Accuracy and acceptability of second-dialect performance on American television

REMCO KNOOIHUIZEN
Rijksuniversiteit Groningen

This article analyses a case of second-dialect performance as an idealised instance of second-dialect acquisition, without mitigating factors such as access, analytical ability and motivation. It focuses on the Australian English and American English speech of three young Australian actors. An acoustic analysis of their short-vowel systems shows that they can successfully adapt to perform in an American English accent, but that their second-dialect system is less stable and more variable than their native system.

A foreign-accent rating experiment on the actors' American English with American English judges shows that the actors on average are thought to sound slightly less American than the native American English-speaker controls. The discrepancy between the acoustic accuracy and listener acceptability may be explained by judges attending to different features from those included in the acoustic study.

This study of second-dialect performance shows what is maximally possible in second-dialect acquisition. Given the difference between the two measures of success, studies of second-dialect acquisition would benefit from including subjective measures in addition to acoustic accuracy.

1 Introduction

An actor's success depends on the extent to which they can convincingly portray an identity other than their own. Actors have many tools at their disposal to this end, both visual and behavioural; one of these tools is their use of language, for example, by speaking a different dialect from their own (see, e.g., Hodson 2014). Such performance in a second dialect may not always be on target, but while, for example, Bell & Gibson (2011: 569) acknowledge that inaccuracies may be accidental, their focus is on strategic inaccuracies that add an extra layer to the performance. For many second-dialect performances, however, and indeed for many audiences, inaccuracies need not be layered and a successful linguistic performance hinges purely on hitting the target.

Reviews and online comments often praise or criticise the accuracy of a (staged) second-dialect performance. The second-dialect performances of Hugh Laurie, as Dr Gregory House in *House*, and David Tennant, as the Tenth Doctor in *Doctor Who*, are generally seen as highly accurate and successful. But if accuracy and inaccuracy can be strategic, it is important to understand how they behave when they are merely accidental.

This article reports on a two-part study of American English second-dialect performance by three Australian actors. An acoustic analysis of the short vowel systems in their first and second dialects is presented in section 3; a listener acceptability experiment on the American English performance data is presented in section 4. The results from the acoustic analysis and the holistic acceptability experiment are then compared (section 5) and contextualised in the field of second-dialect acquisition and performance.

2   Background

2.1   Second dialects in performance

The sociolinguistic literature has paid considerable attention to the use of dialect in performance. This includes dialect performance in music (Trudgill 1983; Beal 2009; Coupland 2011), on the radio (Coupland 2001) and in film (Hodson 2014; Bucholtz 2011). Often, a dialect is performed because it is felt to be appropriate for the context (e.g. the American English accent in early songs by the Beatles; Trudgill 1983), because it conveys a form of authenticity (as in, e.g., the Arctic Monkeys songs; Beal 2009) or because it evokes a particular 'feel' (Welsh-accented radio broadcasts (Coupland 2001) or New Zealand advertisements (Bell 1992)). Studies of dialect performance, then, tend to investigate questions related to this authenticity – including the effect of more or less successful performance. The question of success in performance leads to further questions about how dialect performance relates to (second-) dialect acquisition.

The question of authenticity is not straightforward, and what is seen as authentic is based on subjective greyscale judgements on multiple parameters (Hodson 2014: 224–30). While it may be argued that a performance is inherently deauthenticating (see Coupland 2009: 284), this need not be the case. Coupland (2009: 294) discusses how the Artic Monkeys 'successfully perform authenticity', and how 'speakers' agentive control of a repertoire of meaningful feature[s] and styles' can be seen as performance (authenticity that is performed); Bucholtz (2003: 408), Coupland (2009: 298) and Hodson (2014: 235) claim that authenticity ultimately is constructed within the context of the performance. Although accuracy of the dialect performance is often seen as a measure of success and of authenticity, this context-based construction means that '"accuracy" is irrelevant' (Coupland 2009: 298).

What makes authenticity an even murkier issue in the current study is that the performed dialect is not a non-standard variety, as is often the case. Like David Tennant's Doctor Who, performed in a second-dialect version of Standard Southern British English, the performed General American English accents constitute a standard variety of English. And indeed, Coupland's measures of authenticity can apply to Standard English as much as they can to non-standard English (Hodson 2014: 225–6). In popular perception, however, this type of dialect performance is about 'losing'

rather than 'putting on' an accent; it is about no longer being perceived as a linguistic Other and becoming invisible (or inaudible).

2.2 Acquisition and performance

If an actor performs in a second dialect, they need at some level to have acquired that dialect. This raises the question of whether actors' second-dialect acquisition differs in any way from acquisition outside a performance context, or whether the two can be seen as equivalent. Siegel (2010: 64–6) is quite adamant that D2 performance is separate from naturalistic D2 acquisition, and that studies of second-dialect performance do not inform our knowledge of SDA. And of course, many factors that are relevant in SDA are neutralised in D2 performance – predominantly social factors, but perhaps also linguistic ones. But to avoid throwing out the baby with the bath water, it is worth exploring where the differences lie between acquisition and performance, and discussing where the value of performance data may lie for studies of second-dialect acquisition.

One difference between performance and acquisition lies in the type of knowledge or competence required. Performance, according to Siegel (2010: 219), 'requires only explicit knowledge or conscious awareness of some of the features of the D2, and the ability to imitate them accurately'. This means that performance data give no evidence of 'automatic processing of underlying mental computational procedures' (2010: 74), the implicit competence that is evidence of true acquisition. And indeed, Siegel's discussion of dialect coaching (2010: 196) describes very explicit methods of learning: awareness raising, guided listening and imitation, starting with 'signature sounds' (what we might call stereotypes) and progressing to less obvious differences between the D1 and the D2. We also see predominantly conscious, cognitive learning in Trudgill's (1983: 145–58) discussion of second-dialect performance according to Le Page's 'acts of identity' model (Le Page & Tabouret-Keller 1985), in which success (read: accuracy) is determined by:

1. the extent to which we are able to identify our model group;
2. the extent to which we have sufficient access to [the model groups] and sufficient analytical ability to work out the rules of their behavior;
3. the strength of various (possibly conflicting) motivations towards one or another model and towards retaining our own sense of our unique identity;
4. our ability to modify our behaviour (probably lessening as we get older).

Implicit learning, which Siegel appears to view as true acquisition of a D2, is described as a subconscious, automatic process, in which learners are not aware of their own learning – even if the process is not quite implicit or subconscious at every level (Dörnyei 2009: 138). Some explicit knowledge does seem to be required, however: a learner is thought to only be able to acquire a feature if they are to some extent aware of it (the *noticing* hypothesis; Dörnyei 2009:

164) and most SDA studies fail to clearly distinguish the two types of processes (Siegel 2010: 74). In an SDA context, Nycz (2016) has shown that awareness of D2 features is not a necessary requirement for dialect change, but that it does influence the process: speakers showed enhanced levels of adoption or rejection of a D2 feature they were aware of compared to a feature that had not been noticed.

All in all, the difference in learning styles between learning for performance and true acquisition looks to be quantitative rather than qualitative, with performance typically showing more explicit and acquisition showing more implicit learning methods, but neither necessarily relying exclusively on one or the other.

Another important difference between second-dialect performance and acquisition is motivation. In acquisition, there is often an integrative motivation. Canadian learners of D2 New York English, for example, 'deliberately manipulat[ed] their use of salient dialect features' in order to show or hide their Canadian identity (Nycz 2016: 70). The motivation for second-dialect performance is often more instrumental: a character's identity does not necessarily influence an actor's identity. Even though the 'acts of identity' model mentions motivation in the context of identity, this is probably to be interpreted loosely. The use of American English by the Beatles (Trudgill 1983) was not to convey a personal identity as Americans, but a professional and therefore more instrumental identity as pop musicians with little bearing on who they were as individuals. Identity, therefore, is likely to be a more important motivational factor in acquisition than in performance.

