“It is impossible for an Englishman to open his mouth without making some other Englishman hate or despise him.”
George Bernard Shaw
(Wardhaugh and Fuller 2014: 24)

“AMERICAN VS. ENGLISH”

US AMERICAN AND BRITISH ENGLISH SPEAKERS’ ATTITUDES TOWARD PHONOLOGICAL AND LEXICAL FEATURES IN ACCENT TAG VIDEOS

Abstract

The study examined two YouTube Accent Tag videos to reveal how pairs of British and American speakers reacted to each other’s and their own varieties as they pronounced words from a list and answered several questions by offering their lexical choices. Accent Tag videos represent a novel way for lay people to be involved in science by offering their language varieties and opinions, accumulating data in unprecedented numbers in the history of dialectology and also creating a rich source for various types of linguistic inquiry. The results showed a marked difference between the manners in which standard-accented British and American speakers evaluated both their own and their interlocutor’s speech. The two RP-accented British speakers were more prone to criticise the two mainstream accented American (General American) speakers’ speech but did not receive criticisms from their interlocutors. Further, neither of the British speakers disparaged their own speech, while one American speaker did. The study also identified some disfavoured features of American English, mainly phonetic differences in comparison with RP, including the lower unrounded LOT-vowel [ɑ], T-flapping, and the flat BATH-vowel [æ] (Wells 1982). The findings of these case studies support the hypothesis that in the game of “American vs English”, RP-accented British English is generally rated higher than mainstream American (or GA) in both groups of speakers.
Keywords

Accent Tag, American, British, GA, RP, attitudes, citizen sociolinguistics

While living in the USA, I noticed an advertisement for a media newsreader’s job opening, for which “a British accent” was specified to be advantageous. As my accent was perceived to be changing toward American at the time, this was surprising, and it made me wonder if British English was generally liked more than American English by the native speakers of these varieties. The reverse situation seemed unlikely, namely that in Britain one would look for a newsreader with “an American accent”. Yet this ranking of the language or its accent was not in line with the general influence of American culture on the world.

The issue of British versus American has not yet been placed into the focus of existing research, although often times borne out by perceptual studies. Perceptions of regional English varieties, both national and international, were examined in previous research (e.g., Giles 1970; Ball 1983; Cargile & Giles 1997: 195; Hiraga 2005; Coupland and Bishop 2007), and they generally concluded that speakers of English, both native and non-native, rated standard-accented British English the most attractive, more so than standard-accented American English speech. However, none of the studies explored the issue from the American perspective. Further, they were not based on real-life discourse. Moreover, they did not reveal clearly if it were some sounds (and then precisely which sounds), or other characteristics, such as voice quality, intonation, or rate of speech, that made subjects react in a positive or negative manner. In summary, no study to date has examined how standard-accented British and American speakers naturally react to specific words and phonological features in one another’s speech. Using lay people’s published videos on YouTube (a new method in linguistics), the present investigation is designed to fulfil this gap. Its goal is to discover if standard accented American speech is indeed rated lower than standard accented British speech by speakers of American and British English. Considering the economic, cultural, and political power of the USA in the past century, such a result would be surprising.

Vlogs are ideal vehicles for the investigation of semi-natural spoken discourse since they capture the entirety of behaviours and processes, including language use, paralinguistic cues, gestures, mimics, and other verbal and nonverbal cues. Videos are rich in contextual cues, which may be employed in the pragmatic interpretation of such phenomena as the repetition of a word with exaggerated pronunciation and different intonation, accompanied by various facial expressions. In addition, the selected Accent Tag videos (a special group of vlogs focusing on
language, explained below) can be described as both structured and natural. They are structured to some extent because they are related to a linguist’s dialect survey (to be described below), and therefore better suited for systematic research. Also, they are quite natural since they are just one of the many videos routinely produced by vloggers.

