Jasmine and the Bee: Spelling word-initial \[t^h\] in English

Gerald R. McMenamin  
California State University, Fresno

Lindsay N. Kerr  
Georgetown University

Jasmine was an American seventh grader whose first language (L1) is Punjabi and second language (L2) is English. A good speller, she tied for first place at her regional spelling bee. She maintains that she then correctly spelled the tie-breaking word: tomography. After requesting an instant video replay of her word, the judges declared Jasmine’s spelling incorrect: \[d-o-m-o-g-r-a-p-h-y\].

We examine three questions: Is Jasmine’s presumed-high word-recognition ability consistent with such an error? Did her L1 interfere with her L2 pronunciation? Did an L1-influenced pronunciation cause the judges to perceive her word-initial /\[t^h\]/ as [\[d\]]? After inquiring of Jasmine about the spelling event itself, we examined the considerable research on Indian English, spelling, speech production, and speech perception; we also audio-recorded Jasmine’s spoken language at various style levels to assess the degree to which Punjabi may be influencing her English pronunciation. The first question remains without a confirmed answer, though our assumption is that it is ‘no’. Work on Indian English indicates a probable ‘yes’ to the latter two questions. Spectrographic analysis of Jasmine’s word-initial English \[t^h\]-sounds demonstrates the shorter voice onset times associated with the nonaspirated /\[t^h\]/-sounds of Punjabi and Indian English. It is probable that Jasmine pronounced the ‘t’ \[t^h\] of tomography with little or no aspiration, and that the judges heard the near-unaspirated [\[t\]] as a [\[d\]]. Jasmine’s experience demonstrates that the Scripps National Spelling Bee’s sole reliance on dictionary pronunciations of Standard American English might well be adjusted to accommodate contestants who speak a nonstandard, or a standard but non-American, variety of English.*

Keywords: Scripps National Spelling Bee, dialect and spelling, Indian English, aspiration, voice onset time, voiceless stops

1. **Introduction.** This is the story of a spelling bee participant, whom we pseudonymously call Jasmine, and her encounter with a policy of the Scripps National Spelling Bee (SNSB) that permits only standard word and letter-name pronunciations as specified by the dictionary. Jasmine spelled the word-initial letter <\[t\]> of ‘tomography’ with what was very likely an L1-influenced initial [\[t\]]-sound, which the judges interpreted as a mistaken letter <\[d\]>. Jasmine was thereby eliminated from the regional competition due to a pronunciation not found in Webster’s Third New International Dictionary, which the SNSB rules prescribe as ‘the primary source used to craft the pronunciations’ for the Bee (Scripps National Spelling Bee 2013), thereby requiring standard pronunciations for the English words and letter-names used in the contest.

2. **The American spelling bee.** The history of the spelling bee is rooted in American nationalism, that is, a reaction against England following the American Revolution (McArthur 2011:6), and the subsequent intent to help American English take root by selecting and improving regional varieties of English in America:

The cultural practice of spelling bees emerged in the United States within a nexus of social, political, pedagogical and linguistic factors roughly linked to the emergence of American English as a unitary category tied to America’s political identity after the Revolutionary War. (McArthur 2011:1)

In fact, McArthur (ibid.) cites Noah Webster as framing ‘the establishment of a “uniform standard of elegant pronunciation” as a nationalist imperative’. What started out

* The authors gratefully acknowledge the helpful suggestions made by faculty and students of the Fresno State Linguistics Club, particularly Sean Fulop and Chris Golston, and they are especially appreciative of the significant insights afforded by the Language referees.
historically, then, to be community-oriented spelling bees have since become well-organized media events, reaching the national level in the form of the Scripps National Spelling Bee (McArthur 2011:1). The stated learning goal of the SNSB is correct English usage. Scripps describes the Bee as ‘a cultural celebration of the English language’, saying that ‘this event spotlights our mission to inspire children to improve their spelling, increase their vocabularies and develop correct English usage that will help them all their lives’ (Scripps National Spelling Bee 2013). The constraint of correct English usage appears to be the basis for the SNSB’s explicit rule requiring standard pronunciation, which is, without doubt, important to the SNSB and to every participating contestant. Such a requirement, however, presents a significant public-policy challenge, one that transcends Jasmine’s predicament: while the reach of the SNSB into the linguistic lives of millions of students is immense, provision for the linguistic diversity of its many participants is not an integral part of contest protocols.