2.3 Measuring success in SDA and performance

Acquisition and performance also differ in what constitutes 'success'. Successful performance is often judged holistically (Hodson 2014: 228), even though raters may be able to point at specific linguistic features or at non-linguistic features such as the rater's awareness of the performance actually being in a second dialect (Hodson 2014: 231–3). Similarity to real-world dialects plays a role, but is not the determining factor in the success of a performance (Hodson 2014: 236). Studies of second-dialect acquisition, by contrast, have a much more particular measure of success, comparing learners' use of specific D2 features to native speakers of the second dialect, or charting the loss of D1 features. Often, the features investigated are the 'signature sounds' that dialect coaching focuses on: the loss of *t*-flapping (Chambers 1992) and the acquisition of *t*-glottaling (Tagliamonte & Molfenter 2007) by Canadian learners of England English; the loss of Canadian raising (Nycz 2013a, 2016) or the acquisition of the Philadelphia split-TRAP pattern (Payne 1976, 1980). Here, similarity to real-world dialects is in fact the determining factor. Of course, quantitative comparison of the use of stereotypical features also occurs in studies of performance; see Trudgill (1983), which focuses on specific features of American English (dubbed the 'USA-5 model' by Simpson 1999).

2.4   The value of performance data

It is clear, then, that second-dialect performance and acquisition cannot simply be equated. Performance is likely to involve relatively more explicit learning (speech coaching) and a different, more instrumental motivation than 'real' second-dialect acquisition. These are not either/or questions, however: motivations can be multiple, and explicitly learned skills can be internalised and thereby become more implicit (Dörnyei 2009: 151–8). Optimistically, we can state that training for performance constitutes the maximum of what can be acquired naturalistically. The instrumental motivation is not counteracted by a desire to retain a personal identity, and we can study what is trained, rather than the inevitably smaller set of skills that are internalised and retained.

Another clear value of performance data is that much of it is in the public domain, and therefore readily available and accessible. This applies not only to the performed second-dialect data, but often also to actors' first-dialect data, e.g. from other work or non-performed speech. Given these advantages of second-dialect performance data, it is worth asking what this can tell us about what is maximally possible in second-dialect acquisition, and about how such success is measured.

2.5   Factors influencing success in SDA

Compared to second-language acquisition, the literature on second-dialect acquisition in naturalistic settings is extremely small. Siegel (2010) aims to give a comprehensive overview and arrives at only seventeen studies published in English, many also on SDA in English. More recently, the field has been advanced by Nycz (2013a, 2013b, 2015, 2016). These studies give a range of different results, but it is clear that the success of SDA varies depending on individual and linguistic factors, and that SDA goes hand in hand with an increase in variability.

An individual factor that is often believed to constrain a person's ability to acquire a second dialect is the age at which they start learning the dialect: age of acquisition. The idea that pre-adolescents can acquire a faithful copy of the speech they hear around them, while post-adolescents cannot is fundamental to Labov's (2007) distinction between 'transmission' and 'diffusion': the latter, which has an element of SDA in it, often introduces changes to the original input. Labov (2007: 383) is explicit about this issue when he writes that 'it would be helpful to know more about the limitations on children's ability to learn new dialects and on adults' inability to learn them'. Work on age of acquisition within SDA, however, is much less categorical than Labov's assertion. Younger learners do appear more successful at SDA than older learners (e.g. Payne 1976, 1980; Chambers 1992; Tagliamonte & Molfenter 2007; Siegel 2010: 84–92), but the role of adolescence as a watershed period between ability and non-ability of successful SDA is much less clear; there is even some evidence that the latest age at which successful SDA can be expected is as young as 7 (Siegel 2010: 92). Studies that show successful SDA after age 14 all focus on lexicon and morphology only (Kerswill 1994; Ivars 1994; Omdal 1994). The overall evidence for the influence of

age of acquisition on success in SDA suggests that someone who starts acquiring a D2 before adulthood *will* learn *some* things, but not necessarily consistently.

Linguistic factors, too, influence success in second-dialect acquisition. Chambers (1992) formulates eight generalisations, or 'principles', on these linguistic factors. The generalisations deal predominantly with linguistic level and rule complexity; the three principles that are relevant to the current study on pronunciation are given below:

3. Simple phonological rules progress faster than complex ones.
5. In the earliest stages of acquisition, both categorical rules and variable rules of the new dialect result in variability in the acquirers.
7. Eliminating old rules occurs more rapidly than acquiring new ones.

Rule complexity appears to be the clearest linguistic predictor of success (Principle 3). A phonological rule, here, is to be understood as a generalisation of what a learner has to do to produce D2 output from the original D1 forms (or from a generative point of view, underlying D1 representations). Chambers does not give a clear metric for rule complexity, but from the examples he discusses, we can distinguish at least three types:

a. categorical substitution of phonemes;
b. change of distribution of phonemes;
c. phonemic split.

The easiest type is when a learner must categorically substitute each instance of a D1 sound with a different D2 sound. A Northern English learner of Standard Southern British English must acquire a more diphthongised FACE vowel [eɪ] instead of their D1 realisation [eː]. (For the sake of argument, these realisations are stereotypical generalisations and gloss over any variation that may exist in the D1 and D2.) Because the distribution of this vowel is identical in both accents, this is a relatively simple rule.

A more complicated type of rule is when a learner must change the distribution of two sound categories that exist in both the D1 and D2, but have different distributions. Both Northern English and SSBE have /æ/ and /ɑː/ phonemes, but the example learner would have to change the distribution of TRAP=BATH /æ/ vs PALM /ɑː/ to TRAP /æ/ vs BATH=PALM /ɑː/. This is a more difficult rule, because the substitution /æ/ > /ɑː/ only applies in some cases, but not in others.

Finally, the most complex type of rule is when a learner must acquire a D2 phonological split of a D1 phoneme. An example here is the split of Northern English STRUT=FOOT /ʊ/ into STRUT /ʌ/ vs FOOT /ʊ/. This is similar to the previous type, but the extra complication is that /ʌ/ is a *new* phoneme for the learner, not a pre-existing one as /ɑː/ was.

The easier processes can generally be expected to be acquired more successfully, although this is not a hard and fast rule.

Chambers's Principle 7 holds that it is easier to eliminate old (D1) rules than to acquire new (D2) rules. This may be the case for some of the allophonic and
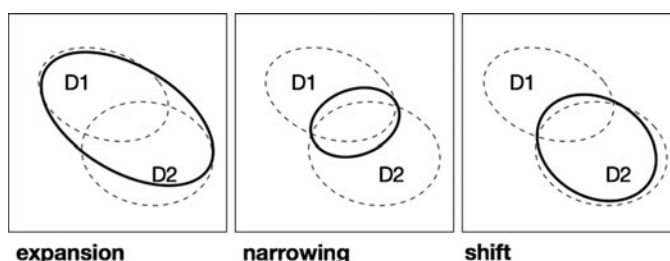
Figure 1. Schematic representation of three learner strategies for the acquisition of vowels in a second dialect

phonotactic rules in his study (*t*-flapping and rhoticity), but other examples are less convincing. Chambers sees 'low vowel merger' (LOT=THOUGHT) as a D1 rule that his speakers of Canadian English have to undo, but 'vowel backing' (TRAP≠BATH) as a new D2 rule to be acquired. On LOT=THOUGHT, Chambers (1992: 696) explicitly says that '[p]olylectally, the C[anadian] E[nglish] merger is innovative,' but surely we must assume that SDA happens in a historical linguistic vacuum — and to a synchronic acquirer of a second dialect, 'low vowel merger' and 'vowel backing' are essentially the same: acquisition of a new phoneme and redistribution of a set words over this new phoneme and the single pre-existing phoneme that was previously used in all these words. Unlike *t*-flapping and rhoticity, the figures for these two phonological splits in Chambers (1992) are not meaningfully different, suggesting that we do indeed need to look at the process of acquisition synchronically. (Compare also Nycz's (2015: 470) comment on the role of theoretical analysis in SDA.)

