Nepalese Linguistics, Journal of Linguistic Society of Nepal; ISSN: 0259-1006(Print)

INTELLIGIBILITY OF NEPALI ENGLISH VOWELS: RATIONALES FOR TEACHING PRONUNCIATION

KRISHNA PRASAD CHALISE
Central Department of Linguistics, T. U.
krishna.chalise@cdl.tu.edu.np

(Received: 08 Sept., 2025; revised: 06 Oct., 2025; accepted: 05 Nov., 2025; published: 26 Nov., 2025)

This paper has investigated the possible challenging vowels for the mutual intelligibility between NE (Nepali English) and GAE (General American English) based on internal and external masking and functional load. It is based on the data from secondary sources. The finding is that there is a remarkable intelligibility gap between NE and GAE because of internal and external masking, and phonetic distinctiveness between the corresponding vowels in them. Pronunciation is crucial for intelligibility, comprehension, and overall language proficiency. Still, it is often viewed as incidental to other language skills and is not included in textbooks, the evaluation system, and teacher training programs. For intelligible pronunciation, the listener should easily identify the phonological content of the speaker. For this purpose, they should activate both analytical skills (identifying and discriminating the phonetic features) and motor skills (articulating the phonetic features). This achievement is possible through rigorous pronunciation training based on the findings from empirical research.

Keywords: Speech perception, intelligibility, pronunciation teaching, F1 distance, functional load

1. Background

Pronunciation or the phonetic aspects of a language are important for intelligibility, comprehension, and general language ability. But the significance of pronunciation for successful communication has been overlooked in TESOL classes in Nepal. Pronunciation is often viewed as secondary or incidental to other language skills, such as grammar and vocabulary. This devaluation is demonstrated by the lack of attention placed on it in textbooks, the evaluation system, and teacher training programs.

The situation is almost the same globally. Derwing and Munro (2015) say that the study of pronunciation has been "marginalized," and as a result, instructors are left to draw on intuitions with little direction or guidance. Generally, English teachers overlook pronunciation, citing that they are not skilled at teaching it, do not teach it, or consider it unimportant because it is not assessed in examinations. Additionally, they often claim that tools for teaching pronunciation are unavailable. Macdonald's (2002) study in Australia shows that the curriculum does not contain a pronunciation component to be taught, and no way to assess pronunciation achievement. The issue of identity and a lack of suitable materials for pronunciation teaching are the reasons for the situation.

The curricula neglect pronunciation because people believe that exposure to real-world listening materials during language learning will automatically improve pronunciation. According to Sardegna and Jarosz (2023), researchers have falsified the belief. Several studies have justified that teaching pronunciation effectively raises pronunciation awareness and encourages more deliberate speech production and perception, all of which help L2 learners pronounce words and utterances more correctly.

Pronunciation errors can cause major consequences by causing misunderstandings and

communication breakdowns. Arguing for the importance of L2 pronunciation, Hurtado and Estrada (2010) claim that pronunciation is potentially one of the "most important aspects in learning a second language" because it improves communication as well as it creates a positive image of the learner's abilities. Shaw's (1981) study found that 47% of Singaporeans, 59% of Indians, and 62% of Thai bachelor's final-year students reported that pronunciation was their weakest among the four skills of English. But among them, 71% of Singaporeans, 71% of Indians, and 88% of Thais wanted speaking to be their best English skill. Intelligible spoken English is essential in professional settings, such as client meetings, presentations, job interviews, and marketing. Effective pronunciation skills can help applicants connect more effectively with stakeholders, clients, and coworkers, which can enhance job prospects and promote opportunities for advancement.

Acquiring native-like pronunciation in L2 is unattainable (Flege et al., 1995); therefore, achieving intelligible pronunciation is the objective of acquiring good pronunciation competence (Szyszka, 2023). Intelligibility denotes the degree of match between a speaker's intended message (coding) and the listener's comprehension (decoding), whereas comprehensibility refers to how much effort the listener needs to put into understanding an utterance, and accentedness describes the extent of L1 accent detected in L2 speech (Derwing & Munro, 2015).

2. Methodology

This paper is based on secondary data. The formant values of the Nepali vowels are from Chalise (2022), the formant values of the Nepali English (NE) are from Chalise (2025), and the formant values of General American English (GAE) are from Hillenbrand, et al. (1995).

