University of Reading

English intonation in storytelling: a comparison of the recognition and production of nuclear tones by British and Hong Kong English speakers

Article

Accepted Version

Hudson, T., Setter, J. ORCID: https://orcid.org/0000-0001-7334-5702 and Mok, P. (2022) English intonation in storytelling: a comparison of the recognition and production of nuclear tones by British and Hong Kong English speakers. English World-Wide, 43 (3). pp. 357-381. ISSN 0172-8865 doi: 10.1075/eww.21035.hud Available at https://centaur.reading.ac.uk/100443/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1075/eww.21035.hud

Publisher: John Benjamins

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

English intonation in storytelling: a comparison of the recognition and production of nuclear tones by British and Hong Kong English speakers

Toby Hudson (University of Oxford, UK), Jane Setter (University of Reading, UK), and Peggy Mok (Chinese University of Hong Kong)

Abstract

This paper presents data for a tightly controlled recognition and production study of English language intonation in reading by speakers of British English and second language learners of English in Hong Kong. We demonstrate a relatively high correlation between the scores for the two studies when data are separated by utterance type (statement, echo, WH-question, etc.). Our finding that this cohort of English learners performs better at production of nuclear tones than in the corresponding recognition study when both are judged by a template for British English adds support to the claim that the perception-production link, a theory that production is contingent on perception, is not borne out by the empirical study of learners of World Englishes. Data collected for the British English speakers give insight into a changing intonational phonology, while Hong Kong data indicate differences in intonational categories, a different distribution of tones, and possibly tonal innovation.

Keywords: intonation, production, perception, L2, Hong Kong English, emergent phonology, nativization

1. Introduction

Intonation, often referred to as the melody of speech, is an intrinsic part of the linguistic system of languages which make use of it. In English, for example, it is used in turn-taking, to manage conversational behaviour, to indicate grammatical meaning, to draw the listener's attention to new or important information, and to show various affective meanings, such as emotion or a speaker's attitude towards events present in the conversational context. It is one of the earliest-acquired aspects of a linguistic system for first language (L1) speakers (Berkovitz 1980; Wermke et al. 2016), and is phonemic, meaning the variations in intonation patterns used are particular to and meaningful in a

given language. It cannot, therefore, be presumed that another language with similar patterns uses them for the same linguistic meanings.

Intonation in English has been comprehensively described and categorised in a number of texts and research studies. From a pedagogical point of view, and as far as British English is concerned, books such as O'Connor and Arnold (1973), Brazil (1994), Bradford (1988), Tench (2015) and Wells (2006) offer approaches to teaching and learning intonation which attempt to capture its major uses and offer plenty of contextualised practice activities. However, one only has to glance at the list of possible meanings for one of O'Connor and Arnold's (1973) ten tunes to realise intonation is not as straightforward or simplistic as pedagogic texts might have one believe. For example, the pattern described as "The Switchback" – basically, a fall-rise – is given the following set of attributes (ibid.: 170):

> In STATEMENTS: grudgingly admitting, reluctantly or defensively assenting, concerned, reproachful, hurt, reserved, tentatively suggesting; (in echoes) greatly astonished.
> In QUESTIONS: (in echoes) greatly astonished; otherwise, interested and concerned as well as surprised.
> In COMMANDS: urgently warning with a note of reproach or concern.
> In INTERJECTIONS: scornful.

This may explain why intonation has a reputation for being one of the most difficult aspects of a second language (L2) to learn and teach; with so many meanings attached to one pattern, how can a learner be sure he or she is making the right meaning? Chun (1998: 61), for example, claimed that intonation is "seemingly impossible" to learn, and Taylor (1993: 2) explained that teachers believe it is "not teachable, and not learnable either". The fact that it has this reputation can put English language teachers off from venturing into this area in the first place. Pronunciation *per se* is often

neglected in English language teaching; intonation, therefore, is rarely tackled, and research on this area is scarce (Derwing & Munro, 2015).

There is ample evidence that the intonation patterns of L2 speakers of English vary in comparison with L1 speakers. Previous research, however, usually focuses on production differences or meaning interpretation by L2 speakers (e.g. Grabe et al. 2003; Setter, Wong and Chan 2010), with little attempt to ascertain whether L1 and L2 speakers know what the appropriate intonation patterns should be in a particular context and are able to use them. Furthermore, it is often assumed that L2 learners' patterns are different simply because of their L1 backgrounds. Are the differences in intonation patterns produced by L2 speakers due to L1 influence, or does the lack of knowledge and understanding about which patterns are appropriate also have a role to play?

This paper presents evidence from two studies, one on perception and one on production, to investigate British English (BrE) and Hong Kong English (HKE) speakers' knowledge and use of intonation patterns in a story-telling task. We start by looking at whether listeners are able to select suitable intonation patterns in certain contexts throughout the story before investigating their production when reading out the narrative passage. While results have been disseminated for the perception and production experiments individually (Mok et al. 2016; Hudson, Setter and Mok 2019), we now compare the data from the two studies to see whether there is any relationship between the perception and production of intonation patterns in the two contexts.

While there are intuitive grounds for assuming a connection or a correlation between the perception and production capabilities of an individual, and this notion is supported by a prevalent model of speech learning (Flege 1995), it has been challenged, for instance, by every evidential instance of perception without corresponding production (Isbell 2016). Additionally, we might suppose that the picture would be more complex in L2 acquisition of prosody, since intonational categories such as fall, rise and level pitch are universal, but the function of the categories is language-specific, not to

mention the interplay of the linguistic and paralinguistic functions of intonation. Puga et al. (2018) demonstrated – using the same procedure to test listeners' understanding of the suitability of tones in various contexts – that German learners of English performed significantly better in production than in what we are referring to as perception, i.e. at odds with the traditional assumptions of a perception-production link. In the present study we are interested to see whether the same pattern can be seen with the Hong Kong learners of English.

We have chosen these two varieties of English for various reasons. First, HKE is an emergent variety of Post-colonial English (Setter, Wong and Chan 2010; Schneider 2007), but very little work has been done on the intonation of HKE. Second, English is learned from an early age in Hong Kong, but the curriculum has little, if any, emphasis on English prosody (stress, rhythm, intonation). While there are L1 English teachers from countries such as Britain, Australia and the United States operating in many schools and kindergartens in Hong Kong, meaning there will be exposure to different intonation patterns in conversation, the lack of formal instruction means HKE speakers may not develop a knowledge of how intonation patterns are used in various linguistics contexts, whether this knowledge is intrinsic or extrinsic. Third, intonation in the L1 of most HKE speakers – i.e. Cantonese – operates very differently. Typologically, Cantonese is a tone language, usually described as having six lexical tones, three of which have shorter allotones (Matthews and Yip 2013). There is intonation in Cantonese, but intonational pitch variation is constrained to preserve lexical tones and not used to indicate attitudinal or discourse meanings to the same extent as in English. Instead, Cantonese makes use of a rich inventory of sentence-final particles (Matthews and Yip 2013). It is therefore of interest to see if HKE speakers notice which English intonation patterns are appropriate in different sentential and attitudinal contexts in a story-telling context and can identify and use them, and also whether there are consistent emerging patterns in the variety.