Scripps reported that the final SNSB competition of 2013 began with more than eleven million students participating in local spelling bees just like the one Jasmine entered (Scripps National Spelling Bee 2013). And, as it has since 1976, the SNSB allowed students from outside the United States to compete, resulting in a highly diverse field of national-championship finalists:

This year, 281 spellers will converge from eight countries: the Bahamas, Canada, China, Ghana, Jamaica, Japan, South Korea and the United States. Competitors hail from all 50 states, the District of Columbia, American Samoa, Guam, Puerto Rico, the U.S. Virgin Islands and Department of Defense Schools in Europe. (Scripps National Spelling Bee 2013)

In addition, the SNSB website reports that it is considering an international competition, an indication, indirect as it is, that it considers diversity in the contestant field to be important. Yet, such manifest interest in diversity is controverted by the circumstances of the Bee itself. First, contest rules specify Webster’s dictionary as the pronunciation authority, thereby disallowing accommodation for regional or social variation. Second, there is very likely de facto adjustment for such variation in pronunciation, depending on where a local or regional bee is held, as well as on the varieties of English spoken by participants and judges. And third, there are many participants whose accents will not be related to regional or social variation in English but are the result of influence from a non-English L1 instead.

Jasmine’s English is influenced by both Indian English and Punjabi. Her case, as reported here, illustrates the potential for linguistic discrimination on the basis of language variation, and is likely indicative of general spelling bee practice, given the SNSB rules. We are hopeful that the SNSB will consider policy changes that would be designed to accommodate the distinct varieties of English that accompany the social, cultural, and geographical diversity of present and future participants. Although increasing contestant diversity is unlikely to present an issue for the SNSB’s new national-level emphasis on morpheme and word recognition, there appears to be a present need for policy change related to regional and social accents of English, whether standard, nonstandard, or influenced by a non-English L1.

3. THE CONTESTANT. Jasmine is an American schoolgirl who was in seventh grade at the time of this competition. She was born, raised, and educated in the United States. Her first language and that of her family is Punjabi, and she communicates within an extended social network of Punjabi speakers (family, friends, religion, and travel) both in the US and on extended family visits to India. Jasmine acquired English as a second language in her school and community and is now bilingual in Punjabi and English. As
it happens, she has taken a continuing interest in spelling; as a sixth grader she began to compete at the school and regional levels of the SNSB. Preparatory to the 2013 national competition, Jasmine won her local school spelling bee and from there moved directly on to compete in a central California regional bee. She was one of the last two finalists when she was asked to spell *tomography*. After she spelled the word, the judges stopped the bee, reviewed the video, and finally declared that Jasmine had misspelled the word by starting it with a `<d>`: `<d-o-m-o-g-r-a-p-h-y>`. Jasmine took second place, which was a real achievement, but by not placing first she lost the opportunity to represent her region at the SNSB finals in Washington, DC.

It is the judges, not the speller, who have the last word, so to speak, at a spelling bee. Yet the panel’s need for an instant video replay of Jasmine’s performance gives a clue to the ensuing problem: Jasmine, in her own quiet but steadfast manner, maintains that she *spelled* *tomography* with an initial `<t>`: `<t-o-m-o-g-r-a-p-h-y>`. While the decision of the spelling bee judges is final, a close examination of their decision in Jasmine’s case may help her in subsequent years and may also inform other contestants who use a variety of English distinct from that of the dictionary or of the judges when pronouncing English sounds and letter-names. It may also serve as a catalyst for discussion of SNSB practices that might be reviewed with an eye to allowing all American students (as well as those from other countries) real and equal access to spelling bee achievement in the United States.

**4. Assumptions and questions.** To focus then on the matter of Jasmine and the regional bee where she was eliminated from competition, we make three reasonable assumptions: (i) the misspelling was not an error of oral performance, that is, not the oral analog of an accidental ‘typo’ in written language; (ii) the speller truly recalls that she spelled the word with an initial `/t/`; and (iii) the judges truly heard the word spelled with an initial `/d/`. In the absence of an audio recording that would support spectrographic analysis (and Jasmine’s parents being reluctant to request the recording of the bee itself), we can only corroborate Jasmine’s pronunciation by examining her postproduction. Then, assuming the judges heard a `[d]` and she likely said `[t]`, we can question why she would have said `[d]` but thought `/t/`. If, however, her postproduction indicates a likelihood of her saying `[d]`, then we could question what acoustic cues were present or absent in her `[d]` production of `/t/`. The problem, then, is this: What could have caused the judges to hear the first letter of her spelling of *tomography* as `[d]`? This issue can be best understood by narrowing the inquiry to three specific issues.

- Would such a practiced speller’s knowledge of English word formation have allowed her to actually misspell the word?
- Could the sound system of her first language have interfered with her pronunciation of the English `/t/-sound, including how she named the letter `<t>` while spelling the word?
- Could the listener-judges’ perception of the `/t/-sound as something different from what the speaker-speller intended have been due to the speller’s accent as influenced by Indian English or by Punjabi?

