

Vietnamese is spoken by about 86 million people, which makes it the 17<sup>th</sup> largest language community in the world. It is part of the Austro-Asiatic language family, and is by far the most widely spoken of these languages. It has borrowed a large portion of its vocabulary from Chinese, thanks to an early period of Chinese domination, but it is otherwise linguistically unrelated. It is generally described as having three dialects: Hanoi in the North, Ho Chi Minh in the South, and Hue in the center. The three dialects are mostly mutually intelligible, though Hue is said to be difficult for speakers of the other two dialects to understand.

The northern speech...is marked by sharpness, or choppiness, with greater attention to the precise distinction of tones. The southern speech, in addition to certain uniform differences from northern speech in the pronunciation of consonants, does not distinguish between the *hoi* and *nga* tones; and, it is felt by some to sound more laconic and musical. The speech of the Center, on the other hand, is often described as being heavy because of its emphasis on low tones.<sup>1</sup>

Vietnamese uses the Latin alphabet, with additional diacritics to indicate tones.

Vietnam itself is the world's 13<sup>th</sup> most populous country, and the 8<sup>th</sup> most populous in Asia. It became independent from Imperial China in 938 BCE. Since 2000, it has been one of the fastest-growing economies in the world.

### Oral Posture

Oral or vocal tract posture is the characteristic pattern of muscular engagement and relaxation inherent to a given language or accent. It is, if you will, the “home base” for the articulators—the set of actions they favor and the shape to which they tend to return. Together with prosody—rhythm, stress, inflection—oral posture is what gives an accent or language its overall “feel.”

There is some jaw protrusion in Vietnamese speakers, and the jaw position is quite high. The back and root of the tongue are held nearly rigid, with the root occasionally retracted. The tongue favors a cupping position in the middle of the mouth [ɐ], but tends to bunch somewhat along the midline, when forming other vowels.

---

<sup>1</sup> <http://www.vietspring.org/language/intro.html>

The most dramatic aspect of Vietnamese oral posture, for an English speaker, is just how friendly Vietnamese speakers are with their glottises. In Vietnamese:

- All final stop-plosives [p, t, k] are glottally-reinforced [ʔp̚, ʔt̚, ʔk̚]; that is to say, they are accompanied by a simultaneous closure of the vocal folds (glottal stop)
- All vowel-initial words begin with a glottal stop [ʔ], regardless of whether or not they are following a pause.
- There are two glottalized stops [ʔb, ʔd], usually realized as voiced implosives.
- Two of the six tones that serve to distinguish words from one another are glottalized or creaky-voiced, and therefore involve a stop or compression at the vocal folds.

All of this extensive use of glottal closure and compression means that Vietnamese speakers are ready to glottally-stop, reinforce, or compress at any instant. A great deal of this glottal focus carries over into English.

### Prosody

Prosody, or intonation, is the music of a language or accent. It comprises rhythm, stress, and pitch movement (inflection). It is an essential part of accent work—if accurate sounds and oral posture are grafted onto an underlying prosody that is unchanged, it won't sound like the target accent.

Vietnamese is an “isolating” language. The vast majority of words are one syllable long, and they tend not to change form much for cases, tenses or grammatical agreement. This means that Vietnamese speakers tend to break English speech up into syllable units, often separated by glottal stops. The breath pattern is thus very different from that of native English speakers.

Vietnamese is a tone language. It has six tones, which are part of the lexical identity of a given word. To make things even more complex, the tones are not purely pitch contours, but include variable glottal settings. So [mɛ̚-ɰ] is distinguished from [mɛ̚-ɰʔ] not just by the fact that the latter starts and finishes slightly higher in pitch, but also by the fact that the former is breathy and the latter is glottalized (& often uses creaky voice). Both of these phonation types (glottalized or creaky and breathy) are apparent in Vietnamese English.

The six tones make Vietnamese sound very musical, if staccato. This too may carry over into speaking English. Vietnamese English tends to be very broken up, with short intonation phrases, an abundance of stresses, recognizably Vietnamese tones, and a great many glottal

stops. Another difference between Vietnamese and English is that Vietnamese is a syllable-timed language, whereas English is a stress-timed language. In a syllable-timed language (Spanish is another example), all the syllables take roughly the same amount of time to say. In a stress-timed language, syllables lengthen and reduce according to whether or not they are stressed. It is very hard for speakers of syllable-timed languages to get the hang of stress-timing, so this feature will persist, strengthening the staccato rhythm of Vietnamese English