Finally, Principle 5 holds that SDA leads to variability in the acquirer whether or not the D2 is variable on a specific point. This is not surprising, as acquisition is inherently characterised by variable success at all stages of the acquisition process (Larsen-Freeman 1997). In the case of phonetic variables, where no two realisations of a phoneme will share the exact same acoustic properties (formant frequencies, duration, voice onset time, etc.), this means that we should expect considerable variation in SDA — and it is not clear whether this variation derives from the D1, the D2 or the acquisition process. (This is not to say that variation, including social constraints on variation, cannot be acquired successfully; see Tagliamonte & Molfenter 2007.).

In the context of accommodation, believed to be a major mechanism in SDA (Siegel 2010: 70–3), three different patterns have been identified by Bigham (2010) for how speakers deal with D1 and D2 variation. These are visualised in figure 1. Some learners showed a larger range of variation compared to their D1 to include a larger number of D2-typical realisations not possible in the D1. Other learners appeared to have restricted their D1 variation to include only those realisations that are also acceptable in the D2. Finally, some learners shifted all their realisations in the direction of the D2 while maintaining a similar range of variation. The different outcomes are likely to be

the result of different, possibly subconscious, acquisition strategies, but arguably only the latter outcome, shift, constitutes true acquisition of the D2.

3   Phonetic study

3.1   Methods

3.1.1   Informants

The speakers in the phonetic study are three young Australian actors: Tom Green, Dena Kaplan and Tim Pocock. All three had roles in the Australian television series *Dance Academy* (ABC 2010–13), and the American television series *Camp* (NBC 2013). *Dance Academy* follows a group of young dancers at a Sydney ballet school. The Australian setting is prominent in the show, with frequent images of iconic Sydney cityscapes; the actors use Australian English in this show. *Camp*, which revolves around young (trainee) staff at a summer camp, is clearly set in the United States, even though the show was recorded in Australia. There are numerous American cultural and place name references, cars drive on the right, and the actors use American English. In terms of second-dialect performance, then, Australian English is the actors' D1 and American English is their D2.

Unlike the majority of dialect performance discussed in the literature, the performed (second) dialect in *Camp* is not a non-standard variety of English. The target appears to be a standard American English designed not to add an extra layer, but to convey normality to an American audience despite an Australian cast.

Green (b. 1991) was born in Australia and grew up there, and the majority of his work is Australian, although *Camp* was not his first American production; he also plays a role in an internet-based American series. Kaplan (b. 1989) was born in South Africa and moved to Australia at age 7. Her role in *Camp* is her first American role. Pocock's (b. 1985) dialectal history, finally, is a bit more diverse, having grown up in South Africa and Ireland before moving to Australia. As with the others, the majority of his work is Australian, but Pocock has a previous North American role in the 2009 movie *X-Men Origins: Wolverine*.[1]

Although two of the three informants can be argued *not* to have Australian English as their D1, for the purposes of this article it is unproblematic to regard it as such nonetheless. The measurements for Kaplan and Pocock's D1 performance show no meaningful differences from those for Green, and also auditorily they use Australian, not South African or any other English. In both Australian and American English, the actors use a 'general' rather than a 'broad' or specific regional accent (Wells 1982: 470, 594). The actors received guidance on their American English accents from a

---

[1] Author biographies are based on Wikipedia and IMDB: https://en.wikipedia.org/wiki/Tom_Green_(Australian_actor), www.imdb.com/name/nm2964015/ (Green); https://en.wikipedia.org/wiki/Dena_Kaplan, www.imdb.com/name/nm2752988/ (Kaplan); https://en.wikipedia.org/wiki/Tim_Pocock, www.imdb.com/name/nm2992710/ (Pocock). All information current as of September 2016.

speech coach.[2] From their biographies, the actors' work in Australian and American English does not appear to have been concurrent.

3.1.2 *Audio quality*

The source of the data used in this study was digital recordings of *Dance Academy* and *Camp* in the .mp4 file format. This is a lossy form of digital compression, meaning that not all acoustic information that would be available in an uncompressed format (.wav) is retained in the audio file. Lossless extraction of the audio track from the .mp4 file to .wav allows for acoustic analysis, but the loss of data in the initial lossy conversion will still have an influence. De Decker & Nycz (2011) have shown that different recording devices and formats do affect formant measurements. However, within-device and within-format comparison is unproblematic, especially for the first two formants, which are least affected by audio conversion. The third and higher formants show greater deviations, and De Decker & Nycz (2011: 54) advise against using these measurements.

3.1.3 *Data selection and coding*

This study focuses on the actors' pronunciation of short vowels. Previous SDA research on this pair of dialects, albeit with the D1 and D2 roles reversed, focused primarily on the more salient long vowels and diphthongs GOAT, FLEECE, FACE and PRICE in addition to KIT and rhoticity (Foreman 2000; Siegel 2010: 34–6), but as successful second-dialect acquisition presumably means attaining accuracy in both salient and non-salient variables, the short vowel system should give an insight into what is possible to acquire. The short vowels were also chosen because the acoustic study originally served as a teaching tool, and working with long vowels and diphthongs would be too difficult for students of this level.

All normally stressed tokens of short vowels uttered by the three actors were isolated from eight second-season episodes of *Dance Academy* and five episodes of *Camp*.[3] To avoid breaking effects, tokens before /l/ and /r/ were excluded, also if the liquid was in the onset of the following syllable. Tokens obscured by background noise or music were not included in the data set. An overview of the number of tokens for each speaker in each lexical set is given in table 1. The first and second formants of all tokens were measured manually in Praat (v. 5.3.04, Boersma & Weernink 2012) at peak intensity using default settings.

It is clear that numbers are only sufficient to discuss the actors' performance in the KIT, DRESS, TRAP, STRUT and LOT lexical sets. The lower token counts in the other sets do allow us to make impressionistic judgements about (i) the acquisition of a different distribution of phonemes in the LOT, CLOTH and THOUGHT lexical sets in American English, and (ii) the possible retention of the TRAP/BATH differentiation from the actors'

---

[2] I thank Aviva Dassen for procuring this information from Dena Kaplan via Twitter (https://twitter.com/xAviva/status/475210740492951552, 7 June 2014; https://twitter.com/DenaAmyKaplan/status/475398714832351233, 8 June 2014).

[3] These episodes were, for *Dance Academy*, season 2, episodes 3–8, 12 and 13; for *Camp*, episodes 1, 4, 6, 8 and 9.