**5. Indian English.** Responses to the above questions are very likely related to what we hear as the speller’s Punjabi-influenced Indian English accent. Indian English is a recognized and established standard variety of English, accepted as such worldwide (e.g. Kachru 1983, Sailaja 2009, and Sedlatschek 2009). And while Indian English demonstrates its own internal variation within and around South Asia, it is often seen
from afar as a singular variety that contrasts with Standard American English at all linguistic levels, including the production of the vowels, consonants, and suprasegmentals, all of which make up its distinctive accent. The complex variation present in Indian English is due to region, ethnic group, and proficiency: specifically, the influence of Indian languages, the high degree of regional variation in the four major language families of India, and the fact that Indians learn English from Indians, whose English is influenced by the sound systems of local languages (Brozbă 2011:1). And more to the point, voiceless stops tend not to be aspirated (i.e. have short voice onset times) in syllable-initial position (Brozbă 2011:5), and ‘many speakers, under the influences of their first language, still use retroflex [ʈ] and [ɖ] for [t] and [d]’ in words like today, terrific, and demand (Brozbă 2011:6).

An even more relevant point to be made here, however, is that as standardized and accepted as Indian English is around the world, it can become a liability when used in the context of another standard variety such as American English: ‘a distinctive accent may become a detriment to the speaker when it affects the intelligibility of an utterance and may result in a reduction in both social and professional interaction and progress’ (Awan & Stine 2011:998). This is precisely the problem Jasmine faced.

6. Previous work. We examine below the considerable research on word formation and spelling, speech production, and speech perception as they relate to the questions above, all of which present possible reasons for why this speller may have gotten the spelling correct but the answer wrong: the speller’s acquired ability to recognize English morphemes, her Punjabi-influenced production of English /t/, and the judges’ perception of a Punjabi-like English /t/ as something phonetically close to English [d].

6.1. Word and morpheme recognition. Learning to spell in English is not simply a matter of committing long word lists to rote memory. Skilled spellers acquire over time the ability to combine two strategies to successfully spell English words: (i) Phonetics: the sound-letter correspondences for some sounds, and (ii) word recognition: the identification of word parts recognized as meaningful units, that is, morphemes: ‘It is clear that to spell English successfully one must have a high degree of sensitivity to both the phonological and morphological structure of words’ (Brengelman 1970:132).

For example, phone, phonic, and phonetic share a single root morpheme that is realized in each word by one of three different vowels, [o], [a], and [ə].

The underlying competence of a skilled speller, therefore, includes mastery of those sound-letter correspondences that do exist in English, coupled with the ability to identify the morphological structure of words. This morphological competence is especially important for enabling speakers of different social or geographical varieties of English to become good spellers of English. Children who speak different varieties of English (like Indian English) are able to become good spellers precisely because the English spelling system includes representation of some surface features of pronunciation and also because they recognize the morphemes underlying phonological structures, where even the same morpheme is not always pronounced or spelled in the same way. The word in question, tomography, has three morphemes, as seen in Table 1.

<table>
<thead>
<tr>
<th>MORPHEME</th>
<th>MEANING</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>tom(o)</td>
<td>‘slice’</td>
<td>graph</td>
<td>y</td>
<td>[Noun]</td>
</tr>
</tbody>
</table>

Table 1. Morphology of tomography.

One might safely surmise that it is unlikely that a speller at the competence level of Jasmine would not recognize the Greek morpheme tom(o)- (as in atom, anatomy, di-
chotomy, tome, entomology, etc.). Even if she did not happen to recognize tom(o)-, it is also unlikely that she would replace it (in thought) with the candidate morpheme dom(o)- (as in dome, majordomo, or domicile); she likely would instead either immediately reject dom(o)- because it has to do with ‘house’ or ‘home’ and nothing to do with tomography, or would reject dom(o)- because she correctly presumes that domography is simply not an English word. A speller’s word-recognition ability will come into play regardless of how the word is pronounced:

Thus it is possible for the same spelling system to be entirely adequate for dialects which may sound conspicuously different. It does not follow automatically from the fact of dialect divergence that a common spelling system is impossible or that speakers of some dialects will have greater difficulty with spelling than others. (Brengelman 1970:133)

Although it cannot be stated as a fact without further observation, it is likely that, for Jasmine to have reached the level of spelling proficiency needed to be a spelling bee finalist, she had to have a well-developed and already internalized understanding of English word formation, together with the ability to readily apply this knowledge to a word containing two common Greek morphemes. After all, Jasmine ‘attacked’ the spelling of this word without requesting definition or clarification.

6.2. The /t/-sounds of English and Punjabi. The /t/ is a voiceless stop consonant in English and Punjabi, both languages requiring a complete closure of the vocal tract and no vibration of the vocal folds. There are, however, differences between the two languages in point of articulation and aspiration. The single /t/ of English is apical-alveolar and aspirated word-initially: [tʰ]. In contrast, Punjabi has four /t/-sounds: apical-dental /t/, retroflex /ʈ/, aspirated apical-dental /tʰ/, and aspirated retroflex /ʈʰ/.

By way of review, the English /t/ is aspirated, especially when it occurs word-initially before a vowel in a stressed syllable. With an aspirated or ‘plosive’ stop, the vocal tract is momentarily closed; pressure builds up behind the closure and is released with an explosion of air (Laver 1994). The relatively long interval between the release of the stop and the start of the vowel favors simultaneous aspiration in English, which is strong in word-initial stressed syllables.