Intelligibility analysis relies on the position of the vowels in the acoustic vowel space, level of masking caused by acoustic distance (F1 difference), and functional load (FL) as outlined by Koffi (2021). If the F1 distance between two vowels within the same natural class (high, mid, low; or front, central, back) exceeds 60 Hz, they are perceived as distinct vowels. If the distance is less than 60 Hz, their perception is confusing. The vowels with a 41-60 Hz range are perceived as slightly similar vowels, 21-40 Hz as similar, and 0-20 Hz as identical. The degree of similarity influences the level of masking, with higher similarity causing more masking. When contrastive vowels in a language or between languages are similar, they tend to mask each other and impede intelligibility. Conversely, similar vowels in different languages or dialects can enhance intelligibility. The overall level of intelligibility also depends on the relative functional load (RFL), which represents the proportion of words in a language's vocabulary that rely on contrast between two sounds. If two vowels are contrastive but have a low RFL, intelligibility level remains high; however, if their RFL is high, intelligibility level decreases.

3 Level of vowel intelligibility in Nepali English

Vowel intelligibility problems in NE arise because of two factors. The first factor is the number difference of vowel phonemes in Nepali and English. Nepali has six basic vowels, whereas English has twelve (may vary in different varieties). Because of mother tongue interference, the NE speakers, if they are not trained in the English vowel system, try to map the twelve English vowels to the six Nepali vowels. As a result, the acoustically close English vowels cluster together in the acoustic vowel space. When the contrastive vowels are acoustically very close in a language or variety in itself, they mask each other. According to Koffi (2021), pairs of vowels that belong to a natural group, like front, back, central, high, or low, produced by the same speaker or group of speakers are acoustically so close that it is hard to perceive their distinctiveness. This is called

internal masking. The second factor is that the contrastive vowels in English are acoustically very close to the non-corresponding contrastive vowels in NE and vice versa; as a result, the distinct words produced by the speakers of English are perceived by the speakers of NE as the same word and vice versa. This condition is called external masking (Koffi, 2021).

3.1 The level of internal masking

Labov et al. (2013) state that the F1 difference between two vowels is a crucial acoustic criterion for differentiating between perceptually similar and different vowels. If the F1 difference of two vowels is less than 60 Hz, they are perceptually similar vowels, and if the difference is greater than 60 Hz, they are perceptually distinct vowels. Based on this criterion, if the corresponding meaning-distinguishing vowels have an F1 difference greater than 60 Hz, they are perceived as distinct vowels and cause intelligibility problems. According to Kiffi (2021), the F1 distance of 0-20 Hz indicates identical vowels, 21-40 Hz suggests very similar vowels, 41-60 Hz indicates similar vowels, and 61 Hz and above denotes distinct vowels. The assessment of the F1 distance between and among the NE vowels that belong to the same natural class is presented in Table 1. The shaded cells indicate the pairs with internal masking.

Table 1: F1 distance of the NE vowels

		Female speakers' F1 comparison											
Male speakers' F1 comparison		i	ɪ	e	æ	ɑ	ʌ	ə	ɝ	ɔ	o	ʊ	u
	i	-	19	137	414	507	341	308	281	332	105	84	51
	ɪ	4	-	156	433	526	360	327	300	351	124	103	70
	e	121	125	-	277	370	204	171	144	195	32	53	86
	æ	365	369	244	-	93	73	106	133	82	309	330	363
	ɑ	391	395	270	26	-	166	199	226	175	402	423	456
	ʌ	325	329	204	40	66	-	33	60	9	236	257	290
	ə	303	307	182	62	88	22	-	27	24	203	224	257
	ɝ	221	225	100	144	170	104	82	-	51	176	197	230
	ɔ	263	267	142	102	128	62	40	42	-	222	248	281
	o	35	39	86	330	356	290	268	186	228	-	21	54
	ʊ	50	54	71	315	341	275	253	171	213	15	-	33
	u	30	34	91	335	361	295	273	191	233	5	20	-

The analysis of the state of internal masking is presented in Table 2.

Table 2: Internal masking in NE

Pairs	M	Range	Mask.	F	Range	Mask.	RFL%
i-ɪ	4	< 20	Com.	19	< 20	Com.	95
ɑ-æ	26	21-40	Mod.	-	-	-	76
ʌ-æ	40	21-40	Mod.	-	-	-	68
ʌ-ɝ	-	-	-	60	41-60		nill
ʌ-ɔ[1]	-	-	-	9	< 20	Com.	65
ɔ-ɝ	42	41-60	Sligh.	51	41-60	Sligh.	nill
ʊ-o	15	< 20	Com.	21	< 20	Com.	12
u-o	5	< 20	Com.	54	41-60	Sligh	51
u-ʊ	5	< 20	Com.	20	< 20	Com.	7