Table 1. *Number of tokens for each speaker, per lexical set*

	Tom Green		Dena Kaplan		Tim Pocock	
	AusEng	AmEng	AusEng	AmEng	AusEng	AmEng
KIT	55	49	37	39	31	36
DRESS	73	49	58	38	37	41
TRAP	93	83	68	53	58	54
STRUT	52	43	32	32	34	30
LOT	33	43	20	34	27	29
CLOTH	6	4	6	7	2	4
THOUGHT	6	7	5	5	1	9
BATH	7	8	6	7	5	9
FOOT	10	9	8	10	9	5

Table 2. *Reference formant values, in Hz, for five short vowels in two accents of American English (Hillenbrand et al. 1995; Hagiwara 1997) and Australian English (Cox 2006)*

	US Midwest		US California		Australia	
	F1	F2	F1	F2	F1	F2
Male speakers						
KIT	427	2,034	418	1,807	332	2,336
DRESS	580	1,799	529	1,670	467	2,085
TRAP	588	1,952	685	1,601	695	1,763
STRUT	623	1,200	574	1,415	743	1,386
LOT	768	1,333	710	1,221	584	1,040
Female speakers						
KIT	483	2,365	467	2,400	402	2,697
DRESS	731	2,058	808	2163	537	2,400
TRAP	669	2,349	1,017	1,810	886	2,014
STRUT	753	1,426	847	1,753	941	1,563
LOT	936	1,551	997	1,390	708	1,182

D1 when speaking their D2 (see below). The infrequent FOOT vowel will be ignored in the remainder of this article.

3.1.4 *Defining the D2 target*

The measure of success in this phonetic study is the similarity of D2 realisations to D2 native-speaker performance. Table 2 (visualised in figure 2) gives reference values for the first and second formants of the five vowels in the accents of three relevant areas: the Midwestern United States (Hillenbrand *et al*. 1995), Southern California (Hagiwara 1997) and Sydney, Australia (Cox 2006). The Midwestern US accent roughly represents a non-local 'General American', while the Californian data may be relevant as Kaplan's character in *Camp* attends Stanford University. The (adolescent)
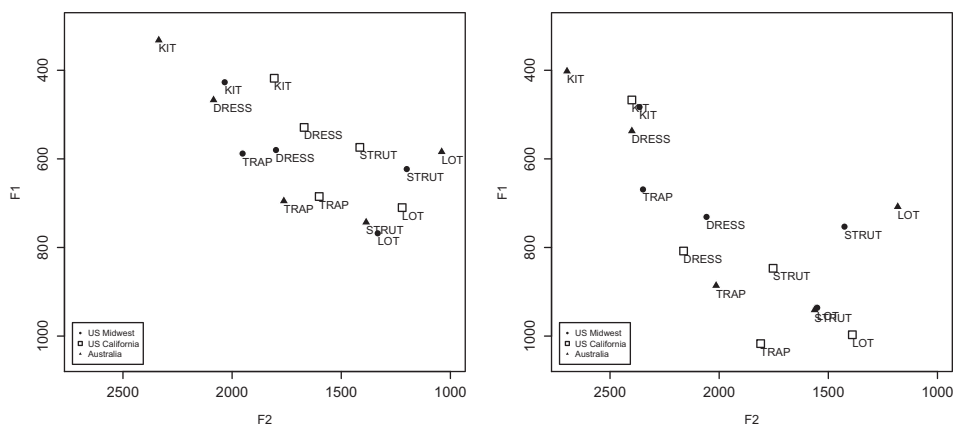
Figure 2. Reference formant values, in Hz, for five short vowels in two accents of American
English (Hillenbrand *et al.* 1995; Hagiwara 1997) and Australian English (Cox 2006). Male
speakers on the left, female speakers on the right

data for Sydney clearly function as a baseline D1 reference. Given the inter- and,
especially, intra-speaker variability in formant values, it is not fruitful to formulate
(normalised) D2 target values for F1 and F2, but the reference figures do give a clear
general idea of how the speakers need to change their speech in order to sound more
D2-like. For the five vowels in question, this involves the following changes:

KIT   AmE KIT is lower and further back than its AusE counterpart. The AmE
realisation is similar to AusE DRESS.

DRESS   Also DRESS is lower and further back in AmE than in AusE, but the lowering
and backing required stops well short of AusE TRAP.

TRAP   The changes necessary for TRAP are more complicated. Midwestern US TRAP
is higher and more fronted than the AusE vowel, which is stereotypically
thought of as raised, but the Californian vowel is further back and (for female
speakers only) considerably lower. TRAP raising appears to be spreading from
New England to other varieties of AmE (Clopper *et al.*, 2005); in California,
however, TRAP is raised only before nasals, but is lowered and backed in all
other contexts (Eckert 2008; Podesva 2011). In other words, the direction and
complexity of the necessary change depend on the exact identity of the D2.
Simultaneously, in AusE, the stereotypically high front vowels are currently
lowering in a drag chain led by TRAP (Cox & Palethorpe 2008). There is
some evidence, though inconclusive, that variables undergoing change in the
D1 are more salient, and therefore easier to acquire in the D2 (Siegel 2010:
121).

STRUT   The AmE STRUT vowel is raised compared to AusE, and is equivalent in
height to AusE LOT. The two American reference accents do not give clear
information on whether STRUT is slightly further front or slightly further back
than the AusE vowel.

Table 3. *Lexical incidence of phonemes in five lexical sets in Australian and American English (based on Wells 1982)*

	TRAP	BATH	LOT	CLOTH	THOUGHT
Australian English	/æ/	/ɑː/	/ɒ/	/ɒ/	/ɔː/
American English	/æ/	/æ/	/ɑ/	/ɔ/ ∼ /ɑ/	/ɔ/ ∼ /ɑ/

LOT AmE LOT is lower and further front than in AusE, and can more or less be equated with AusE STRUT. (Although F3 is not taken into account in this study, the speakers would need to unround their LOT vowel, suggesting an AmE LOT target very similar to AusE STRUT.)

All these concern simple substitutions of a D1 realisation by a D2 realisation and, although the phonetic near-identity of some D2 phonemes with different D1 phonemes may be a complicating factor, there is little to suggest that the actors should be unsuccessful in producing the altered sounds. We may therefore expect a relatively high measure of success. This is different, however, when it comes to the distributional differences between Australian and American English, for which we unfortunately have rather patchy D2 performance data (see table 3).

The first of these differences concerns the TRAP and BATH lexical sets. These are distinguished in Australian English, with the exception of a subset of BATH labelled DANCE by Wells (1982), which patterns with (and is coded here as) TRAP; American English does not distinguish the sets. As this change involves a merger for those with D1 Australian English, it is not expected that the actors will have difficulty accurately pronouncing the BATH set. A possible complicating factor, the sameness in Australian English of BATH and START, is unlikely to be problematic, as there is a clear orthographic cue <r> identifying the START words.

The second difference is more difficult to acquire, as it involves a redistribution of lexical sets across phonemes. In Australian English, LOT and CLOTH use the same vowel, while THOUGHT is different; in American English, CLOTH is identical with THOUGHT while LOT is different. In large areas of the United States and Canada, however, THOUGHT/CLOTH and LOT are undergoing or have undergone a merger, so that all three are /ɑ/. From an acquisitional perspective, this is clearly the easier choice.

3.2 Results

The average measurements for the three speakers in both accents are given in tables 4, 5 and 6, visualised in figures 3, 4 and 5. The tables and figures allow us to determine what kinds of changes the speakers have made to their speech in their D2; the standard deviations (ellipses in the figures) are a measure of variability in production.[4] An overview of these types of changes is given in table 7.