Punjabi has a series of unaspirated voiced and voiceless stops along five points of articulation: labial, dental, retroflex, palatal, and velar. Four of these are plosives, the palatals being affricates. Since Punjabi lacks the aspirated voiced series of other areal languages, it is then characterized by four triads of plosives, each triad having a voiced unaspirated, voiceless unaspirated, and voiceless aspirated stop, for example, /d/, /t/, /tʰ/. Therefore, voice onset time in Punjabi (like other three-plosives series languages) ranges from prevocalic /d/, to the short delay of /t/, and on to the longer lag of /tʰ/ (Hesselwood & McChrystal 1999).

6.3. A note on voice onset time. Voice onset time (VOT) was first introduced by Lisker and Abramson (1964) as a means of providing a physical realization for differentiating the crosslinguistic articulatory differences of stop consonants. VOT is a measure of the amount of the time it takes for vocal folds to vibrate after a stop has been released, which is done by measuring the time from the release of the stop to the start of the waves produced by glottal vibrations of the vowel that is to follow. The time it takes to produce a stop, as measured by VOT, will differ from language to language. Lisker and Abramson’s study showed mean crosslinguistic differences in VOT for producing language-specific stops. Some languages, such as Korean, have a mean VOT of 94 ms for aspirated /t/, while a language such as Armenian has a mean VOT of only about 60 ms. Unaspirated stops in other languages (e.g. Thai) demonstrate even shorter VOTs. VOT measurements will also differ depending on place of articulation. Cho and Lade-
foged (1999) maintain that VOT will be longer the farther back the closure or the more extended the contact area is, and that the VOT will be shorter if the articulator moves faster. This means that the VOT for English alveolar /t/ may be longer than for the Punjabi dental /t/ or retroflex /ʈ/.

In addition, English stops /b, d, g/ are labeled as voiced, yet voicing may not occur until after the stop has been released if the consonant is word-initial, so it is oftentimes the aspirated component of the English /p, t, k/ that causes English speakers to perceive them as different from /b, d, g/. The VOT, as a measure of aspiration of the voiceless stop /t/, is, of course, longer because aspiration causes a longer delay before periodicity starts.

Cho and Ladefoged (1999) break stops down into three categories: unaspirated, slightly aspirated, and aspirated. Wholly unaspirated stops generally last around 30 ms, and slightly aspirated stops about 50 ms, and note that even stops with slight aspiration may be perceived as voiced. A fully aspirated stop in English can last longer than 90 ms. What this means for the purposes of the present study is that the VOT of an aspirated stop of an English speaker is rarely less than 50 ms, causing the shorter VOTs of a Punjabi-influenced unaspirated production of English /t/ to resemble English /d/.

6.4. VOT and the /t/-sounds of English and Punjabi. Two things then become important for understanding the present issue. First, Punjabi divides the ‘available’ VOT space into thirds, while English divides it in half, thereby increasing the potential for interference between the unaspirated voiceless /t/ of Punjabi and the voiced /d/ of English. Second, the issue is not only the division of onset duration of Punjabi into three parts that affects production and perception of its plosives, but also the point at which aspiration starts on the VOT continuum. In Iverson and Salmons’s (2003) paradigm, English is an ‘aspiration’ language, wherein voiceless stops are heavily aspirated and voiced stops are not voiced at all. In contrast, Spanish, for example, is a ‘voice’ system, wherein voiceless stops are unaspirated and demonstrate a wide range of VOT variation. This is likely what prompts Luchini and Chiusano (2009:58) to observe in contrasting English and Spanish that aspiration in English is ‘an essential feature to achieve intelligibility’. While English and Spanish both divide up the VOT duration into two parts, their respective VOTs for /d/ and /t/ split in different places. Spanish prevoices voiced stops but allows wide-ranging VOT variation in voiceless stops (0–100 ms). English, by contrast, allows more variation in voiced stops but requires a VOT of greater than 40 ms in voiceless stops.

In sum, then, while the significant potential for /t/-/d/ interference in the English of the L1-Spanish speaker likely relates to a singular difference in VOT position for voiceless stops, the potential for interference is even greater in the English of the L1-Punjabi speaker due to two factors: VOT position for voiceless stops and differences in the division of the VOT continuum.

Granted, although this may be a strong theoretical basis for explaining the differences between contestant and judges in Jasmine’s case, it remains at the level of just a possibility for a couple of reasons: it may not happen all the time (e.g. as in Jasmine’s speech), and Punjabi has aspirated voiceless stops also. The /t/-sounds of Punjabi are well described in Karamat 2001 and Heselwood & McChrystal 1999. The differences between the /t/-sounds of the two languages, as described above, are summarized in Table 2.