Hence this study will investigate how standard accented British and American YouTubers rate each other’s and their own speech in a playful competition performed in the Accent Tag videos, with the aim to find out which variety is preferred by the speakers. More specifically, it will explore the self-evaluation and other-evaluation of the natural speech of two pairs of British and (North) American speakers who took the Accent Tag challenge together. A further goal is to reveal some specific phonological and lexical issues that commonly receive attention or are possibly judged in a positive or negative manner. It is expected that despite an increased sensitivity toward politeness issues due to the publicity of the videos, British speakers will rate their own variety higher than that of Americans. Based on the personal experience described at the beginning of this essay, American speakers are expected to have a positive evaluation of British English, while ambivalent toward their own variety. In addition, evaluations of specific features of British and American accented speech shall be explored.

This study will employ the term “standard” for both British and American model varieties and accents, and for accent it will also use GA (General American) as a matching American counterpart for British RP (Received Pronunciation). These terms have been debated but also widely used in the available literature; see Kretzschmar (2004) concerning GA, Wolfram and Schilling (2016) with regard to standard, and Lippi-Green (2012: 46) on accent. The British and US American accents were called “standard” and “North American” in Giles (1970), “RP” and “Network English” in Hiraga (2005), and “Queen’s English” or “Received Pronunciation” and “North American-accented speech” in Coupland and Bishop (2007), respectively. However, a large amount of the description below will contain how the vloggers react to one another’s speech and therefore the routine distinction will be made between the “British” (or “English”) and the (US) “American” speakers and their language in the videos under scrutiny.

**Literature Review**

Most people have an opinion about language use around them, but do not record it and make it public. Normally, such an opinion disappears with the fleeting moment. However, ever since
the Accent Tag videos were started in 2011, many people have been able to voice their opinions and make them permanent on YouTube and Tumblr. Rooted in linguist Bert Vaux’ novel online dialect survey of American English (Vaux 2002, Vaux and Golder 2003, more in Szitó 2020), these Accent Tag (Accent Challenge, or Accent Tag Challenge) videos reflect popular interest in linguistic variation, at the same time covering phonological and lexical varieties of English from all over the world. Despite their use of the narrow term accent, the videos offer a wide range of spoken language material for students of lexical and phonological variation, discourse and style. The present study will focus on how the vloggers evaluate their own and their interlocutor’s speech as they pronounce words from a list and answer several questions requiring words.

Accent

The question that makes standard British accent more attractive than standard American accent is not easy to answer. After all, accent is defined as a phonological variety of language, and if it is just that, then it is hard to see how one sound, or a stress pattern, could be judged more pleasing than another. However, accent also carries social weight. As a “bundle of distinctive intonation and phonological features” (Lippi-Green 2012: 332), the accent betrays much about the speaker, especially about their early socialization. Perhaps this is why people are sensitive to accents. Through stereotype formation, accent helps people orient themselves among the countless strangers that they meet or hear in their lives, supplying listeners with enough conventionally coded information to classify speakers (Wells 1982: 30). On the basis of accents, listeners are able do as much as form an opinion about the speaker’s political stance, “their general intelligence, their reliability, even (for someone the hearer cannot see) their handsomeness or beauty” (ibid. 30). In this way, speakers with a standard, close-to-norm accent will have an instant advantage over others as soon as they open their mouths. Nevertheless, in the case of two standard, hence socially highly valued, accents in two societies located in different continents, the question as to whether one is more attractive may even appear absurd.

Despite this absurdity, it has been suggested that British speakers have a negative attitude toward American English. Hiraga (2005: 289) wrote that “(i)n Britain, there seems to be an ingrained, cross-generational aversion to American English.” Several studies were conducted on accent perception, many of which built on Giles (1970), but the results did not bear out a consistently negative attitude. For instance, Giles (1970) found that British listeners
presented with the speech of a speaker appearing in the guise of various accents, evaluated North American as second to standard British English when ranking communicative content, and third in prestige. Only on the aesthetical scale was North American perceived to occupy a lower place, somewhere in the middle of the hierarchy of varieties of speech. In other terms, it was considered less pleasing. In a study similar to Giles’, Hiraga (2005) concluded that American (named Network English in the study) fared high, right below RP, in the dimensions of status and solidarity, whereas nonstandard accents, especially urban accents, were placed lower in the hierarchy. Both studies emphasized that the results can be explained by a link between accents and social class structure, but only in Britain is this formulated as either stigmatized or prestigious. Hiraga (2005) also claimed that American (i.e., Network English) was able to reach such a high ranking because it was outside the social class structure of British society.