---

[4] Statistically significant differences between D1 and D2 formant values are indicated with asterisks in the tables, but these results may not actually offer any useful information. Acoustic production in the D2 does not need

Table 4. *Mean first and second formant frequencies (in Hz, standard deviations in brackets) for Tom Green's performance in Australian and American English*

Lexical set	F1 (AusE)	F1 (AmE)	F2 (AusE)	F2 (AmE)
KIT	456 (59)	483 (97)	2,034 (205)	1,988 (233)
DRESS	587 (82)	623 (83)	1,858 (171)**	1,773 (176)
TRAP	715 (116)	709 (127)	1,680 (214)*	1,762 (222)
STRUT	724 (97)	697 (121)	1,356 (140)	1,359 (137)
LOT	692 (78)***	775 (108)	1150 (152)***	1296 (153)

Table 5. *Mean first and second formant frequencies (in Hz, standard deviations in brackets) for Dena Kaplan's performance in Australian and American English*

Lexical set	F1 (AusE)	F1 (AmE)	F2 (AusE)	F2 (AmE)
KIT	516 (80)***	611 (128)	2,147 (289)	2,110 (231)
DRESS	701 (77)	727 (129)	2,003 (227)***	1,851 (194)
TRAP	861 (86)	901 (171)	1,831 (169)	1,799 (198)
STRUT	814 (120)	806 (97)	1,603 (128)**	1,469 (211)
LOT	794 (76)	824 (129)	1,376 (124)	1,413 (167)

Table 6. *Mean first and second formant frequencies (in Hz, standard deviations in brackets) for Tim Pocock's performance in Australian and American English*

Lexical set	F1 (AusE)	F1 (AmE)	F2 (AusE)	F2 (AmE)
KIT	436 (43)***	486 (47)	2,107 (226)***	1,920 (169)
DRESS	585 (75)	626 (96)	1,866 (131)	1,733 (167)
TRAP	725 (111)	730 (105)	1,752 (140)*	1,682 (156)
STRUT	716 (115)	663 (94)	1,385 (129)	1,385 (146)
LOT	656 (110)***	775 (98)	1,149 (130)*	1,244 (119)

Table 7. *Overview of changes made to vowel pronunciation apparent from the results*

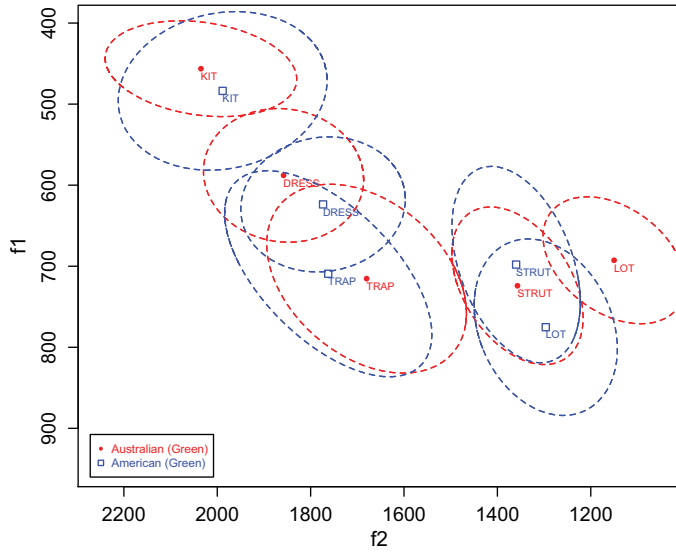|	Green	Kaplan	Pocock
KIT	expansion	shift, expansion	shift
DRESS	shift (stability)	shift, expansion	expansion
TRAP	shift (stability)	expansion	shift (stability)
STRUT	expansion	shift, expansion	narrowing
LOT	shift	expansion	shift

Figure 3. (Colour online) Mean first and second formant frequencies (in Hz) for Tom Green's performance in Australian and American English. Ellipses are drawn at one standard deviation from the means.
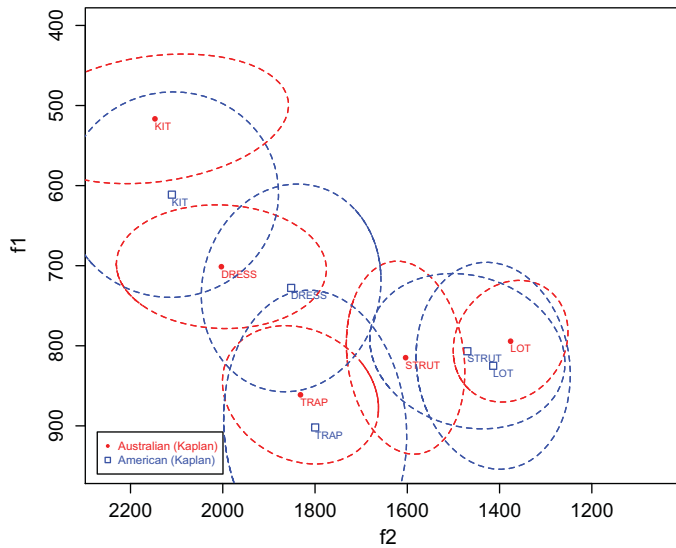

Figure 4. (Colour online) Mean first and second formant frequencies (in Hz) for Dena Kaplan's performance in Australian and American English. Ellipses are drawn at one standard deviation from the means.
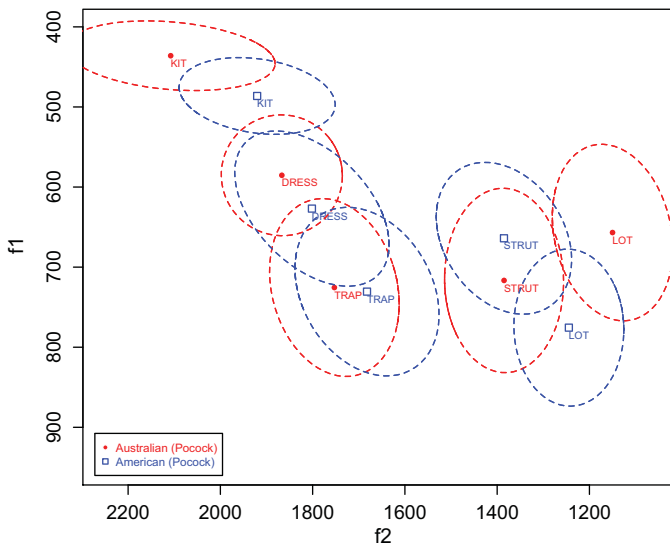
Figure 5. (Colour online) Mean first and second formant frequencies (in Hz) for Tim Pocock's performance in Australian and American English. Ellipses are drawn at one standard deviation from the means.

All three speakers have changed all five vowels in the expected direction, and in that sense have done what is necessary for an AmE performance. The degrees of difference, however, differ between speakers. Green's mean formant values in the D2 remain relatively close to those in the D1, with the exception of LOT. Kaplan makes much clearer changes, but is less stable in her AmE as the standard deviations for her D2 vowels are much larger than those for her D1 vowels. Pocock shows a more stable D2 that also visually is more clearly distinguished from his D1.

Figures 6, 7 and 8 allow us to examine the phonological changes the speakers have to apply in an American English accent, although some of the speakers have very few tokens for the relevant lexical sets and the conclusions are necessarily tentative. The first of the phonological changes is the (synchronic, acquisitional) merger of the TRAP and BATH lexical sets, and (in the case of a Californian English target) the introduction of allophonic raising of /æ/ before nasals. These data are presented in the left panels of the relevant figures. It is clear that all three speakers have successfully merged the TRAP and BATH sets, but evidence for the allophonic raising pattern is less clear. Kaplan's data cannot give evidence for the pattern, as all TRAP and BATH tokens in her data were pre-nasal. Green appears to have an incipient nasal pattern: although the high and front realisations are always pre-nasal, not all pre-nasal /æ/ tokens are high and front. Pocock