Where BrE is concerned, we are interested to see whether the predicted intonation patterns, based on the literature, are in fact selected and used by this population, or whether speakers do something else.

It is possible, for example, that uptalk (Warren 2016) may affect speaker choices, i.e. that speakers might select or use a rising intonation in some contexts where the more traditional models predict falling patterns, based on older varieties of BrE. The present methodology will enable us to test the validity of existing descriptions as far as use in a story-telling context is concerned.

Previous studies on intonation in HKE have shown that there is an over-generalisation of intonation patterns associated with the nuclear tone; i.e. speakers tend to use fewer tones in more contexts than BrE speakers. For example, Bolton and Kwok (1990) reported that HKE speakers would apply a falling tone to all statements, and a rising tone to all questions. However, Setter, Wong and Chan (2010) observed that, in map-task and conversational contexts, there was an increase in the use of the rising tone in statements, and Cheng, Greaves and Warren (2008) and Lin (2008) argued that, while most questions had a rising tone in HKE, WH-questions are produced with a falling tone. It has also been demonstrated that the frequency of use of different tone types in HKE is different from BrE. In BrE, Cruttenden (2014) indicates that the falling tone is the most common (50%), followed by the rise and fall-rise (40% together). Mompeán-Gonzáles and Monroy-Casas (2010) reported the following proportions of nuclear tones in British television commercials: 75% falling, 13% rising and 2% falling-rising, the remaining 10% being level and rise-fall. These findings are of interest, but should be treated with caution, as the speech in television commercials is far from natural. In HKE, both Setter, Wong and Chan (2010) and Cheng, Greaves and Warren (2008) found that the level tone was the most common (49.3% and 43.3% respectively). This may be because the speech collected was more natural than that used in Cruttenden's (2014) or Mompeán-Gonzáles and Monroy-Casas's (2010) studies; level tones in natural speech are common and often associated with hesitation phenomena (Cauldwell, 2002). Cheng, Greaves and Warren (2008) reported that the second and third most common tones are the fall (37.1%) and the rise (5.7%), whereas Setter, Wong and Chan's (2010) data found that the rise was the second most common tone (24.4%), with the fall coming a close third (23%).

There is a lack of fall-rise tones in HKE in comparison with BrE; Setter, Wong and Chan (2010) reported only 2.71% in their data. It is commonly accepted that, in BrE conversational contexts, where it is categorized by Brazil (1994) together with the rise as a "referring tone" (Brazil 1994), the fall-rise is the default referring tone, used to indicate given information and polite requests. It is also used to show disagreement in a more face-saving way than simply using a fall (Brazil, Coulthard and Johns 1980). In American English, however, this tone is much less high-incidence (Bolinger 1998). One suggestion is that speakers of Post-colonial varieties deliberately avoid using it in order to distance themselves from their former colonial rulers (David Deterding, p.c., August 2011). This shows a level of awareness of the use of intonation patterns which might be difficult to attest.

The rise-fall is uncommon in BrE – Mompeán Gonzáles and Monroy Casas (2010) only found that 3% of nuclear syllables in speech from television commercials involved this tone – but it is hardly ever used in HKE. Cheng, Greaves and Warren (2008) reported 0.016% of occurrences. Setter, Wong and Chan (2010: 40) found only 0.59% of tones were rise-fall, and noted that, unlike BrE, the use of the rise-fall in HKE was not associated with indignation, sarcasm, surprise or being impressed (Roach 2009), but simply to add emphasis. It may be the case, then, that HKE speakers simply have no knowledge of the function of the rise-fall in other varieties of English such as BrE. The lack of knowledge of certain functions of intonation in BrE may also partly explain other differences between the HKE and BrE speakers mentioned above, but, so far, few studies have investigated whether L2 learners have explicit awareness of what the appropriate intonation patterns are to signal meaning effectively in English discourse.

Why is it important to study intonation? Where international communication in English is concerned, Jenkins' Lingua Franca Core (2000) considers intonation patterns, or tunes, to be unimportant, but nucleus placement – also known as sentence focus and tonicity – is identified as essential for clearly indicating meaning. It is likely that the global influence of varieties such as British and American English is declining in favour of Post-colonial and learner varieties, and so the use of intonation in

older varieties may be losing its relevance. By looking at the intonation choices of English language speakers from different backgrounds, we can learn more about how meaning is made in different varieties, and this in turn can help us to be more sensitive to usage differences, and not assume that a speaker using a pattern one variety associates with a negative affective meaning realizes this is how an utterance is being perceived.

In order to investigate HKE and BrE speakers' understanding and use of intonation patterns in a controlled narrative context, we adopted two methods. To look at understanding, we followed He, van Heuven and Gussenhoven (2012) and conducted a forced-choice intonation patterns selection task. This allows us to directly assess whether HKE and BrE speakers have knowledge of the appropriate intonation patterns in given narrative situations throughout the task. To look at the use of intonation patterns in a narrative task, we collected data from a randomly-selected subset of the participants in the selection task reading the same story as used in the selection task. We then compared the results from the studies to see if there was any relationship between the speakers' knowledge of the appropriate patterns to use at various points in the narrative with their own production choices. Our research questions were as follows:

1. Are the patterns of perception and production of intonation in HKE and BrE speakers in a story-telling task similar to published descriptions?
2. What is the relationship between the production and perception of intonation in HKE and BrE speakers respectively?

Hypotheses:

1a. The patterns of perception and production among HKE speakers in a story-telling task will be dissimilar from published descriptions of English intonation.

1b. The patterns of perception and production among BrE speakers in a story-telling task will be similar to published descriptions of English intonation.

2a. There will be no relationship between perception and production performance among HKE speakers.

2b. There will be a strong relationship between perception and production performance among BrE speakers.

It should be noted from the outset that we are looking at speech in a very narrow context, i.e. story-telling. Our findings can inform research into intonation but they will be confined to intonation in story narratives, like Mompeán-Gonzáles and Monroy-Casas's (2010) work is confined to intonation in acted television commercials. We have selected story-telling as our domain for several reasons, including that it is relatively easy to control and to predict patterns of intonation as a point of reference for BrE, based on existing descriptions based on well-established traditions, such as O'Connor and Arnold (1973) and Wells (2006).