These results were somewhat contradicted by Coupland and Bishop (2007), who conducted quantitative online research into stereotyped attitudes of accent labels. Their results showed much lower ratings of the label “North American accented English” than in the previous two studies described above, in which subjects reacted to speech samples rather than labels. In Coupland and Bishop (2007), North American accented English was ranked below several British local varieties, unlike in the previously mentioned studies: a little higher for prestige and lower for pleasantness. Although the dimensions and the levels of rating differed and sometimes contradicted in the studies, it is possible to conclude that, viewed from the British perspective, American accented English was perhaps only slightly lower in status than British accented English, but it was perceived less beautiful – although still much better than nonstandard urban accents associated with the working class.

The American perspective has nevertheless been unexplored: no study has so far focused on Americans’ reactions to British English. The studies in perceptual dialectology mainly investigated US American varieties and its speakers, and those in folk linguistics examined non-native accents. The assumption that Americans may somehow conform to the above mentioned negative British attitude was, however, stated in several studies. For instance, Cargile and Giles (1997: 195) claimed that “an American may think a stranger ’cultured' and ’refined' simply because his or her accent is deemed British.” Wolfram and Schilling (2016: 13) also noted that “Americans still assign positive value to British dialects … It is difficult to say exactly why Americans look upon British English so favourably, but one possibility is a
lingering colonial effect.” Further, in a study of Australians’ perceptions of English accents, RP was also called “the pan-Anglophone standard” (Ball 1983: 177).

Within their own country, however, Americans do not seem to be concerned much about prestige. Writing about dialects, Wolfram and Schilling stated that

(f)or the most part, Americans do not assign strong positive or prestige value to any particular dialect of American English. The basic contrast in the United States exists between negatively valued dialects and those without negative value, not between those with prestige value and those without. (2016: 13)

In the USA, students of dialects found that the standard was not as much connected to social class structure as in Britain and could rather be characterized with the lack of stigma. According to Wolfram and Schilling (2016: 13), “if a person’s speech is devoid of socially stigmatized structures, then it is considered standard or ‘mainstream’.” Niedzielski stated that standard is a variety that is “least likely to call attention to itself” (2002: 210).

It appears from the above that the concept and evaluation of standard versus regional accents is different in Britain and the USA. In Britain, RP is associated with high prestige, education, and upper-class speech, while “regional accent is a social stigma” (Giles 1970: 225). At the same time, standard American accent, the “accentless accent”, is not homogeneous and

...can best be characterized as what is left over after speakers suppress the regional and social features that have risen to salience and become noticeable. Decisions about which features are perceived to be salient will be different in every region, even different for every speaker, depending on local speech habits and the capacity of speakers to recognize particular features out of their varied linguistic experience (Kretzschmar 2004: 262).

Outside the countries hosting the examined standards, but still within the Inner Circle of historically native speakers of English (Kachru 1992), perceptions of these two standards vary. For instance, New Zealanders scored British (RP)-accented speakers high on scales of intelligence but low on likeability in a study by Wilson and Bayard (1992, as cited in Cargile and Giles 1997: 195). In another study, Ball (1983) found that Australians associated RP with very high competence and confidence, but with moderate social attractiveness, whereas East Coast US American, “a potential rival to RP as the international standard” (Ball 1983: 177), with high confidence, moderate competence, and low social attractiveness. Even non-native speakers tended to rank RP the highest and standard American somewhere below it, depending
on the dimensions employed in the study (e.g., Jenkins 2009). In summary, standard British and standard American speech may receive different evaluations in different dimensions both within the Inner Circle countries and outside. This issue may deserve more attention in the future.

Method

Accent tag videos collectively shortened and modified the original dialect survey to make it suited into a few minutes of video showtime. In most of them, there was a list of words to be read and several questions to answer, available (in varieties) on the web.

It has been suggested that Accent Tag videos are eminently suitable for examining attitudes to language. These have several advantages. First, the videos are numerous, estimated to be well over 400,000, and are still being created. Second, these videos have been inspired by a spontaneous, popular interest in language variation and are constantly available for linguistic analysis. Goudet (2018) calls studying vlogger content “a new approach to field work”. In it, the linguistic data are not elicited by a linguist but offered by the informants at their own initiative. Although such data are less systematic and reliable, this sort of research is more valid than experiments.