to be statistically significantly different from that in the D1 for it to be meaningfully different in perception, and vice versa. As the Australian realisations for our speakers were quite different from those in Cox (2006) (table 2), it is also unlikely that a comparison of the D2 realisations to the American reference values would yield interpretable results.
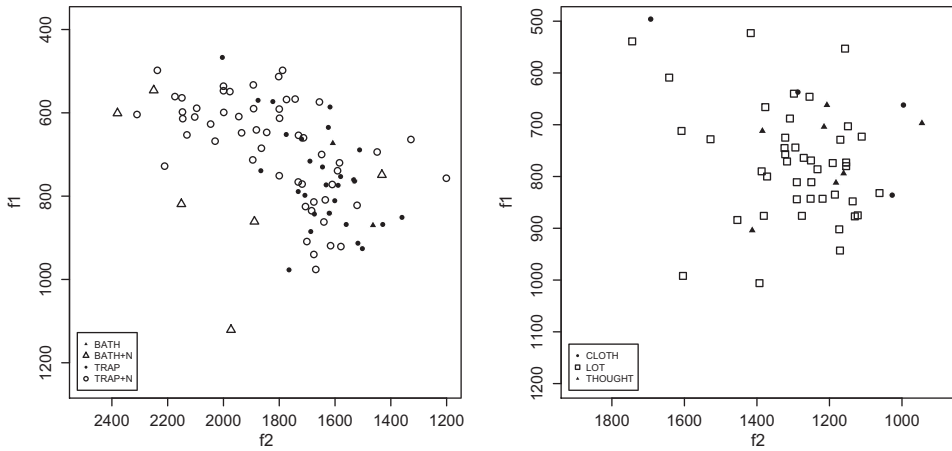
Figure 6. Tom Green's realisations of the ʙᴀᴛʜ and ᴛʀᴀᴘ lexical sets (left) and the ᴄʟᴏᴛʜ, ʟᴏᴛ and ᴛʜᴏᴜɢʜᴛ lexical sets (right) in American English

Figure 7. Dena Kaplan's realisations of the ʙᴀᴛʜ and ᴛʀᴀᴘ lexical sets (left) and the ᴄʟᴏᴛʜ, ʟᴏᴛ and ᴛʜᴏᴜɢʜᴛ lexical sets (right) in American English.

does show a clear nasal pattern, although he only employs the front–back dimension, and not the height dimension as well.

The other phonological change concerns the ʟᴏᴛ, ᴄʟᴏᴛʜ and ᴛʜᴏᴜɢʜᴛ lexical sets, shown in the right panels of figures 6, 7 and 8. Here especially, low token counts obscure the picture somewhat. Green's American English appears to have the innovative three-way merger between these sets. Kaplan and Pocock's American English performances do so, too, but a bit less clearly. In Kaplan's data, the more fronted realisations are ʟᴏᴛ, perhaps showing the traditional American pattern. Pocock's data, on the other hand, show that if anything, ᴛʜᴏᴜɢʜᴛ may be a bit further back than the other two sets, which is more in line with his D1 Australian English pattern.

Figure 8. Tim Pocock's realisations of the BATH and TRAP lexical sets (left) and the CLOTH, LOT and THOUGHT lexical sets (right) in American English
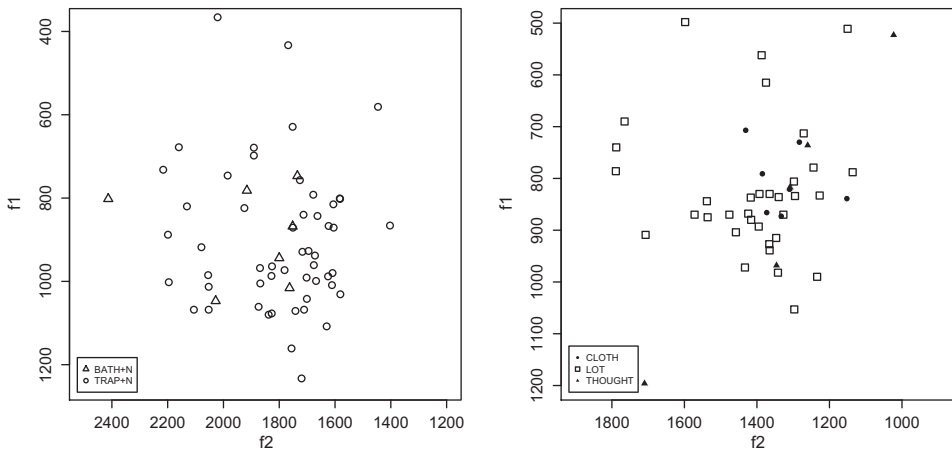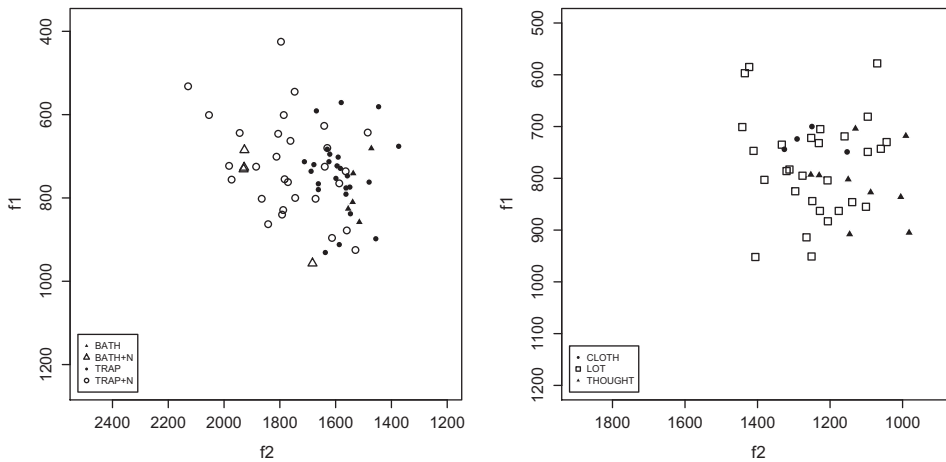
3.3   Discussion

These results suggest that the three actors differ in their success at performing in American English. Based on the direction of changes made and the relative size of standard deviations in the D2 compared to the D1, Pocock seems to be most successful of the three in his American English performance. While Green is consistent in his D2 speech, he has not changed much from his D1; Kaplan does clearly make the required changes, but her D2 is much less consistent. In terms of Bigham's accommodation types (table 7), the speakers show different patterns. Where Kaplan seems to rely predominantly on strategies resulting in expansion, Pocock uses a range of strategies that may ultimately allow for greater success. The accommodation types do not appear to pattern according to vowel category. We would need a much clearer idea of the range of variation in the target accent to be able to say what strategy would be the best choice. This would most likely be a different strategy for different vowels, and given inter-individual variation even in the D1, also for different speakers.

The expected difference in acquisition between phonetic and phonological changes is borne out by the data. The merger of TRAP and BATH, expected to be easy to acquire, was indeed acquired without any problems, while the change in distribution of vowels over the LOT, CLOTH and THOUGHT sets was acquired much more variably. The split of a phoneme, as is required with the nasal pattern in /æ/, is thought to be most difficult but is acquired with reasonable success, to the extent that the data allow us to investigate this.

4   Foreign accent rating

If the acoustic study shows moderately successful second-dialect performance by the three Australian actors, at least in the short vowel system, the next question is whether their performance is acceptably American for their predominantly American audience.

Table 8. *Distractor items in the foreign accent rating*

Speaker	Home state	Series	Episode (Year)
Carla Gugino	Florida	*Political Animals*	s01ᴇ04 (2012)
Paul James	DC	*Greek*	s01ᴇ06 (2007)
Taylor Lautner	Michigan	*Cuckoo*	s02ᴇ03 (2014)
Julia Louis-Dreyfus	New York	*Veep*	s02ᴇ01 (2013)
Meaghan Jette Martin	Nevada	*10 Things I Hate About You*	s01ᴇ14 (2010)
Jason Schwartzman	California	*Bored to Death*	s01ᴇ05 (2009)

To investigate this, a foreign accent rating (FAR) experiment was performed in which American native-speaker judges rated the actors' Americanness.