2. Method for Study 1: Recognition of tonal suitability

2.1 Participants

We recruited 40 (5 male, 35 female) speakers of Hong Kong Cantonese, aged 17 to 25 years old (mean age 19.25 years; St Dev ±1.72 years). They were all university students in Hong Kong and they participated in the experiment for course credit. Nine of the participants had experience of living abroad for no more than one year (one month to 12 months), and none reported speech or hearing difficulties. All participants had learned English as a second language since childhood. 30 of them had the highest grades (5**/5*) in English in the Hong Kong Diploma of Secondary Education Examination (the public entrance examination for Hong Kong universities), which are equivalent to IELTS 7.51-7.77 and 7.16-7.32, respectively, according to the official conversion table produced by the Hong Kong Examinations and Assessment Authority.[1]

[1]https://www.hkeaa.edu.hk/en/recognition/benchmarking/hkdse/ielts/

We also recruited 25 speakers of British English for comparison (three males, 22 females). They were all undergraduate university students, aged between 19 and 34 years (mean age 21.16 years; St. Dev. ±2.76) years, and participated as the research-led teaching element of a module on Global Englishes. Three of them had experience of living abroad for no more than one year (three months to twelve months). None of them reported speech or hearing difficulties.

2.2 Materials

A short story was created and used in the experiment (see Appendix 1). The story is 523 words long and contains different types of sentences (e.g. questions; statements). 28 sentences and part-sentences in the passage were selected as test items. The sentence types used as test items and the numbers of each sentence type and their nuclear tones are shown in Table 1.

Table 1. Perception study: numbers of test items.

Sentence type	**Nuclear tone**	**Number of items**
Statement	fall	3
Continuation	level/rise	3
Statement question	rise	3
Echo question	rise/fall-rise	3
Yes/no question	rise	3
WH-question	fall	3
Closed tag	fall	3
Open tag	rise	2
Checking tag	rise	1
Sarcasm	rise-fall/fall	3
Checking	fall-rise/rise	1

Participants were recorded reading the story **before** (here, “Study 2”) they did the intonation selection task.

Five possible English nuclear tones were examined in this experiment: fall, rise, fall-rise, rise-fall and level. Recordings of the words with each of the five tonal patterns used in the experiment were produced by a female speaker of BrE (the second author), who had previously read the entire passage in a narrative style to ascertain which were the likely tones on the test items. It was decided not to synthesise the different tonal patterns as we wanted the stimuli to sound natural. The nucleus was on the same syllable in each tonal pattern.

The assigned 'correct' nuclear tones for each sentence type are based on the second author's performance in the initial reading, and standard descriptions of BrE intonation (Cruttenden 2014; O'Connor & Arnold 1973; Wells 2006).

2.3 Procedure

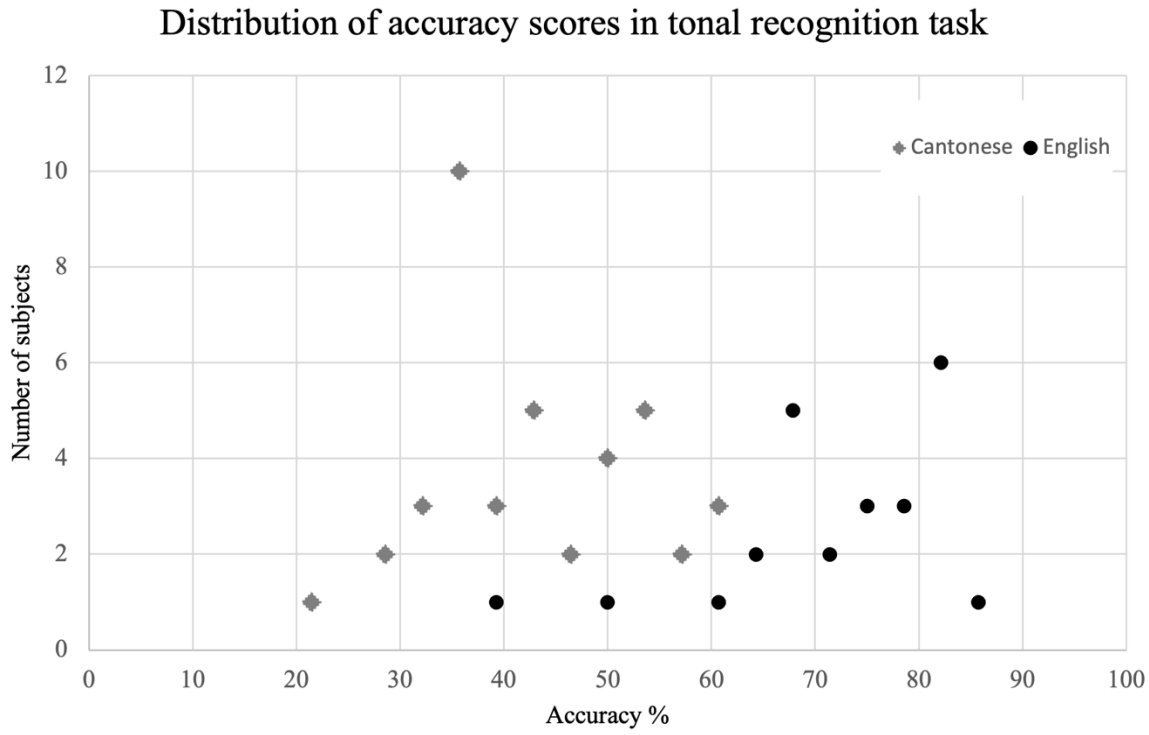
The experiment was conducted using PowerPoint. There were 44 slides in the presentation. 42 slides contained the experimental passage, 28 of these with test items. Each non-experimental slide shows one or more sentences from the story together with a recording which started automatically. Subjects were notified if the next sentence was a test item, which was shown in a separate slide. The five possible answers were played, each containing one of the five nuclear tones which were labelled A, B, C, D, and E. The sequence of five nuclear tones was presented in a different order for each test sentence. Subjects were asked to choose the most appropriate answer and write down the corresponding letter on an answer sheet. A printed version of the story was provided. Subjects could re-play any part of the passage, including the possible answers, during the experiment.

3. Results for tonal recognition study

The overall accuracy of each subject was calculated. Accuracy was based on the selection of test sentences containing the assigned nuclear tone for each sentence type in Table 1[2]. It should be noted that we are using the term 'accuracy' here as shorthand to indicate whether the participants selected the expected tonal pattern. In some cases, more than one tonal pattern was deemed appropriate based on the existing descriptions of English intonation mentioned above.

