *Do you like her?* or to hint at a continuation of the speech, to show interest in a topic, to show suspicion about other, and to encourage someone to do something. People rarely try to speak if the person they are talking to is still speaking with a high pitch; if they do, it could be considered an interruption (Kenworthy, 1987). However, if people actively participate in the conversation, they may use a high pitch to express *yes*, *fine*, and *okay*. Furthermore, if they wish to be polite, they use a rising intonation, e.g., *Give it to me, please*.

English speakers use fall-rise [↘↗] intonation for uncertainty, doubt, and reservation. So, it is fair to say that a fall-rise nucleus often conveys the idea of *but*, for instance, *Bob will undoubtedly win the presidential election*. Contrarily, they use rise-fall [↗↘] intonation to show (dis)approval or to convey that the speaker is impressed. On the other hand, they use level intonation [→]for sharing routine information or making neutral assertions. (Roach, 2009)

In an English conversation, speakers often want to foreground certain things and background others. In such cases, the intonation is higher on what the speakers want to underscore, or the speakers use rising and falling intonation sharply so that the important words stand out (Kenworthy, 1987), which often carry more semantic weight. Such words are nouns, verbs, and adjectives. The function words like articles, auxiliaries, and prepositions carry less weight, yet speakers could also destress content words if it takes more time to pronounce stressed syllables (Dalton & Seidlhofer, 1994).

If English speakers presume something is not essential or has already been known to listeners, the intonation goes down, and the new information is emphasised. Ladefoged (1982) argues that "the topic of a sentence is less likely to receive the tonic accent than the comment which is made on that topic" (p. 94) because the comment is unpredictable. Accordingly, a falling tone indicates shared information, and a rising tone indicates new information (Cauldwell & Hewings, 1996). Additionally, English speakers use a non-falling tone when they start a topic and a falling tone when they want to close the topic.

Indeed, intonation allows English speakers to communicate their purpose (Verdugo, 2005). However, the choice of tone that non-native speakers use is different, implying native and non-native speakers use distinct intonation patterns when they speak English (Verdugo, 2005). Although intonation is not as difficult as stress for Nepalis, they tend to use a narrow pitch range, which can cause intelligibility problems because it can function as a level tone and indicate uncertainty in English (Snow & Balog, 2002; Verdugo, 2005). Indeed, when speakers want to demonstrate certainty, they must use falling intonation to prevent communication breakdowns and unintelligibility between interlocutors (Verdugo, 2005). Though the context is vital to navigating the speakers'

purpose, the contextual rules of English intonation could pose a problem for Nepalis. As Cauldwell and Hewings (1996) point out, spontaneous discourse may scantily follow the norm of new information being expressed with rising and shared information with falling intonation. Accordingly, there is always a possibility that speakers might use a level tone for familiar information (or to repeat ritualistic details) and a rising tone for the last item in a list expressed earlier if they want to add a new item (Cauldwell & Hewings, 1996).

Discussion and Implications

The current analyses reveal huge differences between English and Nepali pronunciation. Because the differences are inherent to the two languages, teachers of Nepali ELs can focus on making their pupils aware of those variations and prepare them through carefully designed lessons and assessments. So, Nepali ELs must get ample opportunities to become aware of the distinctions in pronunciation between their home language and English and receive sufficient training to enhance intelligibility. However, awareness and mechanical training would not suffice as they could overwhelm ELs with theory. Put differently, EFL and ESL teachers must use varieties of research-based activities and strategies proposed in different sources (Roach, 2009) and create an authentic learning environment, where ELs can meaningfully use and internalise the target language pronunciation. Yet the environment should allow ELs to concentrate on problematic pronunciation issues. Thus, teachers should carefully devise activities prioritising learning tasks, giving learners sufficient time to listen and practise the problems, providing them with timely feedback, and assessing their progress (Kenworthy, 1987). Also, teachers must allow learners to use their schema as suggested by Vygotsky (1978), who argued that higher order cognition is achieved when teachers scaffold learning from what learners already know to what they need to learn.

This study will serve as a resource to many. Educators, in specific, will get insights into Nepali ELs' English pronunciation issues to revisit curricula, educational materials, lesson plans, assessments, and pedagogical approaches in practice to address the oral communication issues of Nepali ELs in Nepal and other relevant situations. This study will also inform them about the potential areas of English pronunciation that could lead to miscommunication and misunderstanding with Nepali ELs. Researchers could find it as a useful model to conduct a comprehensive and comparative analysis of pronunciation elements across languages and extend this study by collecting data from ELs.

Conclusion

We examined elements of English pronunciation that could cause intelligibility issues for Nepali ELs. The current analyses suggest that Nepali ELs may need help to use both

segmental and prosodic features of English intelligibly due to the differences in pronunciation features between English and Nepali. Nepali has far less sound inventory than English in terms of segmental features, contributing to the intelligibility challenges. Specific to consonants, English has more voiceless consonants than Nepali whereas the latter has contrastive consonants absent in the first. Also, while Nepali has fewer vowels and diphthongs than English, it does not have triphthongs at all, adding to the complexity for Nepalis to use oral English intelligibly. Finally, while English and Nepali both share specific suprasegmental characteristics like word-stress, rhythm, connected speech, and intonation, they do not complement each other because of the differences in their purpose, contexts, and functions.