Third, the comments make it possible for vloggers and their audiences to engage in a dialogue, which offers even more data than the video content alone. Together with other YouTube videos, Accent Tag videos are part of a new culture called participatory culture (Androuotsopoulos 2013, Jenkins et al.) generated by Web 2.0 “social networks” (Rymes and Leone-Pizzighella 2018: 11). A further asset of these videos is their personal voice as the makers connect personal history to their speech. Rymes (2020: 9) also calls this method “citizen sociolinguistics”, which is the “practice of paying attention to how people talk about language.” Such citizen science methods have recently become increasingly popular in various fields of education and science (e.g., Schouten et al. 2019). It is this new method that the present research used in exploring lay people’s attitudes on varieties of speech, generally studied by perceptual dialectology and folk linguistics (Preston 1993, Long and Preston 2002).

When creating Accent Tag videos, YouTubers usually follow their regular video making practices, in that they expressly wish to entertain their viewers and attract more subscribers. They often produce what Androuotsopoulos (2013) calls a “spectacle”, a show which they
carefully prepare by setting the scene, by working on their appearance (clothes and makeup), and by editing the recording before publishing it. During the time this article was in preparation, the most viewed Accent Tag videos typically involved two or more people, engaged in conversation over linguistic problems, often teasing each other. Some popular Accent Tag videos reached hundreds of thousand viewers, and the most popular, also discussed below, was seen by over one and a half million viewers at the time of writing.

The present research has at least three advantages generally missing or unusual in the study of folk attitudes and perceptions of language. One (mentioned above) is that that the evaluations are carried out spontaneously, rather than under experimental conditions. A related advantage is that this method does not force participants to use metalanguage, that is, produce explicit verbal expressions about language, and as a natural alternative, allows for the expression of attitudes nonverbally, reactions such as subtle facial expressions of surprise or the imitation of a feature with an exaggerated phonology and prosody. In exploring such a sensitive issue, the option of not having to verbalize things becomes even more important, considering that opting out and off-record comments (hinting) are much less face-threatening than going on-record, that is, verbalizing a potentially face-threatening act in politeness research (Brown and Levinson 1987). A final advantage is that this method makes it possible to identify specific features or segments of language that are evaluated, rather than long stretches of speech, maps, or labels for accents or other varieties, which on the one hand fail to distinguish smaller scale differences subsumed within that label or stretch of speech and, on the other hand, obscure the precise feature or segment eliciting a negative or positive reaction. This is in accordance with McKenzie and Osthus, who suggest that in future studies of accent, “in-depth micro level research would help researchers to better understand the ways in which particular aspects of speech elicit specific types of speaker ratings” (2011: 111). The authors also suggested that “electronic social media” should be exploited for the analysis of folk attitudes to language variation. These case studies of Accent Tag videos will therefore also be able to demonstrate the potentials of a relatively new method called “citizen sociolinguistics” (Rymes and Leone-Pizzighella 2018), in folk linguistics and perceptual dialectology.

Naturally, the present research will also demonstrate the limitations of the method. The findings of case studies are not generalizable: they are best used for raising issues for further explorations on more systematically organized data. Also, videos on the internet may be removed or made inaccessible by their makers. Further, the structure of videos and the availability of data relevant for linguistic inquiry are not controlled by the researcher at all, even
less so than in ethnographic interviews. Additional shortcomings may derive from faulty interpretations of pragmatically obscure signals. Despite all the limitations, it must be noted that without the existence of these YouTube videos, it would not have been possible to raise and study the central question of the present essay.

Data collection and analysis

The research was conducted in two phases: a background viewing phase and an analytical, purpose-driven phase. In the first phase, I extensively watched Accent Tag videos to gain general impressions. In the second phase, I re-watched about 500 videos, and created a data base of about 50 in Excel to record names, links, number of views, number of subscribers, types of English, transcription segments, and observations.