6.5. Point of articulation of Indian-English /t/. With respect to the dentalized /t/ in Punjabi, it is generally agreed that a Punjabi speaker of English may dentalize /t/ in English (Dhore et al. 2012). With respect to the retroflex /ʈ/ in Punjabi, the speakers
of many Indian languages, including Punjabi, tend to use their own retroflex consonants in place of English alveolars /t, d, n/, apparently because the English alveolar stops are perceived as closer to the retroflex stops than to the dental stops in Indian languages, according to Ball and Müller (2005:289), who also note that ‘[t]his use of retroflex consonants is very characteristic of Indian English, and the retroflex resonance is very pervasive and can cause intelligibility problems’ (ibid.).

6.6. Aspiration of Indian-English /t/. English aspirated /p, t, k/ are often unaspirated in Indian English, as indicated by Malhotra and Vogelaar (2004), Ball and Müller (2005:289), Luchini and Chiusano (2009), and Awan and Stine (2011:999). This absence of aspiration in Indian English is related to differences in VOT, as noted in §6.3 above. Speakers of Indian English produce significantly lower VOTs after /p, t, k/ than speakers of American English. For word-initial /t/, the VOT is ±40 ms in Indian English vs. ±77 ms in American English. The low (< 50 ms) VOT of Indian English causes a weaker release of the consonant and a quicker start of the following vowel, thus resulting in partial or complete absence of aspiration for initial /t/ in Indian English (Awan & Stine 2011:1001). The weaker release of the voiceless stop results then in reduced intelligibility of the voiceless stop consonants:

Because speakers of Indian English may use a non-aspirated form of the voiceless plosive referred to as ‘voiced aspirated’ (Yavaş 2006), the VOT necessary for the appropriate production of Standard English stop plosives may be altered and intelligibility affected, e.g., ‘pin’ may be perceived as ‘bin’ and ‘pet’ may be perceived as ‘bet’. (Awan & Stine 2011:999)

6.7. Voicing of Indian-English /t/. Loss of aspiration and voicing are related insofar as reduced VOT occurs with loss of aspiration in initial voiceless stops, which can cause them to be perceived as voiced. As has been noted, this concurrent change in voicing and degree of aspiration is an important characteristic of Indian English (Jenkins 2000, Awan & Stine 2011:999). Word-initial voiceless stops in some varieties of Indian English, therefore, demonstrate a complex of phonetic characteristics not found in American English: their reduced VOT relates to less aspiration, and the loss of aspiration affects listener perception in that it may cause /p, t, k/ to be perceived as their unaspirated English parallels /b, d, g/:

pay attention to the feature of aspiration of /p/, /t/, /k/ when these sounds occur in initial accented syllables. … [W]ithout the aid of this puff of air, a listener may find it more difficult to identify these sounds as voiceless. Thus, an unaspirated /p/ may be mistaken for /b/, /t/ for /d/, and /k/ for /g/ in word pairs like, pig–big, teen–dean, cap–gap. (Luchini & Chiusano 2009:58)

6.8. Punjabi-influenced production of English /t/. While it is likely that the Punjabi sound system interfered with Jasmine’s production of the English /t/-sound, it does not follow that it had to interfere as well with her ability to spell that sound.

Jasmine speaks English, but her first language and that of her family and proximate community is Punjabi, and she visits her extended family in India. We know that ‘the
language environment clearly influences developmental speech perception’ (Best et al. 1988:345) and also influences speech production, because even as adults, ‘most individuals perceptually assimilate certain non-native speech sounds with similar ones from the native language’ (Golestani & Zatorre 2009:56). This is confirmed by considerable research; Perez notes, for example, that ‘[o]ne influencing factor [of first language influence on the second language] is the articulatory basis of their first language, which is the physiological articulatory pattern habituated during the process of learning the phonology of the first language’ (Perez 1994:246). Perez (ibid.) also cites as important the influence of the language environment, that is, the variety of the language spoken by parents and other relatives.

Notwithstanding the possibility of Jasmine’s Punjabi-influenced naming of the letter <t>, it is likely that she actually said ‘t’ (although a recording for analysis is not available). While each letter represents a sound, which is made up of a combination of features, a single letter can come to represent a moving target of shifting sound features due, for example, to first-language influences. This is a case of a word-initial syllable with a degree of stress, however, and for such consonants their spelling is to a high degree predictable on the basis of sound alone.

6.9. Perception of a Punjabi-like English /t/ as [d]. Given the likelihood that Jasmine’s production of the English /t/-sound was influenced by her Punjabi L1 or by Indian English, it is likely that her pronunciation interfered with the ability of the judges to hear the sound.

Specifically, the judges could well have perceived Jasmine’s production of [t] as a voiced [d]. Similar to what was noted above, Das and Hansen (2004:344) say: ‘The many variations of plosive stops in Indian languages versus just the two in English can lead to confusion in Indian English speakers, which may create a lack of intelligibility for listeners who speak a different dialect of English’. And Luchini and Chiusano (2009) conclude that listeners may find it more difficult to identify unaspirated /p, t, k/ as voiceless, thus hearing /t/ as /d/, and Karamat (2001) notes that Punjabi speakers may appear to be substituting the voiced stops /b, d, g/ for the voiceless aspirated stops /p, t, k/.