As expected, the BrE group had a higher averaged overall accuracy (72.57%) than the HKE group (42.94%), supporting hypotheses 1a and 1b for 'perception'; i.e. the BrE speakers' results were closer to the published descriptions of English intonation than the HKE speakers.

Figure 1 shows the detailed results. The highest accuracy for the BrE participants was 85.71% (only one participant). Most BrE participants scored in the 70%-80% range and only three had scores of 60% or below.



[2] In the legend for Figures 1 and 2, "Cantonese" refers to HKE speakers and "English" to BrE speakers.

Figure 1. Overall results for tonal recognition study: HKE versus BrE.

In contrast, only three HKE speakers reached 60%. Most were between 30% and 60%, and three of them were below 30% accuracy.

Figure 2 shows the averaged accuracy for each sentence type in the HKE and BrE groups. The BrE group had a higher averaged accuracy than the HKE group for all types of test items, except for 'Continuation'. Compared with the BrE speakers, HKE speakers showed particular difficulty in selecting suitable intonation patterns for WH-questions and some tag questions.

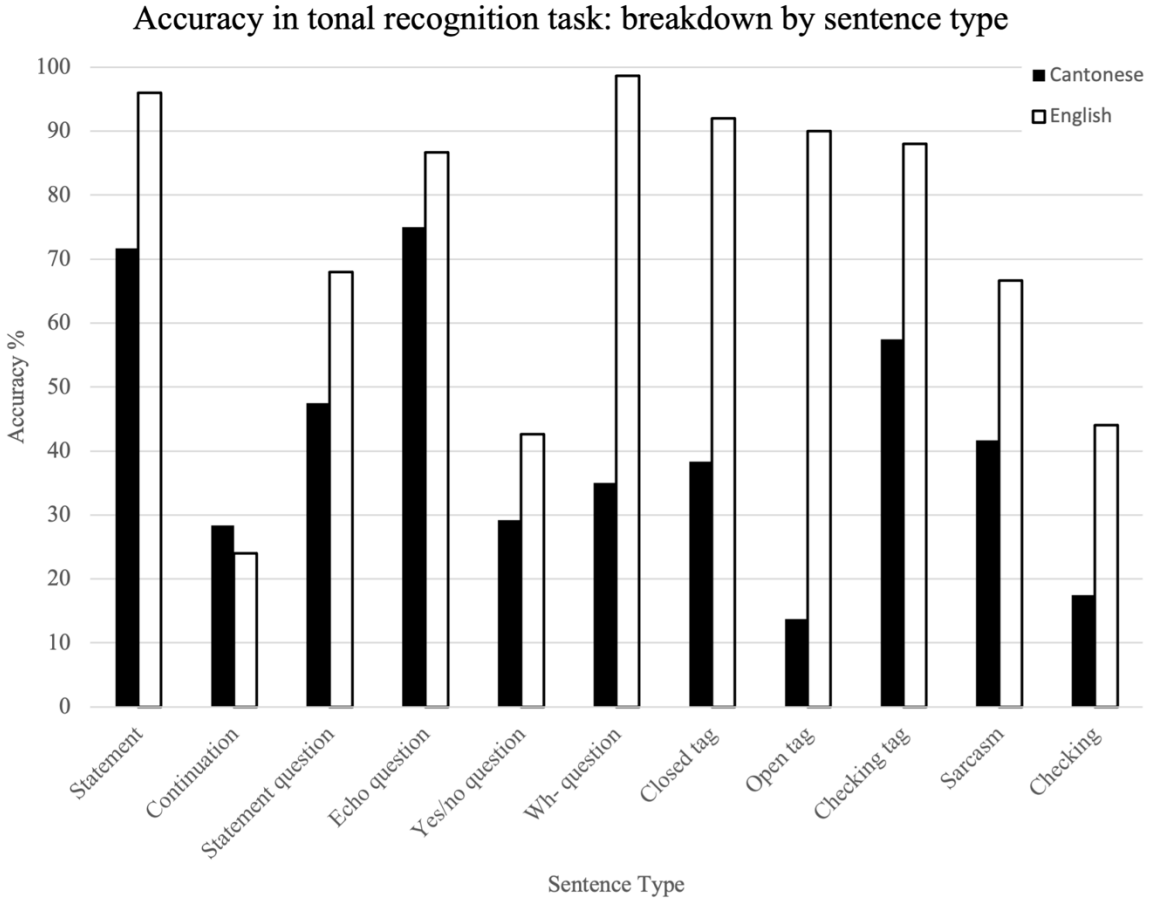


Figure 2. Results for perception study by sentence type: HKE versus BrE.

Table 2 shows the detailed choices for each sentence type; in general, the responses of the HKE speakers are more widely distributed. For 'Statement', the expected intonation pattern is 'fall'. Both BrE (96%) and HKE (72%) speakers tended to select 'fall' as the most appropriate answer, but there is more variation among the HKE speakers. For 'Continuation', their responses are much more mixed. The anticipated tone is 'rise' or 'level'; however, 'fall' is the most commonly selected tone for this sentence type (HKE 49%; BrE 57%), with the 'rise' (HKE 14%; BrE 20%) and 'level' tones (HKE 14%; BrE 5%) selected much less often by both groups.

Table 2. Distribution of choices (rounded percentages) for each sentence type. Target forms are shaded.

Type	Group	Fall	Rise	Fall-rise	Rise-fall	Level
Statement	HKE	72	0	13	12	3
	BrE	96	0	1	1	1
Continuation	HKE	49	14	18	4	14
	BrE	57	20	17	0	5
Statement question	HKE	16	35	35	5	9
	BrE	6	66	26	0	2
Statement question/echo	HKE	18	70	3	8	0
	BrE	24	52	20	4	0
Echo question	HKE	12	13	62	11	3
	BrE	9	45	41	1	1
Yes/no question	HKE	46	29	18	2	6
	BrE	36	43	20	1	0
WH-question	HKE	35	14	18	24	9
	BrE	99	0	0	1	0
Closed tag	HKE	38	33	18	3	7
	BrE	92	1	1	4	1
Open tag	HKE	43	14	11	4	29
	BrE	6	90	0	0	4
Checking tag	HKE	30	58	8	3	3
	BrE	4	88	0	4	4
Sarcasm (1)	HKE	31	16	19	20	13
	BrE	68	0	2	28	2
Sarcasm (2)	HKE	13	10	50	23	5
	BrE	4	0	84	8	0
Checking	HKE	63	20	18	0	0

Type	Group	Fall	Rise	Fall-rise	Rise-fall	Level
	BrE	56	0	44	0	0

For ‘Statement question’, ‘rise’ is the expected choice, while both ‘rise’ and ‘fall-rise’ are acceptable for ‘Echo question’. One of our test items could be construed as either a ‘Statement question’ or an ‘Echo question’; we accept both ‘rise’ and ‘fall-rise’ for this item. The distribution of the five options for these sentence types can be seen in Table 2. ‘Rise’ and ‘fall-rise’ the main choices for both ‘Statement question’ and ‘Echo question’. HKE speakers did not show a preference for one or other of these tones for ‘Statement question’, but preferred ‘fall-rise’ for ‘Echo question’. The pattern is contrary to the BrE speakers, who showed a preference for the ‘rise’ in ‘Statement question’ and a fairly even split between ‘rise’ and ‘rall-rise’ for ‘Echo question’.