For this essay, two Accent Tag videos were selected along several criteria. The videos were both recommended by the YouTube algorithm and Google upon the search words accent, tag, challenge, American, English and British in various combinations. The selection criteria also included popularity (a high number of views as well as subscribers), and that two people were to be seen in the video, one British and one American, communicating with each other. Another criterion was that the selected speakers’ speech should be close to the standard as much as I could estimate (and inasmuch as one may “speak the standard”, which involves idealization). In order to decide this, phonological, regional, and attitudinal clues were exploited (such as when one British English speaker uses the adjectives “posh” and “normal” to describe her own accent in one of the videos), and the origin of the speakers: Michigan (USA) and Oxford (England) provided in the other video. To gather information, several more videos from these two channels (and from a third one owned by the guest) were also watched. All four speakers were young women, as most common in the Accent tag videos.

One of the Accent Tag videos was published by an extremely popular vlogger, Gibi ASMR, with 244 million subscribers on January 24, 2020. On that date, her Accent Tag video reached over 1.6 million views, which was the highest ever recorded number in the literature on Accent Tag videos. Gibi had two channels and published videos three times every week on her ASMR channel. It must be noted that ASMR videos are supposed to make the listeners feel good and relaxed, and this excludes loud, harsh speech and negative content or behaviour. The most immediate influence is that in this video, the speakers talk in very low voices, close to whispering. The Accent tag video was created in 2018. For the occasion, Gibi had a guest,
another channel owner called Creative Calm from England, whom she could not see but hear. However, the viewers can see (and hear) them both due to technical assistance.

The other accent tag video examined here had 540,631 views on the same date as above and had been posted on the channel named after the owners Kaelyn and Lucy (Kaelyn and Lucy 2012). The channel had 222,000 subscribers in 2016, when it was discontinued. Kaelyn came from Michigan, USA, and Lucy was from Oxford, England. At the time of the recording, they were both staying in Michigan.

Results

Gibi ASMR

The first analysis targeted the most viewed Accent Tag video at the time of writing (the first part of 2020). This is a long video of nearly one hour, as opposed to the general Accent Tag video of a few minutes.

At the beginning, the American host, Gibi, asked her British guest to describe her accent, and this opportunity was used by the British speaker for two positive descriptors “posh” and “normal” to characterize her own speech. In a minute, the British speaker reacted negatively to the American pronunciation of vitamin, water (“sounds so strange”), with which the American speaker agreed, saying, “mine sounds less fancy”. Gibi soon complimented on the Englishwoman’s speech and at the same time deplored her own. Thus in just a few minutes’ time, British English was twice praised and American English three times deplored by the vloggers. This may be schematically represented by using verb substitute (pos, neg) to distinguish positive and negative statements as follows:

| 1 Br pos Br | 2 Br neg Am | 3 Am neg Am | 4 Am pos Br | 5 Am neg Am |

The statements contain pos in the place of a verb expressing positive reactions and neg in the place of a verb expressing negative reactions. They can be read in the following way: the British woman reacted positively first to the British variety; in the second instance, the British woman reacted negatively to the American variety, and so on.

At this point, Gibi announced that she would “give points” and continued to give her vote to the Englishwoman. This prompted the Englishwoman to return the points, thus,
generally to become more positive, perhaps for fear of appearing too critical to the viewers (the majority of whom were likely to be Gibi’s subscribers). In addition, because ASMR video makers go out of their way to make their audience feel good, speakers are assumed to be under more pressure than usual to respect the positive face wants (Brown and Levinson 1987) of the other, which in this video were carried out by praising, agreeing, and generally saying pleasant things. In the end, the vloggers awarded the same amount of points (three) to each other. The American speaker received the vote for the pronunciation of tomato, banana, and Doritos, whereas the English speaker received the point for the pronunciation of water, class, and ball. This, however, does not mean that their speech was equally evaluated if we consider all the opportunities when they expressed their opinions.

Overall, the American speaker tended to make negative comments on her own accent, whereas the English woman never said anything negative about her own accent. More precisely, the American vlogger expressed dissatisfaction with her own accent seven times. In addition, the British speaker once verbally agreed with the American's self-deprecation and twice gave criticism at her own initiative: once for the pronunciation of vitamin and water (which in her opinion were “strange”), and later for the word faucet. In conclusion, what differentiated the attitudes toward the two varieties of English was the presence or lack of negative comments, while both varieties received five positive comments (perhaps due to the point assigning initiative mentioned above). The chart below shows that altogether, there were seven negative statements on American (GA) speech, whereas none on British (RP) speech.