Such perception issues are due to the fact that the internal composition of the sound systems of different languages can result in language-specific listener ability to discriminate sounds (Keating et al. 1981). In fact, Keating and colleagues demonstrate that American English speakers are not very sensitive to these /t/ vs. /d/ differences in other languages, due to their differing discrimination of these two sounds with shorter VOTs. For example, native English speakers typically cannot hear the difference between the English alveolar and the Hindi retroflex stop consonants (Golestani & Zatorre 2009:56). And more generally, English-speaking listeners have been found to be more sensitive to aspiration changes as speakers of an aspirating language as opposed to listeners of a voicing language like Dutch, who are more sensitive to prevoicing (Brandmeyer et al. 2012).

There is also the possible effect of the speaker’s own perception on her pronunciation of the voiceless stops. Avery and Erlich (1992) contend that the aspiration of the voiceless stops in Punjabi is much stronger than in English, saying this:

Punjabi speakers, thus, do not hear the aspiration of the initial stops /p/, /t/, and /k/ in English. Instead, they hear voiceless unaspirated stops and produce them as such. English speakers hear the ... Punjabi speakers’ voiceless unaspirated stops as voiced stops, just to add to the confusion. (Avery & Erlich 1992:129)

7. Method. We take three approaches to examining the question of how a correct spelling of tomography could be mistakenly heard as domography by the listener-judges of the spelling bee. First, we asked Jasmine about the spelling event itself to get
her sense of the relative difficulty of the word and how she responded. Second, we considered the research, as outlined above, on spelling, speech production, and speech perception to determine the likelihood of (i) Jasmine’s making a real word-recognition mistake, (ii) Jasmine’s not pronouncing the letter-name <t> in Standard American English, and (iii) the judges not perceiving the letter-name as Jasmine intended. Third, we observed Jasmine’s spoken language to assess the degree to which Punjabi may be influencing her English pronunciation by recording her spoken English at various style levels, informal to formal: natural conversation, reading passage, minimal pairs, and letter naming (spelling).

8. Subjective findings. We spoke with Jasmine and asked for a self-report of her experience at the spelling bee, specifically how she spelled the word and if she considered the word difficult. She responded without hesitation that she spelled the word <t-o-m-o-g-r-a-p-h-y>, and that it was not a difficult word for her. We also observed that Jasmine used the voiceless, unaspirated /t/-sound at least once in this brief, unrecorded conversation, limited as it was by her continuing sensitivity to recalling the experience.

9. Objective findings: Jasmine’s voice onset times. Jasmine’s language was recorded first in natural face-to-face conversation in which thirteen words with initial /t/ were isolated and analyzed for onset length. The multisyllabic words beginning with /t/ from an oral reading of a children’s story (The tortoise and the hare) were also recorded and analyzed, as well as words embedded in a few sentences read in isolation. Jasmine’s word-initial /t/ was also recorded and compared with her word-initial /d/ in four minimal-pair sets, the most conservative style. Finally, she was recorded spelling seven words, and these were analyzed for VOT as her ‘spelled’ /t/-sounds. For comparison, Jasmine’s mother, whose limited English is much more influenced by her native Punjabi, was also recorded reciting the same list of /t, d/ minimal pairs that Jasmine did. In addition, the /t/-pronunciation of a fifteen-year-old L1 American-English-speaking girl, Mary, was also recorded spelling the word toxicants in order to compare her English /t/ to Jasmine’s.

Jasmine’s productions of /t/, as seen in Table 3, demonstrate a range of VOTs from 27 ms to 88 ms across all production tasks: speaking, reading, minimal pairs, and letter naming (spelling). Her shorter VOTs were less aspirated and sounded closer to English /d/ than to /t/. Jasmine appears to maintain a phonemic distinction between /t/ and /d/, however, given that the range of her VOTs for /d/ in minimal pairs was 10–13 ms, even though her shorter VOTs for /t/ (starting at 27 ms) may cause some of her /t/-sounds to be perceived as /d/ by native English speakers.

<table>
<thead>
<tr>
<th>Natural Conversation</th>
<th>Reading Passage</th>
<th>Minimal Pairs</th>
<th>Letter Naming (Spelling)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOT &lt; 50 ms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>word</td>
<td>VOT</td>
<td>word</td>
<td>VOT</td>
</tr>
<tr>
<td>two hundred</td>
<td>27</td>
<td>tortoise</td>
<td>32</td>
</tr>
<tr>
<td>take</td>
<td>40</td>
<td>topple</td>
<td>34</td>
</tr>
<tr>
<td>to</td>
<td>42</td>
<td>tasty</td>
<td>34</td>
</tr>
<tr>
<td>to</td>
<td>44</td>
<td>tortoise</td>
<td>34</td>
</tr>
<tr>
<td>toddler</td>
<td>35</td>
<td>dapper</td>
<td>13</td>
</tr>
<tr>
<td>tumbled</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>43</td>
<td>/t/</td>
<td></td>
</tr>
<tr>
<td>tempted</td>
<td>48</td>
<td>touch</td>
<td>46</td>
</tr>
<tr>
<td>tomography</td>
<td>48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table 3. Continues)
Jasmine’s speech is only somewhat accented, given that the majority of her /t/-sounds are consistent with English pronunciation. However, as Figure 1 demonstrates, about 20–30% of Jasmine’s /t/-sounds (depending on speech style) have VOTs of less than 50 ms, and there appears to be an inverse relationship (at least based on this limited data set) between VOT and speaking style: the slight but progressive decrease in /t/-occurrences with VOTs of less than 50 ms is consistent with an increasing formality of speaking style: conversation (31%), reading passage (26%), minimal pairs (25%), and letter naming (21%).