Note. M=Male, F=Female, Mask=Masking, Com.=Complete, Mod.=Moderate, Sligh.=Slight

In GAE, [ə] does not have contrast with any vowel because it is taken as the weak form of any vowel. The mutual intelligibility problem occurs when NE speakers do not produce the weak form of the vowels, and GAE speakers use the weak form of the vowels in unstressed positions. Similarly, /ɝ/ occurs only preceding the rhotic consonant, so it is in complementary distribution with the other vowels. As presented in Table 2, the most problematic pairs for intelligibility in NE are [i-ɪ], [o-ʊ], [o-u], and [ʊ-u] in the male speech and [i-ɪ], [o-ʊ], [ɔ-ʌ], and [ʊ-u] in the female speech.

Koffie (2021) states that the level of intelligibility also depends on the RFL of the pair of vowels. Based on the RFL, [i-ɪ], [ɔ-ʌ], and [u-o] are the most problematic pairs. [æ] is produced with a high glide [j] followed by [ɑ]; therefore, the intelligibility problem between [æ] and [ɑ] is not because of vowel masking but because of the different quality.

3.2 The level of external masking

[1] Though data does not show, the [ɔ-ʌ] pair is most probably a problematic pair in male speech, too, because [ɔ] and [ʌ] are allophones in Nepali.

The F1 distance between the vowels in NE and GAE is presented in Table 3. The shaded cells indicate the pairs with external masking.

Table 3: F1 distance between NE and GAE vowels

		Nepali English vowels (Male)										
General American English vowels (Male)		i	ɪ	e	æ	ɑ	ɔ	o	ʊ	u	ʌ	ɝ
	i	**19**	15	140	384	410	282	54	69	49	344	240
	ɪ	66	**70**	55	299	325	197	31	16	36	259	155
	e	115	119	**6**	250	276	148	80	65	85	210	106
	æ	227	231	106	**138**	164	36	192	177	197	98	6
	ɑ	407	411	286	42	**16**	144	372	357	377	82	186
	ɔ	291	295	170	74	100	**28**	256	241	261	34	70
	o	136	140	15	229	255	127	**101**	86	106	189	85
	ʊ	108	112	13	257	283	155	73	**58**	78	217	113
	u	17	21	104	348	374	246	18	33	**13**	308	204
	ʌ	262	266	141	103	129	1	227	212	232	**63**	41
	ɝ	113	117	8	252	278	150	78	63	83	212	**108**
		Nepali English vowels (Female)										
		i	ɪ	e	æ	ɑ	ɔ	o	ʊ	u	ʌ	ɝ
General American English vowels (Female)	i	**56**	75	81	358	451	276	49	28	5	285	225
	ɪ	102	**121**	35	312	405	230	3	18	51	239	179
	e	155	174	**18**	259	352	177	50	71	104	186	126
	æ	288	307	151	**126**	219	44	183	204	237	53	7
	ɑ	555	574	418	141	**48**	223	450	471	504	214	274
	ɔ	400	419	263	14	107	**68**	295	316	349	59	119
	o	174	193	37	240	333	158	**69**	90	123	167	107
	ʊ	138	157	1	276	369	194	33	**54**	87	203	143
	u	78	97	59	336	429	254	27	6	**27**	263	203
	ʌ	372	391	235	42	135	40	267	288	321	**31**	91
	ɝ	142	161	5	272	365	190	37	58	91	199	**139**

There is external masking between [ɪ-e], [æ-ɝ], [ɔ-ʌ], [u-o], and [u-ʊ] in both male and female speech. External masking is between [ʌ-ɝ], [ɑ-æ] in male speech and [æ-ʌ], [ɔ-æ], [ɔ-ʌ], [ʊ-o], [ɝ-o], and [ɝ-ʊ] in female speech. However, the masked pairs absent in male speech but present in female speech are also possible masked pairs in female speech, and vice versa. The main problematic pairs based on RFL are [ɪ-e], [ɔ-ʌ], [u-o], [æ-ɑ], and [æ-ʌ]. The problem of [æ] is its articulation as a glide that begins with [j] and glides to [ɑ] or [ʌ].