4. Method for production study

The participants for the production study (carried out **before** the recognition task described above) formed a subset of those who volunteered for the perception study above, 10 from each cohort; all speakers but one were female in each cohort. The same story text (Appendix 1) was used: in this case, it was presented line by line on a screen to the BrE speakers, on a printed page to the HKE speakers; each participant was at liberty to repeat each sentence as many times as desired. (S)he was given time to read through the whole passage before the recording began, and was asked to make sure that (s)he had fully understood the meaning of the passage. The participant was asked to read out the passage as clearly and naturally as possible. Some cues such as punctuation (“?!”) and contextualising words in the narrative such as qualifying adverbs (“sarcastically”) were intended to promote particular intonational patterns.

These readings were captured on to solid state recorders: a Zoom H2 unit (HKE) with internal microphone, and a Roland Edirol R-09 with Rode Lavalier lapel microphone (BrE), with a 44.1 kHz sampling rate (16 bit).

Participants' intonation was analysed first by ear by the first author (a speaker of British English). A tabulation was made for the nucleus position and contour type for each utterance. We consider this the appropriate procedure for such an experiment since pitch movements can be very subtle and may be missed by even the best pitch-tracking algorithms, which are prone to octave jumps and no readings when there is creak voice. However, recourse was made to f0 traces in *Praat* (Boersma and Weenink 2020) for objectivity when it was hard to decide on an auditory classification. A high level of agreement was reached with the second author's independent analysis after joint reconsideration of the more problematic instances, and once allowance had been made for two common cue-trading phenomena, namely that fall-to-mid stands for a fall-rise, and that creaky voice implies a rise (where perhaps a change in vocal register is readily taken to correspond to a pitch excursion). Thus, an instance of a fall into creaky voice stands for fall-rise. As in the perceptual study, so too here a set of text-book templates, based on O'Connor and Arnold (1973) and Wells (2006), and the intuitions of the second author, were taken as canonical, against which the findings could be assessed. The results which follow come from a scoring of participants' productions in comparison with these templates.

5. Results for the production study

5.1 Choice of tune

The overall picture is one of disparity between the two groups (see Figure 3). The BrE speakers, with a 'success' rate varying along a range from 64% to 86% (mean 71%), clearly outperform the HKE speakers whose scores range from 43% to 71% (mean 55%). There is evidently some overlap between the upper quartile of the HKE cohort and most of the BrE group, which is to say that performance was greatly varied in the former, and the 'highest scores' demonstrate considerable BrE-like patterning in their phonology according to our criteria.

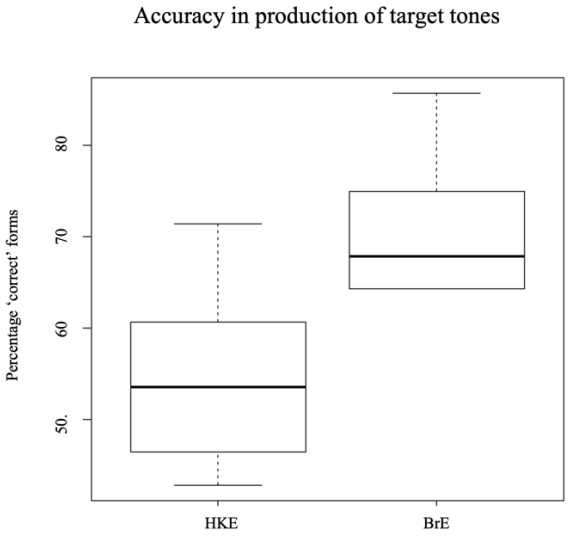


Figure 3. Production of target forms: overall scores for BrE and HKE.

However, the fact that the most competent of the BrE cohort hit a ceiling well below 100% 'success' must alert us to the weakness of such paradigms for assessing competence. Reasons for the limitations of the experimental design are discussed in Section 7. The other unambiguous observation is that the BrE results are heavily skewed towards the lower end of the scale. This means that there is a tight clustering where most participants perform, and a tapering into the higher scores. One interpretation of this is to say that a subset of the speakers is more conservative in their intonation, and that the majority are innovating in certain ways (i.e. showing a similar degree of mismatch with the templates).

Table 3 sets out the choice of tunes for all speakers together, expressed as percentages for ease of comparison between the two cohorts. The utterance types are self-explanatory except for the two variants on sarcasm. 'Sarcasm 1' is a question – "Would you, now?"; 'Sarcasm 2' is a statement. The near-total production of a fall for statements is not remarkable although we might have expected to find high rising tones which are prevalent in spontaneous speech (Warren 2016).

Table 3. Nuclear tones for all speakers (%). Shaded cells indicate target tones.

		fall	rise	fall-rise	rise-fall	level	other
statement	HKE	97			3		
	BRE	93		3		3	
continuation	HKE	47	10	3		40	
	BRE	40	3	27		30	
statement question	HKE	10	70	15		5	
	BRE	5	15	80			
statement question / echo	HKE	20	30	50			
	BRE	10	10	80			
echo	HKE	13	47	27		13	
	BRE	3	37	57		3	
yes/no question	HKE	20	50	23		7	
	BRE	13	23	64			
WH-question	HKE	77	13	3		7	
	BRE	93		3	3		
closed tag	HKE	37	40	10		7	7
	BRE	90	3	3		3	
open tag	HKE	65	10	15	5	5	
	BRE	5	95				
checking tag	HKE	30	70				
	BRE		100				
sarcasm 1	HKE	10	35	15	40		
	BRE	60		5	35		
sarcasm 2	HKE	90			10		
	BRE	60		20	20		
checking	HKE	80	10	10			
	BRE	10		90			

Our discussion of these figures now centres on those conditions where there is a mismatch between the target(s) (shaded cells) and our findings.