4.1   Methods

The FAR followed the methodology presented in Schmid & Hopp (2014: 369–74). Six short (10–15 seconds) fragments of American-accented speech – two for each speaker – were selected as test items. The fragments had to show a relatively uninterrupted stretch of speech by a single speaker (short interruptions by others were removed from the audio file) and a minimum of background noise. The fragments were not controlled for the presence of the five vowels investigated in the acoustic study. Six similarly selected fragments of actors that are native speakers of American English (table 8) functioned as controls. Focusing on American-accented speech and excluding Australian-accented speech from the FAR resulted in a relatively narrow range of accents across test and control items, so that the results would not be distorted by low ratings for clearly non-American accents. This was done as the focus in this article is on the accuracy and acceptability of these performed D2s *qua* second-dialect acquisition. In the television series, by contrast, a single actor continued to use his (broad) Australian English D1. This had the effect that the other actors' D2 American English are assessed as more acceptably American, as they are explicitly contrasted with an accent that is very clearly *not* American. This suggestion to the audience is even stronger as the Australian-speaking character was the (somewhat unsympathetic) leader of a competing summer camp, meaning that Australian English becomes explicitly 'othered' vis-à-vis even imperfect American English accents.

The fragments were presented in a fixed order in a single audio file, with each fragment being introduced by a number spoken by a female voice in an American English accent and a five-second pause between each fragment. The instruction to survey participants explicitly mentioned variation in accents of English and named American English as an identifying characteristic of most Americans. Exactly what constitutes an American English accent, however, was left to the participants to decide. For each fragment, raters were asked to answer two multiple-choice questions:

- Do you think Speaker […] is American?
  – yes
  – no

- How certain are you of your answer?
  - very certain
  - reasonably certain
  - not certain at all

The answers to these questions were converted into a 6-point Likert scale, where 1 is 'definitely not American' and 6 is 'definitely American'. The advantage of this method over direct judgements on a Likert scale is that it forces judges to make a decision instead of choosing one of the middle options on the scale (see Schmid *et al.* 2014: 374).

As Hodson (2014: 231) suggests, knowledge of actor background may influence judgements of their performed accents. The question as it was phrased in the FAR does suggest that some speakers may not be American, but offer no further information. If the task was to identify the Australians speaking a second-dialect version of American English, the judgements might have been different.

Respondents were also asked about their own linguistic backgrounds, in particular about their exposure to varieties other than American English, and whether they could identify the speakers in the judgement task.

4.2   Results

The FAR was distributed via social media in November 2014 using the 'friend-of-a-friend' approach. After removing incomplete responses, responses from judges who spent a considerable time outside the USA, and responses from judges who indicated that they recognised one of the speakers (even if they identified the speaker incorrectly), a total of 154 responses remained. The means of their responses, with standard deviations, are given in table 9.

Table 9.   *Mean scores with standard deviations in the Foreign Accent Rating*

Speaker	Type	Mean	St. dev.
Meaghan Jette Martin	control	5.422	0.995
Carla Gugino	control	5.247	1.195
Julia Louis-Dreyfus	control	5.142	1.228
Tim Pocock (1)	test	5.058	1.238
Paul James	control	4.968	1.398
Tom Green (1)	test	4.929	1.314
Tim Pocock (2)	test	4.844	1.442
Dena Kaplan (1)	test	4.799	1.492
Tom Green (2)	test	4.708	1.490
Taylor Lautner	control	4.286	1.752
Jason Schwartzman	control	4.240	1.657
Dena Kaplan (2)	test	3.942	1.761

It is immediately obvious from the table that the ratings for the Australian test items were not as high as those for the American controls. This is not a hard and fast rule, however: it was especially the female Americans (D1) that received high ratings, while two of the male Americans (D1) scored a good deal lower than most of the Australian (D2) test items. The second fragment from Kaplan was rated especially poorly. With the exclusion of that fragment, however, the test items from the Australian actors had mean scores between 4.708 and 5.058, meaning that they were classified as 'American' with reasonable certainty. It is also possible to establish a hierarchy of relative success: Pocock performs best, followed by Green and Kaplan.

Judges were more in agreement about speakers who *did* sound American than about speakers who sounded less so: the standard deviations for the higher scores are much smaller than those for the lower scores. This negative correlation ($\rho = -0.963$) between means and standard deviations is highly significant ($p = 4.9 \times 10^{-7}$), but standard deviations are likely to be smaller at the extremes of the scale than in the middle.

4.3   Discussion

Two findings stand out from the FAR. Firstly, the Australian speakers of American English were found, on average, to sound less American than the American speakers. A FAR is a holistic rating, and this was not a carefully controlled matched-guise experiment. It is not clear what features, if any, respondents used to come to their assessment of the Americanness of the clips. In this sense, a FAR is a black box similar to the responsive aspect of accommodation (Coupland 1984: 50–1). Hayes-Harb & Hacking (2015) show on the basis of interviews with accent judges that although they use phonetic information (such as segments, rhythm, stress, fluency and intonation) in their judgements, they also comment on imagined social types with no explicit link to phonetic information. While the hierarchy of the actors' judged Americanness (Pocock, Green, Kaplan) matches that of their acoustic accuracy, it is not certain that the accent judges attended to any or all of the short vowels from the acoustic study.[5]

The other finding is that the ratings for the American speaker controls are very clearly gendered.[6] Earlier work on American English suggests female voices are perceived as more standard (northern) than an equivalent male voice (Plichta & Preston 2005: 121, 123), but while this might explain why the female controls are rated as 'better' than the male controls, it does not explain why Pocock and Green are rated higher than Kaplan. Alternatively, what it means to 'sound American' may differ for men and women; if that is the case, Pocock and Green are judged much more similar

---

[5] Correlations were performed post hoc between the FAR rating and measures of pitch (mean pitch, absolute and relative pitch range), vowel quality (F1 and F2 of the five short vowels in Hz and Bark) and rhythm (nPVI measurements; Thomas & Carter 2006). Significant (positive) correlations were found only for F2 DRESS and F2 TRAP, both in Hz and Bark, but as this is not consistent with the reference values from table 2 and figure 2, these findings are likely to be false positives.

[6] An exception to the gendered result is Paul James, whose character in *Greek* was gay and whose performed speech may therefore have been perceived as less stereotypically masculine; see, e.g., Levon 2006.

to the male target(s) than Kaplan is to the female target(s), and the hierarchy of success is even more extreme.

5 General discussion and conclusion

This article has presented an analysis of second-dialect performance by three Australian actors. It was suggested that this analysis can give a realistic impression of what is possible in second-dialect acquisition, as this type of performance mitigates a number of factors that may stand in the way of acquisition: issues of access, analysis and motivation. The acoustic analysis of five short vowels shows that the actors changed their mean vowel productions in the required direction, but resulting in different patterns of variation. Although there is no quantifiable measure of success in the absence of a clearly defined target, the actors appear reasonably successful at second-dialect performance at the phonetic and, to perhaps a slightly lesser degree, at the phonological level. A subjective, holistic assessment of their success in a foreign-accent rating did suggest the actors were found out as non-Americans – or at least as not stereotypically American. Some of the male American speakers used as controls, however, scored even lower on the FAR, and it is uncertain what the judges considered to be an American accent and whether (some of) the features they responded to were the same as those investigated in the acoustic study.

All in all, this study suggests that, at least on the surface, second-dialect performance is similar to second-dialect acquisition in terms of the outcomes of changes made to vowel pronunciation. Given the practical challenges related to longitudinal phonetic studies of second-dialect acquisition — issues of access and being able to control for factors influencing the process — performance may be a very useful proxy for studies of the limits and the mechanisms of accent change. Second-dialect performance is frequent, controlled, and in any case the performances are mediated and therefore much more readily available than data from naturalistic SDA.