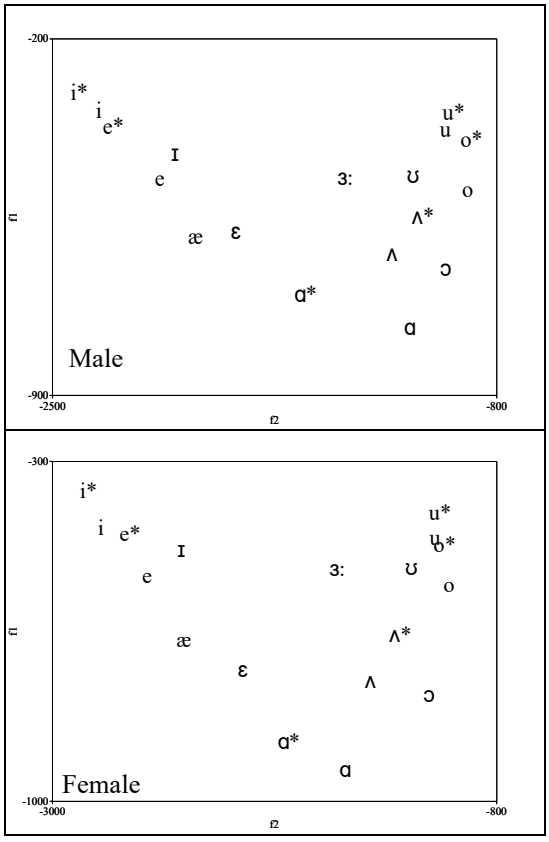
3.3 F1 distance between corresponding vowels

Apart from internal and external masking, the level of acoustic proximity between the corresponding vowels in GAE and NE is equally important for mutual intelligibility. If the corresponding vowels are acoustically the same or very proximate to each other, the level of mutual intelligibility is high.

Comparison of the positions of Nepali vowels (Chalise, 2022) and GAE vowels (Hillenbrand et al., 1995) in the acoustic vowel space is presented in Figure 1. It shows how far the vowels are in the acoustic vowel space. The positional distance suggests the difference in vowel quality, and this difference determines the level of intelligibility. In both male and female articulation, the positions of the corner vowels /i/, /u/, and /a/ are close to one another and are not problematic for intelligibility. But there are several other vowels close to one another, suggesting that they cause intelligibility problems.

In Table 3, the bold-faced figures along the diagonal line present the F1 distance between the corresponding GAE vowels and NE vowels. Based on the F1 differences, the most distinct vowels in GAE and NE are [æ] and [ɝ], and the most similar vowels are [i], [e], [ɑ], and [u]. [ɪ], [o], and [ʊ] are moderately distinct. [ɔ] is similar in male speech but different in female speech, and [ʌ] is similar in female speech but different in male speech.

Figure 1: Nepali and GAE vowels compared



Note. The vowels with * are Nepali vowels

4. Achieving intelligibility

Section 3 presents that there is a remarkable level of intelligibility gap between NE and GAE vowels. This gap was in the speech of the university-level English teachers and graduate-level English students. A higher level of intelligibility gap is expected in the speech of people from other disciplines and areas. The increasing use of English in Nepal in different crucial domains of language use, like education, business, mass media, technology, etc., demands a higher level of mutual intelligibility. In addition, the increasing arena of tourism, foreign trade, foreign employment, and abroad studies needs effective communication in English. This expanding use of English at home and abroad demands better oral communicative skills. Therefore, it is important to minimize the possible intelligibility gap in the English spoken by the speakers of NE. As mentioned earlier, the achievement of better speaking competence through reading literature and learning grammar is not possible. Practice-based pronunciation teaching in the classrooms can be the best solution.

4.1 Rationales for pronunciation teaching

Academic evidence-based discoveries can help improve the quality of pronunciation training in actual classrooms, and there should be a bridge linking pronunciation research and pronunciation teaching in the classroom (Levis, 2018) because the level of intelligibility depends on how the phonological content of the speaker is recognized by the listener (Browne & Fulcher, 2017). For this purpose, the learner should activate both the analytical and motor skills. To develop analytical skills, the learner needs to identify and discriminate the phonetic/phonological features from the segment level to the prosodic level in the L1 and L2. Similarly, to develop motor skills, the learner has to develop the capacity to articulate the phonetic features. Teaching pronunciation is the only means for all the achievements.

4.2 Techniques of pronunciation teaching

The level of intelligibility depends on how well the listener recognizes the speaker's phonological content (Browne & Fulcher, 2017). Therefore, activating both speech perception and production is necessary to achieve clear pronunciation. According to Strange and Shafer (2008), when perceiving an acoustic signal, learners can interpret it correctly only if they have developed mental representations of the target language (TL) sound categories. Thus, their analytical skills should focus on identifying, comparing, and contrasting L1 and L2 sound systems to create new mental representations. Speech production requires not only analytical but also motor skills. Even with good perception and understanding of

L2 input, a learner's TL pronunciation is often imperfect. This can be partly explained by the articulatory habits from L1 that influence L2 phonetic and phonological patterns, and when the patterns from the L1 and L2 do not match. A learner uses an L1 pattern to pronounce an L2 sentence (Zsiga, 2013). In such cases, motor skills must be activated to reduce the mismatch. This indicates that phonetic and phonological research inputs are essential for effective L2 pronunciation instruction.