![Figure 1. Jasmine’s style-specific % of [t]-realizations with VOT < 50 ms.](image-url)
Additionally, as can be seen in Table 3, there were some instances when the amount of aspiration diminished considerably, and her VOT measured around 30 ms. Out of the sixty-five total words in which her initial /t/-sounds were measured, she demonstrated a VOT of less than 50 ms in seventeen (26%) instances. Shorter VOTs usually occurred when she was speaking quickly or when she became frustrated. For example, when she was twice prompted by the interviewer to retry spelling a word, her /t/ began to have a [d]-like quality, with a shorter VOT the more frustrated she became. Once she had the correct spelling, she spelled the word slowly, with slightly exaggerated pronunciation, and her /t/ was once again a clear American English /t/, with a longer VOT. In Figure 2, two of Jasmine’s VOTs can be seen next to that of L1 American-English-speaking Mary, all naming the <t> in the spelling of toxicants. The spectrograms used for Figures 2 and 3 were produced using Praat (Boersma & Weenink 2014).

![Figure 2](https://example.com-figure2.png)

**Figure 2.** Jasmine’s vs. Mary’s spelling of toxicants: VOTs = 27 ms, 67 ms, 91 ms.

Similar effects occurred as Jasmine was reading. If she read a word that was unfamiliar to her and had to go over it a few times, her VOTs started out shorter. Once she had learned the word and could read through it in a more natural style, her VOT was longer, with more differentiation between /t/ and /d/. (The apparent insecurity brought on by this circumstance may even parallel the nervousness of a contestant spelling the tie-breaking word at a regional spelling bee.)

In contrast, the minimal-pair realizations of Jasmine’s mother, who speaks English with a strong Punjabi accent, were more accented. The mother’s /t/-sounds were all unaspirated, with VOTs that were actually shorter than the VOTs for her own produc-
tion of /d/ by about 10 ms. On average, the mother’s VOTs for words beginning with /d/ were around 20 ms, and her VOTs for words beginning with /t/ were about 10 ms.

When Jasmine and her mother are compared, as seen in Figure 3, her mother, with the most accented English, had the shortest VOTs for her /t/-sounds at around 10 ms. In contrast, Jasmine’s less accented English resulted in VOTs that were on average between 50 and 60 ms, but ranged from 32 to as high as 84 ms (shown).

As alluded to parenthetically above, there is also the distinct possibility that, under the stress of performance at the spelling bee, Jasmine’s English pronunciation reverted to something more Punjabi-like. It is known that stress can cause changes in articulation (Murray et al. 1996), and an interesting change in Jasmine’s own pronunciation of word-initial /t/ can be observed. While reading the story of *The tortoise and the hare*, Jasmine at first had difficulty identifying and saying the word *tortoise*, which occurred fourteen times. Then, once comfortable with the word, she proceeded to say it without difficulty. The VOTs of all fourteen instances are given order of occurrence in Table 4, and note that the /t/s of her first two attempts had VOTs of less than 50 ms, and all but one after that had VOTs at or above 50 ms.

**10. Linguistic conclusions.** We have considered three possible explanations for the circumstance that the contestant in this spelling bee recalls spelling *tomography* correctly, but the judges heard *domography*: actual spelling error, L1 interference in the

![Figure 3. Jasmine’s mother and Jasmine saying *town*: VOTs = 10 ms, 82 ms.](image)
speller’s articulatory production of /t/, and the influence of Standard American English (SAE) on the judges’ auditory perception of a non-SAE [t]. We conclude that it is possible but unlikely that Jasmine made an error in spelling. The explanation lies instead in the perfect-storm combination of the speller’s accented production of /t/ and the judges’ subsequent perception of the sound as [d].

**10.1. Possible influence of word/morpheme recognition on Jasmine’s spelling.** Jasmine won her school competition and was one of two finalists in the regional bee; she is an accomplished speller. Since sound-letter correspondence is not sufficient for learning to spell well in English, it is unlikely that a speller at the competence level of Jasmine would not recognize and spell *tomography* based, at least in part, on its morphology. Since, however, we cannot really know what her word-recognition skills are without further study, it is at least possible that she may be using alternate ‘extreme’ spelling strategies, such as committing to memory the SNSB’s word-list dictionary.