The study raises two other points for further research into accuracy and acceptability of second-dialect acquisition and performance. Firstly, acoustic accuracy on a range of features may not be a reliable measure of success, because these features may not be salient or meaningful for listeners' judgements. Secondly, it is not evident in what way detailed (acoustic) and holistic (acceptability judgement) measurements of the degree of an acquirer's or performer's success are related. In this study, discrepancies between both measures are probably for a large part due to subjective judges attending to a range of different accent features, which may or may not include features from the acoustic study. While many studies of second-dialect acquisition (Siegel 2010) have focused on acoustic accuracy, subjective success is perhaps more important in real-life acquisition as the goal there is highly motivated by a desire for acceptance by the community. It therefore seems prudent to include both types of measures of success in future studies of second-dialect acquisition and performance.

Author's address:

Rijksuniversiteit Groningen
Faculteit der Letteren
English Language & Culture
Postbus 716
9700 AS Groningen
The Netherlands
r.m.knooihuizen@rug.nl

References

Beal, Joan C. 2009. 'You're not from New York City, you're from Rotherham': Dialect and identity in British indie music. *Journal of English Linguistics* 37(3), 223–40.

Bell, Allan. 1992. Hit and miss: Referee design in the dialects of New Zealand television advertisements. *Language & Communication* 12(3–4), 327–40.

Bell, Allan & Andy Gibson. 2011. Staging language: An introduction to the sociolinguistics of performance. *Journal of Sociolinguistics* 15(5), 555–72.

Bigham, Douglas S. 2010. Mechanisms of accommodation among emerging adults in a university setting. *Journal of English Linguistics* 38(3), 193–210.

Boersma, Paul & David Weenink. 2012. *Praat: Doing phonetics by computer (version 5.3.04)*. www.praat.org

Bucholtz, Mary. 2003. Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics* 7(3), 398–416.

Bucholtz, Mary. 2011. Race and the re-embodied voice in Hollywood film. *Language & Communication*, 31, 255–65.

Chambers, J. K. 1992. Dialect acquisition. *Language* 68(4), 673–705.

Clopper, Cynthia G., David B. Pisoni & Kenneth de Jong. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America* 118(3), 1661–76.

Coupland, Nikolas. 1984. Accommodation at work: Some phonological data and their implications. *International Journal of the Sociology of Language* 46, 49–70.

Coupland, Nikolas. 2001. Dialect stylization in radio talk. *Language in Society* 30(3), 345–75.

Coupland, Nikolas. 2009. The mediated performance of vernaculars. *Journal of English Linguistics* 37(3), 284–300.

Coupland, Nikolas. 2011. Voice, place and genre in popular song performance. *Journal of Sociolinguistics* 15(5), 573–602.

Cox, Felicity. 2006. The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers. *Australian Journal of Linguistics* 26(2), 147–79.

Cox, Felicity & Sallyanne Palethorpe. 2008. Reversal of short front vowel raising in Australian English. *Interspeech* 2008, 342–5.

De Decker, Paul & Jennifer Nycz. 2011. For the record: Which digital media can be used for sociophonetic analysis? *University of Pennsylvania Working Papers in Linguistics* 17(2), 51–9.

Dörnyei, Zoltán. 2009. *The psychology of second language acquisition*. Oxford: Oxford University Press.

Eckert, Penelope. 2008. Where do ethnolects stop? *International Journal of Bilingualism* 12(1–2), 25–42.

Foreman, Annik. 2000. A longitudinal study of Americans in Australia. In Keith Allen & John Henderson (eds.), *Proceedings of ALS2k, the 2000 Conference of the Australian Linguistics Society*. www.als.asn.au/proceedings/als2000/foreman.pdf

Hagiwara, Robert. 1997. Dialect variation and formant frequency: The American English vowels revisited. *Journal of the Acoustical Society of America* 102(1), 655–8.

Hayes-Harb, Rachel & Jane Hacking. 2015. Beyond rating data: What do listeners believe underlies their accentedness judgments? *Journal of Second Language Pronunciation* 1(1), 43–64.

Hillenbrand, James, Laura A. Getty, Michael J. Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5), 3099–111.

Hodson, Jane. 2014. *Dialect in film and literature*. Basingstoke: Palgrave Macmillan.

Ivars, Ann-Marie. 1994. Bidialectism and identity. In Bengt Nordberg (ed.), *The sociolinguistics of urbanization: The case of the Nordic countries*, 203–22. Berlin: De Gruyter.

Kerswill, Paul. 1994. *Dialects converging: Rural speech in urban Norway*. Oxford: Clarendon Press.

Labov, William. 2007. Transmission and diffusion. *Language* 83, 344–87.

Larsen-Freeman, Diane. 1997. Chaos/complexity science and second language acquisition. *Applied Linguistics* 18(2), 141–65.

Le Page, R. B. & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

Levon, Erez. 2006. Hearing 'gay': Prosody, interpretation, and the affective judgments of men's speech. *American Speech* 81(1), 56–78.

Nycz, Jennifer. 2013a. Changing words or changing rules? Second dialect acquisition and phonological representation. *Journal of Pragmatics*, 52, 49–62.

Nycz, Jennifer. 2013b. New contrast acquisition: Methodological issues and theoretical implications. *English Language and Linguistics* 17(2), 325–57.

Nycz, Jennifer. 2015. Second dialect acquisition: A sociophonetic perspective. *Language and Linguistics Compass* 9(11), 469–82.

Nycz, Jennifer. 2016. Awareness and acquisition of new dialect features. In Anna Babel (ed.), *Awareness and control in sociolinguistic research*, 62–79. Cambridge: Cambridge University Press.

Omdal, Helge. 1994. From the valley to the city: Language modification and language attitudes. In Bengt Nordberg (ed.), *The sociolinguistics of urbanization: The case of the Nordic countries*, 116–48. Berlin: De Gruyter.

Payne, Arvilla C. 1976. The acquisition of the phonological system of a second dialect. PhD thesis, University of Pennsylvania.

Payne, Arvilla C. 1980. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In William Labov (ed.), *Locating language in time and space*, 143–78. New York: Academic Press.

Plichta, Bartłomiej & Dennis R. Preston. 2005. The /ay/s have it: The perception of /ay/ as a North-South stereotype in United States English. *Acta Linguistica Hafniensia* 37(1), 107–30.

Podesva, Robert J. 2011. The California vowel shift and gay identity. *American Speech* 86(1), 32–51.

Schmid, Monika S. & Holger Hopp. 2014. Comparing foreign accent in L1 attrition and L2 acquisition: Range and rater effects. *Language Testing* 31(3), 367–88.

Schmid, Monika S., Steven Gilbers & Amber Nota. 2014. Ultimate attainment in late second language acquisition: Phonetic and grammatical challenges in advanced Dutch-English bilingualism. *Second Language Research* 30(2), 129–57.

Siegel, Jeff. 2010. *Second dialect acquisition*. Cambridge: Cambridge University Press.

Simpson, Paul. 1999. Language, culture and identity: With (another) look at accents in pop and rock singing. *Multilingua* 18(4), 343–67.

Tagliamonte, Sali A. & Sonja Molfenter. 2007. How'd you get that accent? Acquiring a second dialect of the same language. *Language in Society* 36(5), 649–75.

Thomas, Erik R. & Phillip M. Carter. 2006. Prosodic rhythm and African American English. *English World-Wide* 27(3), 331–55.

Trudgill, Peter. 1983. Acts of conflicting identity: The sociolinguistics of British pop-song pronunciation. In *On dialect: Social and geographical perspectives*, 141–60. Oxford: Blackwell.

Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.