<table>
<thead>
<tr>
<th>Gibi ASMR</th>
<th>on American (GA)</th>
<th>on British (RP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>American comments</td>
<td>5 negative</td>
<td>0 positive</td>
</tr>
<tr>
<td>British comments</td>
<td>2 negative</td>
<td>5 positive</td>
</tr>
<tr>
<td><strong>Altogether</strong></td>
<td><strong>7 negative</strong></td>
<td><strong>5 positive</strong></td>
</tr>
</tbody>
</table>

Table 1

Positive and negative comments on British and American speech in the Gibi ASMR video

*The words and phonological features receiving attention*

The video revealed several features of American pronunciation that were judged negatively. Frequently, the negative opinion was expressed by repeating a word in focus with an exaggerated pronunciation and a grimace. The precise sound was inferred from the contextual clues.
Altogether, the words and the sounds in American English shown to be disliked by either speaker were as follows: water ([wɜːrɪ], all three sounds following the initial [w]), vitamin (the i-vowel [aɪ], T-flapping or T-voicing), theatre (T-flapping / voicing, R-pronunciation), class ([æ], perhaps the dark l [l] as well), ball (the lower unrounded LOT-vowel [a]), plaster (the a-vowel [ɑ]) (Wells 1982).

The American speaker commented on her pronunciation of theatre, saying, “it always sounds like I’m saying "A-D-D-R-E" (laughs). It’s kind of silly.” She also emphasized her dislike for her own pronunciation of three words together: vitamin, water, and class, proposing, “they're all like very … almost like harsh. They just … they don't flow very well.” She also compared the two pronunciations of ball, saying “I prefer your ball a lot better. I think our ball is so gross. I don't like it,” to which the Englishwoman replied in agreement, “I know what you mean.” The British speaker also commented on the American word faucet, stating, “You guys are so crazy… honestly.”

The words that the Englishwoman liked in American were tomato (ei – “sounds cool” “fresh, nice”), banana and Doritos (she did not make it clear what was attractive about the latter two).

According to these comments, the speakers could perceive differing aesthetical qualities in different pronunciations of words. If they were negative, they were characterized as “harsh”, “silly”, “gross” and “not flowing very well”. If a pronunciation was perceived as positive, it “sound(ed) cool” or “fresh, nice”. In this video the conversation established the category dimensions of “nice – ugly” and “like – dislike”.

The results may have been positively distorted (i.e., in favour of American English) because of the point-giving system that Gibi cleverly initiated towards the beginning of the video. Nevertheless, it was obvious, even after the first few minutes of the video, that British English was favoured by both speakers over American English.

**Kaelyn and Lucy**

The Kaelyn and Lucy Accent Tag video, created in 2012, is from a YouTube channel named after the American and English owners (respectively). Both speakers were from areas and cities (state Michigan in the USA and Oxford in England) that are often associated with correct speech and the national standard. Because Michigan speakers’ vowels have been undergoing the
Northern Cities Shift (Niedzielski 2002, Labov et al. 2006), Kaelyn’s pronunciation was slightly different from what narrowly counts as standard American, but as Niedzielski (2002) claims, this difference is still largely unnoticed by the public.

Compared to the previously discussed vlog, the Kaelyn and Lucy video was shorter and more focused on the task, and it was substantially different in at least two respects: 1. in the ratio of verbal and non-verbal messages, and 2. in its playful framing (Pichler 2006). In the few minutes of the video, Kaelyn and Lucy preferred to express their opinions or attitudes in a nonverbal fashion, as opposed to the speakers’ tendency toward verbalisations in the previously analysed vlog. Being very close friends, these speakers were able to rely more on other devices such as repetition with modified suprasegmental features (intonation, volume), facial expressions and gestures instead of formulating explicit opinions. Their playful, quick, and brief reactions seemed to be rather of the nature of teasing or genuine surprise than of praising or criticism. What they normally did was pronounce the words in the list, with each taking her turn. While doing so, the second person, often the Englishwoman, stuck in a mimicked pronunciation of the first person, but with lengthening vowels, exaggerating stress, and often with a rising intonation, indicating a possible surprise or disapproval.