### Characteristic sounds

#### Consonants

##### 1. Glottal reinforcement and voiced implosives

- /b, d/ → [ʔb, ʔd] The voiced stops [b] and [d] are always glottalized in Vietnamese (that is to say, they are always reinforced by a simultaneous closing of the vocal folds). Most of the time, the stops themselves are voiced implosives [ɓ, ɗ]. A voiced implosive is produced by moving the larynx downward over a static column of air, setting the vocal folds into vibration. A stop further forward in the mouth is released an instant later. The lower air pressure in the oral cavity created by the lowering of the larynx causes air to rush into the mouth as the stop is released. It is worth trying to perfect these sounds, as Vietnamese speakers will often realize English /b/ and /d/ as voiced implosives [ʔb] and [ʔd]. Even when they do not, they will be pre-glottalized [ʔb] and [ʔd]. This is hard-wired into Vietnamese phonotactics, and is very hard for the Vietnamese speaker to change.

*Big, difficult, back, desk, David, Bruce, doughnut, basic, bold, daisy.*  
 “The big dog bent down to dig a big black dinosaur bone from a ditch.”

- Final /p, t, k/ → [ʔp̚, ʔt̚, ʔk̚]. Word-final unvoiced stop-plosives are glottally-reinforced and unreleased in Vietnamese. This carries over into English.

*Stop, thick, cat, fought, bit, lip, black, top, sick, moat, trap, stock.*  
 “The fat black cat stuck a bit of crack rock in the moat.”

- ##### 2. Word-final /b, d, g, v, z, dʒ/ → [ʔb̚, ʔd̚, ʔk̚, f, s, tʃ]. Vietnamese phonology doesn't allow final voiced stop-plosives or fricatives (though it does have final voiced nasals), so these have a strong tendency to devoice in English.

*Bad, love, stove, dig, plague, bridge, stag, bead, league, badge, rose.*

3. Stressed [p<sup>h</sup>, k<sup>h</sup>] → [p<sup>̄</sup>, k<sup>̄</sup>]. Again following Vietnamese phonotactics, /p/ and /k/ will tend to be unaspirated, even when at the beginning of stressed syllables.

*Past, cat, pick, caught, potential, pancake, kitten, cluster, pitiful*  
“Patrick’s cute kitten caught a piebald possum in the kitchen.”

4. /g/ → [ɣ ~ x]. Vietnamese lacks a /g/ phoneme (though many speakers have it as an allophone), but has a voiced velar fricative [ɣ], which may substitute for English /g/. Word-finally, however, [g] is extremely unlikely. Because of word-final devoicing, unvoiced [x] is likely.

*Brag, bagel, fog, big, saga, tug, flag, mug, Legos, pig, league, fig, sag, rogue.*  
“Maggie’s sagging mug made her feel distinctly piggy.”

5. Consonant cluster reduction. Vietnamese never has more than one final consonant. Vietnamese speakers will have a strong tendency to reduce English final consonant clusters to a single consonant

*Lasts, fists, locks, bolts, closed, lengths, thanked, strikes, beguiled.*  
“Amidst the mists and coldest frosts, with stoutest wrists and loudest boasts...”

6. /r/ → [ɻ]. /r/ (in all positions) in Vietnamese English is often realized as a complex articulation that involves the tip, body, and root of the tongue. The tip is engaged in a retroflex approximant, the sides of the tongue are touching the upper molars, and the tongue root is retracted. The result is a strongly rhotic sound, and a very muscular articulation. It can have a strong effect on the sounds around it. (The symbol I’ve used is strictly unofficial IPA—the strikethrough is sometimes used to symbolize the velar/palatal approximation of the American “molar” /r/, for which there is no other existing IPA symbol.)

“Around the rough and rugged rocks the ragged rascal ran.”  
“Rosie, Rachel, and Arthur reached a record deal with Arista.”

7. The dental fricatives /θ/ and /ð/ are most likely to become [s] and [z] when they are prominent or care is being taken. When the speaker is speaking quickly, [t] and [d] are probably more likely.

"The theme is weak and, therefore, this theater piece is without worth."

8. Coronal consonants (those made with the tongue tip or blade), are a complicated story in Vietnamese. Some are made on dentally [t<sup>h</sup>, ɲ, ʃ, ʒ]. Some are made laminally (using the blade of the tongue, generally with the tip down behind the bottom teeth) [ʃ, ʒ, ʧ]. In Southern (Ho Chi Minh City) Vietnamese, there are two retroflex consonants, with the tongue-tip curling back [ʂ, ʐ]. This is quite a lot of complexity. Listen closely to your chosen source, and let oral posture be your guide.