Pronunciation teaching-learning strategies should be based on the mechanism of achieving intelligible pronunciation. According to Szyszka (2023), there are several stages of acquiring intelligible pronunciation, which are crucial for pronunciation teaching. The first stage is comparison and identification for developing analytical skills. The second stage is production and comparison for developing motor skills. The third stage is to enhance memory function, focusing on remembering or recalling phonetic aspects of an L2, e.g., recalling a teacher's pronunciation. The fourth stage is enhancing the cognitive function by entailing processing phonetic information, e.g., practicing individual sounds first in isolation and then in words. The fifth stage is the development of metacognitive function by planning, organizing, monitoring, and evaluating pronunciation learning, e.g., monitoring one's pronunciation. The sixth stage involves developing social function by promoting cooperation with others, such as teaching L2 pronunciation aspects to others. The seventh stage involves improving affective function by regulating emotions in pronunciation learning, such as developing a sense of humor about mispronunciations. For each of the stages, phonetic and phonological research inputs and technology play a vital role.

5. Conclusion

Pronunciation intelligibility is the most important aspect of oral communication because identification of speech is the initial gate of interpersonal communication. This research investigates that there is a remarkable level of intelligibility gap between NE and GAE speakers. A large body of research has pointed out the benefits of implementing the findings from empirical phonetic/phonological research in pronunciation teaching. The first point to note is that pronunciation teaching is an abandoned field in English classrooms in Nepal. The second point to note is that the tradition of lab-based empirical research and implementation of the findings in classroom teaching is almost nil. Levis (2018) emphasizes that there should be a bridge connecting pronunciation research and pronunciation teaching since academic evidence-based findings can contribute to enhancing the quality of pronunciation instruction in real classrooms.

Finally, teaching pronunciation is a must in English classrooms from the beginning. The schools and universities that teach English as a subject or language need a language lab to enhance the lab-based empirical research and train the students to improve their English language skills. Application of empirical research in the classroom teaching-learning process and rigorous training on listening, speaking, reading, and writing will enhance the language skills of the learners. This technique will remarkably reduce the duration of English teaching and learning to achieve a desired level of competence in the learners.

References

Browne, K., & Fulcher, G. (2017). Pronunciation and intelligibility in assessing spoken fluency. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 37–53). Multilingual Matters.

Chalise, K. P. (2022). *Acoustic analysis of the Nepali speech sounds* [Unpublished doctoral dissertation]. Tribhuvan University.

Chalise, K. P. (2025). *Intelligibility analysis of Nepali English based on the acoustic and durational properties* [A draft research report]. Research Directorate, Rector's office, Tribhuvan University, Nepal.

Derwing, T.M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research.* John Benjamins.

Hillenbrand, J., Getty, L., Clark, M., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5), 3099–3111.

Hurtado, L., & Estrada, C. (2010). Factors influencing the second language acquisition of Spanish vibrants. *Modern Language Journal*, *94*, 74-86.

Koffi, E. (2021). *Relevant acoustic phonetics of L2 English: Focus on intelligibility*. Taylor & Francis Group.

Labov, W., Rosenfelder, I., & Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language, 89*(1*)*, 30-65.

Levis, J. M. (2018). *Intelligibility: Oral communication and the teaching of pronunciation.* Cambridge University Press.

Macdonald, S. (2002). Pronunciation views and practices of reluctant teachers. *Prospect, 17*(3), 3-18.

Sardegna, V. G., & Jarosz, A. (2023). Introduction. In V. G. Sardegna & A. Jarosz (Eds.), *English pronunciation teaching: Theory, practice and research findings* (pp. 3-7). Multilingual Matters.

Shaw, W. D. (1981). Asian student attitudes towards English. In L. E. Smith (Ed.), *English for cross-cultural communication* (pp. 108-122). St. Martin's Press.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. H. Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153-191), John Benjamins.

Szyszka, M. (2023). Intelligibility and situated pronunciation learning strategies. In V. G. Sardegna & A. Jarosz (Eds.), *English pronunciation teaching: Theory, practice and research findings* (pp. 98-114). Multilingual Matters.

Zsiga, E. C. (2013). *The sounds of language: An introduction to phonetics and phonology*. Wiley-Blackwell.