a) Mismatches in both groups:

i. Continuation. We see no commitment to a level tone in this condition. HKE and BrE are as likely to exhibit a fall, or a fall-rise in BrE. There is no clear tone of preference in this condition as far as our method can elicit. We cannot therefore state that HKE are missing a target if no target is apparent in BrE.

ii. Sarcasm 2. Both groups seem to lack a rise-fall target for a sarcastic declarative. This is a tentative interpretation since participants could easily have given a straight reading without a sarcastic interpretation[3].

b) Non-target forms in BrE but not in HKE:

i) Statement question. Adoption of the fall-rise appears to be a genuine innovation from the 20th century RP template. HKE shows a dispreference for turning contours here as elsewhere.

ii) Yes/no question. The same pattern as for statement question.

c) Non-target forms in HKE but not BrE

i) Closed tag. The exceptional 'other' tone is a rise-fall-rise. This complex contour may simply be a case of a rise-fall for emphasis followed by a boundary rise for an interrogative – or a confused reading (it would be exceptional among Englishes to have an intonational *form* which is foreign to British English).

ii) Open tag. Our results confirm expectations in showing no evidence for distinct open or closed tag intonations.

iii) Checking. HKE appears to treat this utterance type as any other statement rather than a category in its own right.

The rise-fall has limited, pragmatic scope in British English, chiefly indicating sarcasm. It is unlikely that its use by HKE speakers here in statement and open tag conditions is motivated by sarcasm. Rather, we suggest that the rise-fall pattern is a means for emphasis. This tallies with previous findings for HKE (Setter, Wong and Chan 2010) and for Singapore English (Deterding 1994). In summary: while the same tunes are in evidence in both varieties, the distribution differs due to certain missing categories in HKE, and different preferences within some categories. For a nativizing variety of English (the third of five phases in Schneider's Dynamic Model; see Schneider 2007), we would expect to see "borrowing from the [indigenous speech community] strand, coining new words using

[3] Our reviewer has rightly observed that there is a mismatch between 'sarcasm 1' and 'sarcasm 2' in the experimental design: the text of the former explicitly states 'Luke said sarcastically'. It is possible that this prompted a higher occurrence of rise-fall by HKE.

strategies of word formation, and adjusting the meaning of existing words to novel environmental conditions" (Schneider 2007: 79). Schneider does not discuss intonation *per se* but his statement here about lexis finds an analogue in our prosodic findings.

5.2 Nucleus position

One potential problem with analysing a second language using a framework designed to describe a particular accent of the target variety is that variation may occur outside the paradigm, such as the presence of more than one tone in an intonation group – indeed the paradigm may not fit at all, if for instance the language variety simply does not anchor a tone to a single nuclear syllable in every utterance (see Deterding 1994: 62). For the sake of simplicity and to facilitate comparison between our two varieties of English we have assumed that the HKE productions have a single nucleus and that, as per the British tradition (O'Connor and Arnold 1973), this is necessarily the final stressed syllable of the utterance. An alternative method would be to pass over all utterances with 'wrong' nucleus placement since they almost certainly have the 'wrong tone', e.g. for "At least, I hope it will" we expect a 'checking' fall-rise on *hope*, but most speakers gave a fall on *will*. While we find speakers marking the 'wrong' word here and there as nucleus, in 6 of the 28 sentences the misplacement is universal or nearly so, and the BrE do the same for two of these. 9 of the HKE speakers read out two intonation phrases for "would you, now?" as though the standard orthography had been misleading.

6. **How do the studies relate to one another (hypotheses 2a and 2b)?**

Correlation by speaker

Figures 4 and 5 show scatterplots for production against perception for each group separately. It is immediately apparent that there is no discernible trend in the HKE data ($r(8)=-.24$, $p=0.5$) and, moreover, while there are some data points which forcibly confirm a correlation, there are others which deny it. For BrE there is a weak positive correlation but again it is not significant ($r(8)=.29$, $p=0.4$).

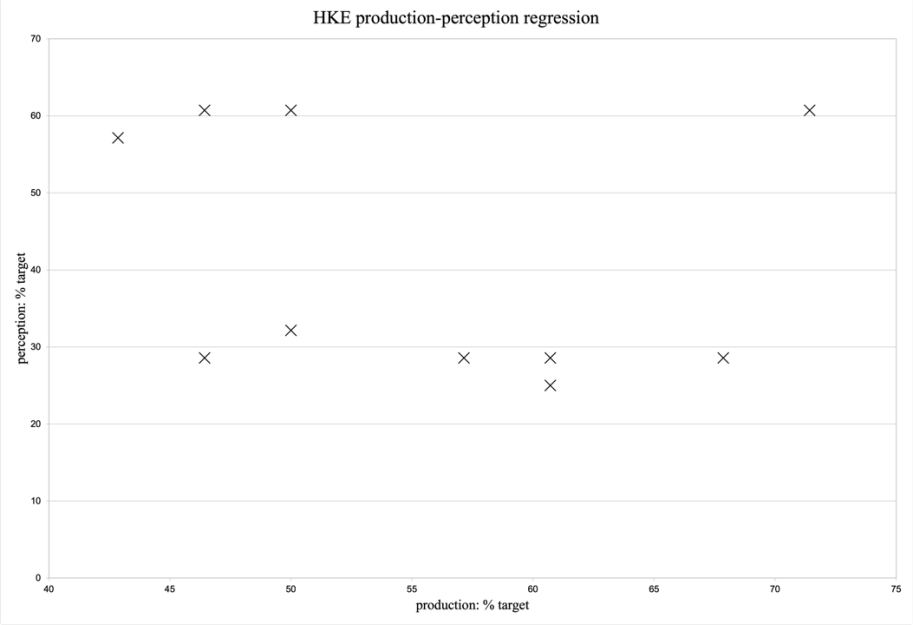


Figure 4. Scatterplot for HKE production scores. Each point represents an individual who participated in both studies.