## Vowels

1. **TRAP/BATH**→[ɐ], **STRUT**→[ɐ ~ ɜ]. Vietnamese has one open, central vowel, which is mapped onto all three of these lexical sets. This vowel is essential enough to the oral posture of the accent that we might call it a *focal sound*: a sound which is integral to the feel of the accent and can therefore help you find the oral posture. Note that **STRUT** words spelled with an 'o' (*love, money, son, some, front*) might acquire some rounding and be realized more like [ɜ].

*Back, have, can, cab, thank, path, master, hand, slander, tax, laugh, abbey, batch, rag, mass, half, command, scratch, plant, bag, answer, fast, dance, mad.*

"The dancing master, a bad man, had a bag half-filled with rats and scratched bananas."

*Cut, fun, some, flood, stuck, fussy, sun, bust, come, stuff, love, mull, smother, lunch, butter, cluck, trust, son, wonder, flung, young, mulch, blunder, front.*

"The sultan's unruly son, an ungrateful runt from day one, strutted his sulky stuff under the sultry sun while his wonderful mother began to come undone."  
"Neither blood nor money could frustrate her love for her country."

2. PRICE→[p̥ɜː]. This vowel is very similar to the TRAP/BATH/STRUT realization, just a bit further forward and generally longer. It starts a bit further back in the mouth than most Americans' PRICE vowel. It is smoothed to a monophthong.

*Ripe, life, rise, time, pint, island, write, knife, fight, sign, type, ninth, Christ, reprisal, trifle, bible, mines, hives, fives, finer, bite, title, gibe, type, lies, ice, lithe, wily, pikes, bind, vile, buy, like, child, sigh, pie, high, kind, shine, tide, lime, dive, python, piling, try, aisle, why, mind, dye, miles, scythed, primal.*

“Mike shanghaied the high-priced supercollider in a primal drive to survive – and even thrive – by means of this vile crime.”

“The python's slimy thighs broke out in hives every time he lied.”

“Try as he might, Michael couldn't find a silent hybrid to use in the crime.”

3. GOAT→[ɔ̟ʊ ~ ɛ̟ø]; GOAL→[äw]. The GOAT vowel is most often a diphthong with a slightly raised [ɔ̟] onset. When strongly emphasized, it has the potential to turn into a very interesting vowel indeed: [ɛ̟ø], where the tongue stays put in at a close-mid vowel between front and central and the lips round throughout the articulation.

When followed by /l/, as in GOAL, the onset will often be more open and the /l/ will usually be vocalized (absorbed into the vowel, effectively—the tongue tip stays down while the back of the tongue makes a sort of vowel/approximant /l/).

*Row, boat, Joe, cold, oak, mauve, soldier, shoulder, so, though, beau, gauche, oaf, cloves, yokel, ocean, rose, close, tone, mole, rogue, stove, loaf, robe, moan.*

“The lonely ponies jolt slowly homeward through the cold, frozen, snow.”

“The golden mole rolls a round roll to his bolt-hole in the mold.”

“The holy ghost holds that the souls of rogues don't roam alone.”

4. NURSE→[ɚ̥ ~ ɚ̥̟]. The NURSE vowel will be realized as a close-mid vowel somewhere between central and fully back, with or without some degree of lip-rounding.

*Church, firm, bird, earth, earnest, worst, murder, worthy, journal, attorney, colonel.*

“The early bird catches the worm, or so I've heard.”

“I've heard worse curses from stern circus girls, Bernie.”

5. KIT→[i]. Like many languages, Vietnamese has only one close front vowel. It is difficult, therefore, for Vietnamese speakers to distinguish between KIT and FLEECE. They will generally use [i] for both.

*Ship, hint, drink, winter, rhythm, English, women, guilt, stiff, dig, sing, pith, live, fig, pink, flick, still, tin, dish, mist, whiff, thing, fit, limb, pin, nip.*

"The stiff English women became sick on the listing ship."

"This drinking has filled his skin to the brim with silly spirit."

6. FOOT→[u]. Similarly, Vietnamese has only one close rounded back vowel, so it is used for both FOOT and GOOSE sets.

*Bullet, bushes, crooked, hood, push, book, should, fully, wood, put, shook, pulpit, sugar, could, shook, foot, good, pull, butcher, wolf, woman, cushion.*

"The butcher shouldn't put his cooked pudding on the cushion."

"Should you bully and push the butcher when he's cooking over wood?"

7. THOUGHT/CLOTH→[ɔ̌], TAUT/WASH→[ɑ̌].

NORTH/FORCE→[ɔ̌ɣ̌], WARMTH→[ɑ̌ɣ̌]

This looks exceedingly complex, but isn't really. There is a split in the THOUGHT/CLOTH sets and in the NORTH/FORCE sets. If a word in one of these sets is spelled with an 'o', the vowel will be a raised, tongue-root retracted [ɔ̌]. If the word is spelled with an 'a', the vowel will be a tongue-root retracted [ɑ̌]. In the case of the NORTH/FORCE/WARMTH words, the /r/ that follows the vowel will be the braced, retroflex, tongue-root retracted approximant [ɣ̌] described in the consonant section above.