Based on their nonverbal and paralinguistic signals, it was impossible to decode their attitude into the explicit “ugly – beautiful” dimension found in the previous analysis above, and even the “like – dislike” (or positive – negative) dimension seemed risky. Instead of “negative”, I decided to use the word “surprising” to code other-evaluations (but not self-evaluations) since this was what the facial expression suggested. However, the word “positive” was left, since the Englishwoman once made a semi-serious positive evaluation of her own accent, claiming “that’s how you pronounce it!”

The British speaker, Lucy, expressed surprise or amusement over American words and pronunciations seven times, whereas the American, Kaelyn, did the same over British words only twice. Lucy also told a story of requesting water in American restaurants and having difficulties to be understood (“What? (grimacing) "WAH-DER"?”). Compared with the previous video, here Kaelyn, the American also allowed herself to be surprised: once she exclaimed, “Are you kidding? You're lying. Ad-VER-tiss-ment?” In the chart summarizing the reactions, I recorded “surprising” in addition to “negative” in the previous chart. Here, too, the American’s pronunciation elicited more reactions of surprise or amusement than the British speaker’s. There was a lack of self-deprecating comments on the speaker’s own variety.
Kaelyn and Lucy on American (GA) | on British (RP)  
---|---
American comments | 0 negative | 0 positive | 2 surprising | 0 positive  
British comments | 7 surprising | 0 positive | 0 negative | 1 positive  
Altogether | 7 surprising | 0 positive | 2 surprising | 1 positive  

Table 2

Positive and negative / surprising comments on British and American speech in the Kaelyn and Lucy video

The words and phonological features receiving attention

The items evoking surprised reactions from the British speaker on American speech were: water, vase, address (the a-vowel), buoy (the uo-vowel sounding as [ui]), aunt ([aɹ]), theatre, syrup (the y-vowel and the R-sound [ɻ], sounding quite like [ɜ]). Apart from water, all the listed words were mimicked, in an exaggerated fashion, by the Englishwoman, to express her surprise or perhaps some disapproval. In the American woman’s case, it was surprising for her that according to RP, the stress is placed on the second syllable (-ver-) in advertisement, which in turn influenced the vowel and consonant pronunciation in the first three syllables. Incredulously, she imitated the RP pronunciation of the word, pronouncing the -ver- syllable with exaggerated emphasis. In addition, the same American speaker performed an affected imitation of the RP pronunciation of lawyer, pronouncing the first vowel as [ɔ], and dropping the R at the end of the word. My interpretation of the American’s nonverbal behaviour was “Look how fancy this is”, but there may exist other interpretations.

In summary, the pronunciation of water was obviously an item receiving a lot of attention in this latter video, since three segments (also appearing in other words) were different for British and American speakers. These were LOT-unrounding (the a-sound), T-flapping or T-voicing (the t-sound), and R-pronouncing (-the -er sound, usually represented by a single [ɜ] in transcription). The LOT-vowel [a] was also in a lower position, as compared to the British [ɔ] sound. Despite the little overlap in the tasks in the two videos discussed above (see Appendix), it was clear which words and sounds were evaluated.
Conclusion

This study of Accent Tag videos investigated the reactions of two pairs of native, standard accented speakers of American and British English while they were focusing on linguistic issues such as pronouncing words from a list and finding a word or expression for a concept. The unequal distribution of positive as opposed to negative cumulative surprised reactions and comments on each variety show that British English was considered superior by both types of speakers.

Further, it was possible to capture several features of General American accent most frequently criticized. The following words elicited reactions to their pronunciation: vitamin, water, class, plaster, ball, theatre, buoy, aunt, syrup, and advertisement. The most conspicuous example was the pronunciation of water, a word which was discussed in both videos, and generally elicited most reactions in numerous other accent-related YouTube videos, as well. More specifically, the low unrounded LOT-vowel ([ɑ]) (Wells 1982: xviii), T-flapping or T-voicing, and R-pronouncing were the characteristics of General American which were negatively perceived and evaluated in this word (and in others). A further characteristic mentioned more than once was the flat BATH vowel ([ɑː]). All of these are well-known features of General American differentiating it from RP-like British English.