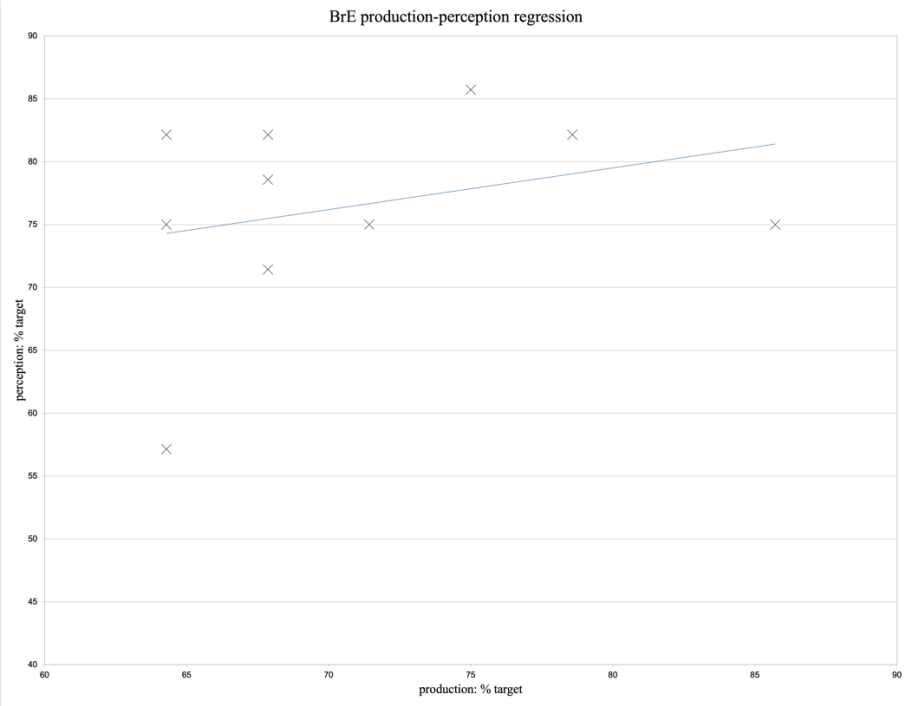


Figure 5. Scatterplot for BrE production scores. Each point represents an individual who participated in both studies.

Correlation by sentence type

A rather strong correlation ($r(8)=.60$, $p<0.001$) is seen between HKE production and perception when the data are separated by sentence type (see Figure 6). Here, each data point represents the average score for a sentence type from the tonal recognition experiment plotted against the average score for the same sentence type from the production study (note that the number of data points is lower than the number of sentences due to overlapping production-perception scores). In other words, smoothing over differences between the varied performance of individuals, we are somewhat able to predict performance in tonal recognition if we know more detailed production performance. Open tags, sarcasm and continuation sit at the low end of the linear trend, statements and echoes at the other. It is interesting to note that a not dissimilar correlation exists ($r(8)=.68$, $p<0.001$) for BrE (see Figure 7), which we take to mean that the way in which these speakers deviate from our presumed targets has a symmetry in the two parts of the experiment. Moreover, for the BrE cohort we again find statement and echo at the high end of the slope, and continuation and sarcasm at the low end. Notably, however, open tag shows a high correlation between production and perception for BrE. Thus the results are showing a complex of some expectations confirmed and thwarted in the same way by both groups but also some strong mismatches as with open tag.

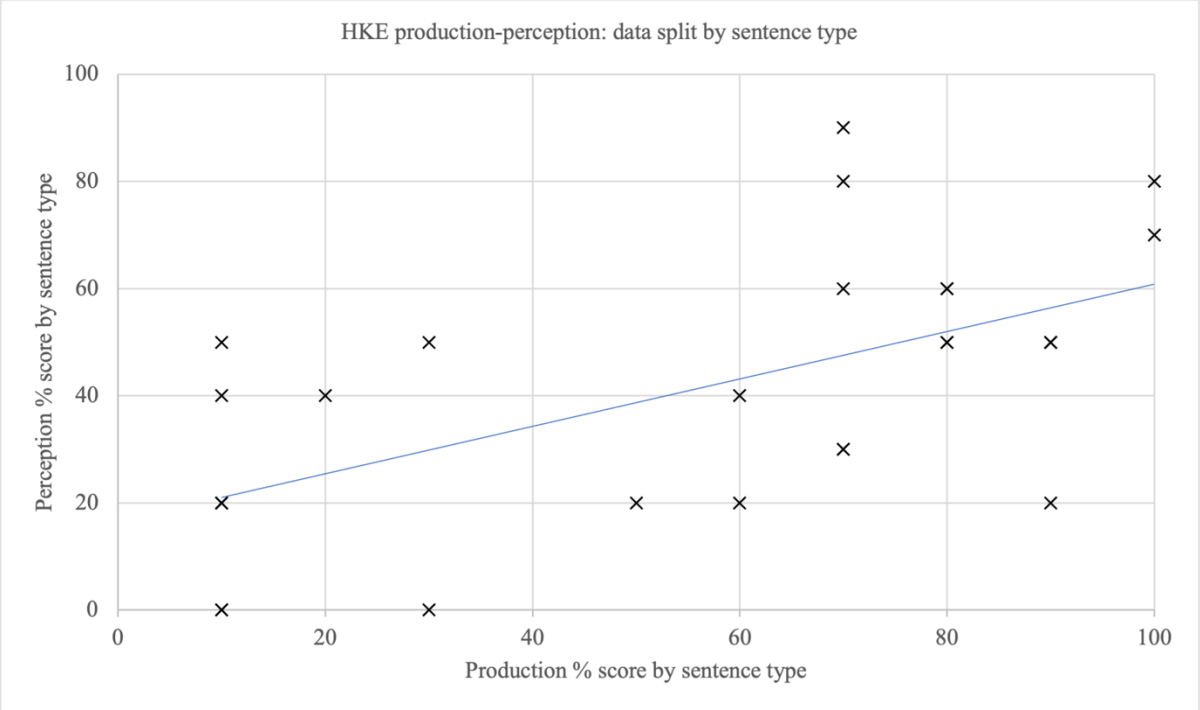


Figure 6. HKE production-perception by sentence type (scores for all speakers averaged). Each data point represents an utterance type.