THOUGHT/TAUT & CLOTH/WASH mixed

*Awe, saw, gnaw, law, cause, paws, lawn, pawn, paul, fought, bought, caught, sought, taught, naught, wrought, lost, cost, frost, song, long, wrong, soft, coughed, moss, dross, all, tall, call, stall, pall, fall, brawl, crawl, trawl, shawl, cloth, broth.*

"Though often called a costly, gaudy, awful loss, Broadway always offers its all."<sup>2</sup>

---

<sup>2</sup> Thanks to Dudley Knight for this sentence.

**NORTH/FORCE & WARMTH mixed**

*Four, war, before, scorn, warren, swarm, lore, pore, pour, or, accord, orb, George, warp, ornament, warn, fortress, swarthy, quarter, warble, reform, flora, resource, mourn, roar, door, Warsaw, wart.*

"Portia's orchestral warbling, pouring forth from her torso, was warm and gorgeous."

7. **INTO**→[ü]. The unstressed version of the **GOOSE** vowel is somewhat responsive to English vowel reduction. Instead of reducing towards a [ə], however, it mostly seems to be affected by losing its lip-rounding. The resulting vowel [ü] is a vowel native to Vietnamese, but very different from any heard from native English speakers. It is thus quite striking. The most common instance of this vowel is in the word *to*.

"To go to or not to go; to be turned into frogs or to be turned into logs..."



## Standard Lexical Sets

Invented and described by phonetician John Wells in his book *Accents of English*, lexical sets are an essential tool for accent description. Originally devised according to the vowel inventories of two "standard" reference accents, RP and "General American" (a mythical beast), the standard sets are as follows:

KEYWORD	RP	"GenAm"	other examples
1. KIT	ɪ	ɪ	ship, sick, bridge, milk, myth, busy, women..
2. DRESS	ɛ	ɛ	step, neck, edge, shelf, friend, ready..
3. TRAP	æ	æ	tap, back, badge, scalp, hand, cancel..
4. LOT	ɒ	ɑ	stop, sock, dodge, romp, possible, quality..
5. STRUT	ʌ	ä	cup, suck, budge, pulse, trunk, blood...
6. FOOT	ʊ	ʊ	put, bush, full, good, look, wolf..
7. BATH	ɑ:	æ	staff, brass, dance, ask, sample, calf..
8. CLOTH	ɒ	ɔ	cough, broth, cross, long, Boston..
9. NURSE	ɜ:	ɝ	hurt, lurk, urge, burst, jerk, term..
10. FLEECE	i:	i	creep, speak, leave, feel, key, people...
11. FACE	eɪ	eɪ	tape, cake, raid, veil, steak, day..
12. PALM*	ɑ:	ɑ:	psalm, father, bra, spa, lager..
13. THOUGHT	ɔ:	ɔ	taught, sauce, hawk, jaw, broad, caught..
14. GOAT	əʊ	oʊ	soap, joke, home, know, so, roll...
15. GOOSE	u:	u	loop, shoot, tomb, mute, huge, view..
16. PRICE	aɪ	aɪ	ripe, write, arrive, high, try, buy...
17. CHOICE	ɔɪ	ɔɪ	adroit, noise, join, toy, royal...
18. MOUTH	aʊ	aʊ	out, house, loud, count, crowd, cow...
19. NEAR	ɪə	ɪɹ	beer, sincere, beard, fear, serum...
20. SQUARE	ɛə	ɛɹ	care, fair, pear, where, scarce, vary...
21. START	ɑ:	ɑɹ	far, sharp, bark, carve, farm, heart...
22. NORTH	ɔ:	ɔɹ	for, war, short, scorch, born, warm...
23. FORCE	ɔ:	ɔɹ	four, wore, sport, porch, borne, story...
24. CURE	ʊə	ɔɹ ~ ʊɹ	poor, tourist, pure, plural, jury...
25. HAPPY	ɪ	ɪ	city, party, money, century, coffee, taxi, calorie...
25. LETTER	ə	ə	mother, brother, daughter, scissors, another, quitter...
26. COMMA	ə	ɝ	sofa, Africa, Alaska,

\*a better choice of keyword for this set might be **SPA**, as some speakers articulate the /l/ in *palm* with a velarized or "dark" [ɫ] (with or without raising the tongue-tip to the alveolar ridge), resulting in a different vowel.