It was also found that pronunciation differences usually elicited emotional reactions, whereas lexical varieties in which pronunciation did not play an important role hardly ever did. On the contrary, people seemed generally at ease while discussing lexical varieties (as though there had been no risk involved), while they could not help reacting to pronunciations with pride, shame, surprise, and perhaps even disgust – as if there had been so much more at stake. This highlights the strength with which accent and identity are linked.

It is precisely this link between accent and identity that makes the results interesting. Considering the present economic, cultural, and political power of the USA, one may well expect a high degree of linguistic security, reflected in positive evaluations of standard American accented speech. That this is not the case, or at least not so compared with standard accented British speech, may be explained by colonial history, as suggested before (Wolfram and Schilling 2016: 13). If this is true then the past few hundred years’ history of the English language is still somehow present in the collective minds of present-day young speakers of English, and this may also be valid to other countries. Further studies are required to confirm this explanation. Although analysing Accent Tag videos certainly had its limitations and did
not allow much control of variables, it was the most accessible means to address spontaneous reactions of British and American speakers in real discourse.

Works Cited

Androutsopoulos, Jannis. 2013. “Participatory Culture and Metalinguistic Discourse: Performing and Negotiating German Dialects on YouTube”. In Tannen, Deborah and Trester, Anna Marie (eds.) Discourse 2.0: Language and New Media. Washington, DC, Georgetown University Press: 47–70.


Goudet, Laura. 2018. “‘I do not have an accent’: I don’t have an accent: folk approaches to dialectology”. Actes du colloque du (Amis du CRELINGUA), Chambéry: les Amis du CRELINGUA, hal-02118788


Sztó Judit. 2020. “‘I do know that I have an accent’: Identities expressed through English as a lingua franca in Accent Tag videos”. In Kálmán, Csaba (ed.). *DEAL 2020: A Snapshot of Diversity in English Applied Linguistics*, 95-118. https://www.eltereader.hu/media/2020/12/K%C3%A1lm%C3%A1n-Csaba_WEB.pdf


**Appendix**

Word lists and questions in the two Accent Tag (Accent Challenge) videos as they appeared together with the published video.
The overlapping words are highlighted.

<table>
<thead>
<tr>
<th><strong>Word list Kaelyn and Lucy read:</strong></th>
<th><strong>Gibi and Creative Calm read:</strong></th>
</tr>
</thead>
</table>
| Garage, herb, **schedule**, figure, jaguar, lieutenant, **water**, advertisement, vase, route, ballet, **tomato**, leisure, address, buoy, aluminium, aunt, oil, theatre, caramel, fire, sure, crayon, toilet, New Orleans, pecan, both, again, probably, spitting image, Alabama, lawyer, coupon, mayonnaise, syrup, pyjamas, caught | **Round 1:**  
Vitamin **Water** Class **Schedule** Tomato Sauce Issue Banana Bread Ball or Bowl? Nonchalant Vaccine  
**Round 2:**  
Baby carriage  
**Shopping cart**  
iPhone/Androids are:  
Summer, Spring, Winter, and ____  
If you need to buy something, you go to the ____  
Biscuit vs. Cookie (wtf is a Crumpet)  
What you put on a small cut?  
Peanut Butter and _________  
At the sink, the water comes out of the:  
**Trainers vs. Sneakers**  
What’s a banger?  
What schools do you go to in order?  
Chips vs. Fries vs. Crisps?  

**Questions to answer:**  
What is it called when you throw toilet paper on a house?  
What is the bug that when you touch it, it curls into a ball?  
What is the bubbly carbonated drink called?  
**What do you call gym shoes?**  
What do you say to address a group of people?  
What do you call the kind of spider (or spider-like creature) that has an oval-shaped body and extremely long legs?  
What do you call your grandparents?  
**What do you call the wheeled contraption in which you carry groceries at the supermarket?**  
What do you call it when rain falls while the sun is shining?  
What is the thing you change the TV channel with?