Though we identified English pronunciation issues that could cause intelligibility challenges for Nepalis, the problems analysed may not fully represent and reflect the English pronunciation issues of all Nepali ELs as English is a global language, and communication strategies and awareness of these challenges may enhance communication and intelligibility even with these divergences. Second, since interactions between people are primarily held in real-life contexts, the isolated English pronunciation issues described in this paper may have less bearing on people's day-to-day interactions, but if the problems are not addressed timely, they could lead to stigmatisation and stereotyping (Vincze & MacIntyre, 2017).

The attitudes of interlocutors, acceptance of ambiguity in pronunciation (e.g., awareness and exposure to different varieties of English) and understanding the message, and willingness to collaborate positively with different users of English with different accents and pronunciations may compensate for some of these challenges (see Ates et al., 2015). Therefore, future studies need to consider social, political, and cultural issues, not just the linguistic ones addressed in this study. As this paper intends to raise awareness of English pronunciation issues among the Nepali ELs caused by differences in English and Nepali and support EFL or ESL teaching to facilitate learners' communication in international settings, we ask readers to judge these (de)merits cautiously.

References

Acharya, J. (1991). *A descriptive grammar of Nepali: An analysed corpus* (1st ed.). Georgetown University Press.

Ates, B., Eslami, Z. R., & Wright, K. L. (2015). Incorporating world Englishes into undergraduate ESL education courses. *World Englishes*, *34*(3), 485–501.

Bradford, B. (1988). *Intonation in context* (1st ed.). Cambridge University Press.

Brekelmans, G. (2017). The value of phonetics and pronunciation teaching for advanced learners of English. *Linguistica, 57*(1), 45–58.

Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System, 28*, 191–215.

Catford, J. C. (1988). *A practical introduction to phonetics* (1st ed). Clarendon Press Oxford.
Cauldwell, R., & Hewings, M. (1996). Intonation rules in ELT textbooks. *ELT, 50*(4), 327–334.
Dalton, C., & Seidlhofer, B. (1994). Pronunciation (1st ed.). Oxford University Press.
Hammer, J. (1991). *The practice of English language teaching* (1st ed.). Longman.
Jones, D. (1998). *The pronunciation of English* (Definitive ed.). Cambridge University Press.
Kenworthy, J. (1987). *Teaching English pronunciation* (1st ed.). Longman.
Khatiwada, R. (2009). Nepali. *Journal of International Phonetic Association, 39*(3), 373–380.
Koffi, E. (2019). An acoustic phonetic analysis of the intelligibility of Nepali-accented English vowels. *Linguistic Portfolios, 8*, 94–112.
Mishra, P. B. (2011). Nepalese migrants in the United States of America: Perspectives on their exodus, assimilation pattern and commitment to Nepal. *Journal of Ethnic and Migration Studies*, *37*(9), 1527-1537.
Regmi, P. R., Aryal, N., van Teijlingen, E., Simkhada, P., Adhikary, P. (2020). Nepali Migrant Workers and the Need for Pre-departure Training on Mental Health: A Qualitative Study. *Journal of Immigrant and Minority Health, 22*(5), 973-981.
Ladefoged, P. (1982). A course in phonetics (2nd ed.). Harcourt Brace.
Roach, P. (2009). *English phonetics and phonology: A practical course* (4th ed.). Cambridge University Press.
Snow, D., & Balog, H. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua 112*, 1025–1058.
Vanderplank, R. (1993). 'Pacing' and 'spacing' as predictors of difficulty in speaking and understanding English. *ELT, 47*(2), 117–125.
Verdugo, D. R. (2005). The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: Pragmatic effects. *Journal of Pragmatics, 37*, 2086–2115.
Vincze, L., & MacIntyre, P. (2017). Accent stigmatization as a moderator of the relationship between perceived L2 proficiency and L2 use anxiety. *Applied Linguistics Review*, *8*(1), 61–78.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (1st ed.). Harvard University Press.

Authors

Puskar Joshi is a doctoral student in the Department of Educational Psychology. A graduate from Columbia University and The University of Manchester, Puskar taught EFL in Nepal. His research interests include education policy; language and literacy; bilingual, multilingual, ESL, and multicultural education.

Dr. Zohreh R Eslami is a professor and serves as the coordinator of the PhD cohort in Bilingual/ESL program at the Department of Educational Psychology. An internationally acclaimed researcher with expertise in areas of interlanguage and intercultural pragmatics, linguistic and cultural diversity, computer mediated language learning and teaching, cyber-pragmatics, and language and gender construction on social media using

multilingual practices. Dr. Eslami is the managing and co-editor of *Applied Pragmatics* journal.

Dr. Héctor H Rivera is an Associate Professor at Texas A&M University, in the Department of Educational Psychology. He is also the founder and former director of the Center for Child and Community Development at Southern Methodist University in Dallas, Texas. He has also co-authored State level educational curriculum for the State of California Steck-Vaughn California Gateways. In 2019, he received the United States Congressional "Education Champion" Award, during a ceremony at the Dallas Art Museum.