**10.2. Possible influence of Jasmine’s L1-Punjabi on her L2-English pronunciation.** The onus of producing perceptible sounds is on the spelling bee contestant, who must face the judges and pronounce the word before and after spelling it. Rule 10 (‘Speller’s role’) of the contest rules states: ‘The speller while facing the judges makes an effort to utter each letter distinctly and with sufficient volume to be understood by the judge’ (Scripps National Spelling Bee 2013). It is likely that the initial /t/-sound Jasmine produced in *tomography* was a Punjabi-like /t/, unaspirated and possibly dentalized. There is a hint of a style-related distribution of her range of /t/ variation, that is, increasing VOTs as she moves from natural conversation, to reading, minimal pairs, and letter naming (spelling). (We say ‘hint’ because of the paucity of data to support the pattern.) However, it is apparent that the Punjabi-influenced /t/ slips into Jasmine’s speech in circumstances that appear to relate to how fast she is speaking or to her emotional state. Given the data reviewed here, it is perfectly possible that the length and stress of an event such as the spelling bee may have set up a context that would lead her to produce a short, unaspirated /t/ that was heard by the judges as an English [d].

**10.3. Possible influence of Jasmine’s English pronunciation on judges’ perception.** The spelling bee judges rely solely on their auditory perception to determine the correctness of a spelling. Rule 9 (‘Judges’ role’) of the current contest rules states:

<table>
<thead>
<tr>
<th>READING PASSAGE</th>
<th>WORD</th>
<th>VOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>tortoise</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>tortoise</td>
<td>53</td>
<td></td>
</tr>
</tbody>
</table>

*Table 4. Jasmine’s sequential VOTs (in ms) for the initial /t/ of *tortoise* in her reading of *The tortoise and the hare.*

*Table 4. Jasmine’s sequential VOTs (in ms) for the initial /t/ of *tortoise* in her reading of *The tortoise and the hare.*
‘The judges listen carefully, determine whether or not words are spelled correctly and uphold the rules. The decisions of the judges are final’ (Scripps National Spelling Bee 2013). It is likely that Jasmine’s initial /t/ in spelling tomography was perceived by the judges as a voiced [d], due to her accented pronunciation. First, we know that the judges themselves struggled with the pronunciation because they asked to review the video recording. In addition, there is overwhelming research indicating that voiceless stops are less aspirated in Indian English, and specifically in Punjabi-influenced Indian English, making them sound more like the /d/ than the /t/ of American English.

11. POLICY CONCLUSIONS. The fundamental policy question is this: Does the SNSB endeavor to maintain the illusion of linguistic homogeneity in American English in its efforts to help children acquire standard English, or does it come to terms with the reality of linguistic variation by recognizing the linguistic heterogeneity of our increasingly diverse nation and world, and develop policies and rules that allow all English-speaking children real access to the extraordinary benefits of the competition?

We know well that language is inherently variable, and its variability is manifested in the many standard and nonstandard varieties of English spoken throughout the United States and the world. The SNSB might be prompted by the growing diversity of its participants to develop strategies for accommodating accents of spellers whose primary language is not English or who speak a non-American but standard variety of English. Such accommodation may also include regional and social dialects of American English. A speller’s particular accent might thereby not interfere with the correctness of his or her spelling bee responses.

There is no doubt, however, that this is not an easily met challenge. On the one hand, the American spelling bee is by definition an oral exercise, so introducing any kind of writing component might present unacceptable changes to the nature of the bee. And to maintain the oral-response requirement by, for example, introducing a telephone- or radio-alphabet, such as the military’s F-foxtrot, G-golf, H-hotel, I-India, and so forth, presents its own set of problems. On the other hand, the SNSB has already found a way to accommodate special-needs contestants, with rule 4 (‘Special needs’), which specifies: ‘The judges have discretionary power to amend oral and/or written spelling requirements on a case-by-case basis for spellers with diagnosed medical conditions involving sight, hearing, speech, or movement’ (Scripps National Spelling Bee 2013). Although speakers of non-SAE varieties are not special-needs spellers, rule 4 demonstrates the SNSB’s intent and ability to include all children as contestants. And, as Jasmine’s case clearly suggests, it may be time to rethink the issue of dialect variation vis-à-vis the spelling bee, given the diversity of students eligible to become spelling bee contestants. Such a review of best practices as related to the spelling bee could result in achieving what might come to be considered complementary goals: maintaining the essential character of the American spelling bee, and making the bee accessible to all children living and going to school in the United States, as well as to those students from other countries who participate in our national spelling bee.

REFERENCES


BEST, CATHERINE T.; GERALD W. McROBERTS; and NOMateMBA M. SITHOLE. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click

McMenamin
Department of Linguistics
California State University, Fresno
5245 N. Backer Ave. M/S PB92
Fresno, CA 93740
[geraldm@mail.fresnostate.edu]
[lk583@georgetown.edu] (Kerr)

[Received 10 June 2013; accepted 25 April 2014]