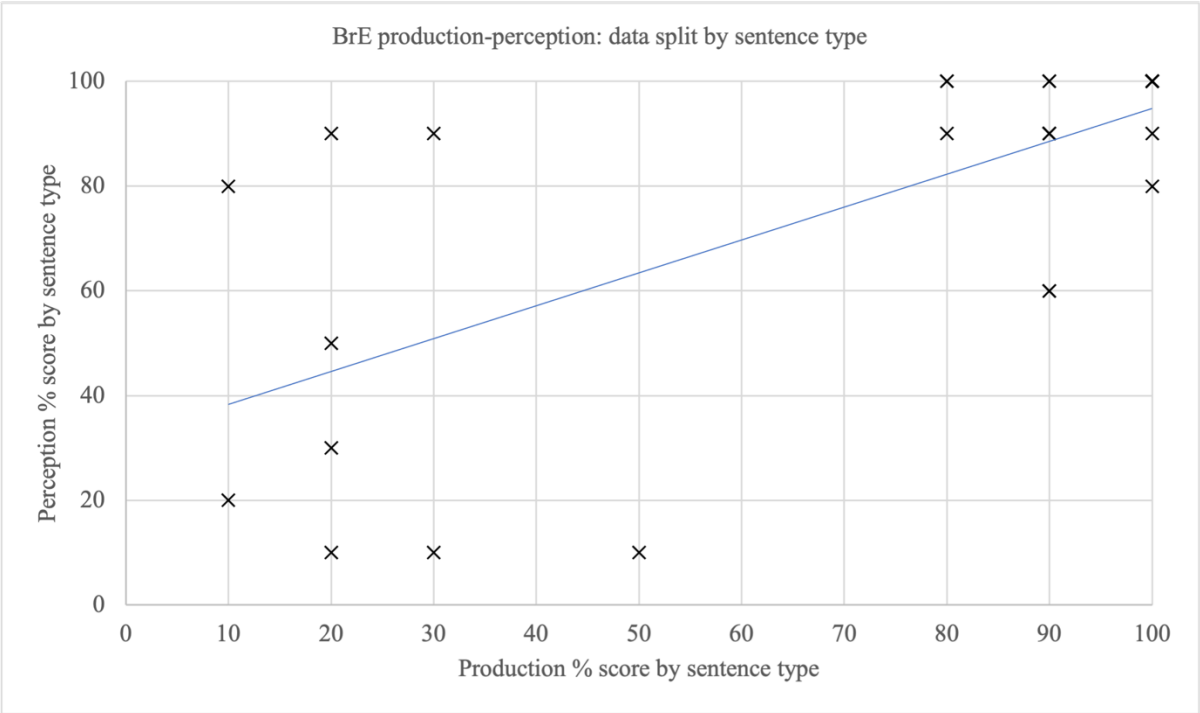


Figure 7. BrE production-perception by sentence type (scores for all speakers averaged). Each data point represents an utterance type.

Global scores

Table 4 summarises the scores for all participants across all sentence types for direct comparison with Puga et al. (2018). For the HKE learners, of the 280 data points (28 sentences * 10 speakers) for each study, (1) is a count of all instances where the individual has given a correct/target response in the production study *and* the perception study, (2) where (s)he has done so in neither, (3) and (4) where (s)he has done so in only one of the studies:

Table 4. Relationship between performance in production and perception tasks, learner data only, alongside data from Puga et al. (2018) for comparison

Correct	N	%	N (Puga et al. 2018)	% (Puga et al. 2018)
(1) Both	82	29.3	224	40.0
(2) None	92	32.9	105	18.8
(3) Perception only	33	11.8	79	14.1
(4) Production only	73	26.1	152	27.1

This summary of HKE learner performance has in common with the German learner data that scenario (4) is approximately double the score for (3). It will be interesting to see how commonly this pattern emerges in future studies involving learner performance data. We have already seen from a lack of clear pattern in the regression for HKE that not all learners conform to this trend, and the strong trend which emerges for sentence type tells us that this overall picture is not to be found for every intonation pattern studied.

7. Conclusion

It is not difficult to see the limitations of pattern-spotting where elicitation of utterances has involved no more direction than is inherent to the text such as punctuation and adverbs. The primary concern is that some readers will at times have failed to comprehend the pragmatics of an utterance especially if they did not 'internalize' the text before delivering it. However, without a design of this type it would be difficult to elicit some of the tones, which only occur in very specific contexts, in a controlled way. We do not claim to have uncovered the same patterns as would be found in spontaneous speech.

Our first study is somewhat more complex than a simple perceptual experiment in that participants are expected to evaluate contextually appropriate stimuli. While we defend this approach for obtaining data useful for our purposes, it is expected that a certain amount of noise in the data comes from the unnaturalness of the situation. The fact that the BrE English speakers hit a ceiling well below 100% in the first study may be as informative of the limits of the methodology as it is of their phonology.

In summary, in comparing our two studies we have demonstrated that the Hong Kong learners attained the targets in production more than twice as much as in perception. This corroborates the finding for German learners of English in Puga et al. (2018) and opposes the notion that perceptual capability necessarily precedes competence in production in L2 speech learning. While this may seem counter-intuitive, the fact that different mechanisms underlie perception and production may well explain what we have found. Production is imitative, and involves muscle memory, whereas choosing appropriate intonation contours by context involves categorical decision-making. We also recognise that most of the utterances called for a rise or a fall, so a speaker who would not produce any contour tones would be likely to hit 50% accuracy in the production experiment; chance level in the perception experiment is 20%. Finally, we ought not to be surprised if prosodic studies do not align with 'segmental' studies which show production to be contingent on perception. For instance, a learner must struggle with a mismatch in vowel space between L1 and L2, but a rise or a fall in pitch is seemingly a more binary affair.

The connection we see between perception and production of intonation is bound up with utterance type, whether linguistic (e.g. statement) or paralinguistic (e.g. sarcasm). Where there is almost no production of the target form, and especially where this corresponds to no recognition of the target in perception, we recognise that there may not simply be a pedagogical challenge, but a different target from that which is documented for British English. Targets may be shaped by, e.g. American or Australian media, and the fact that the British cohort did not conform to the textbook targets is most probably indicative of changes in their phonology. It is interesting to note the lack of high-rising terminals in the British data for a reading context. Likewise, we have confirmed a seeming lack of tone categories for open and closed tags for the Hong Kong speakers, and possibly a reassignment of the rise-fall to emphasis. Although our sample size is too small to warrant strong pronouncements, the findings are in line with expectations for an emergent variety in a stage of 'nativization' (Schneider 2007).

Acknowledgement

The authors would like to acknowledge the assistance of Dr Noor Mat Nayan in the collection of the UK data.

8. References

Boersma, Paul, and Weenink, David. 2020. Praat: Doing Phonetics by Computer [computer program].

Bolinger, Dwight. 1998. "American English". In Daniel Hirst, and Albert Di Cristo, eds. *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press, 45-55.

Bolton, Kingsley, and Helen Kwok. 1990. "The dynamics of the Hong Kong accent: Social identity and sociolinguistic description". *Journal of Asian Pacific Communication* 1: 147-172.

Bradford, Barbara. 1988. *Intonation in Context*. Cambridge: Cambridge University Press.

Brazil, David, Malcolm Coulthard, and Catherine Johns. 1980. *Discourse intonation and language teaching*. London: Longman.

Brazil, David. 1994. *Pronunciation for Advanced Learners of English* (Teacher's Book and Student's Book). Cambridge: Cambridge University Press.

Cauldwell, Richard. 2002. "Phonology for listening: relishing the messy". In *Proceedings of TESOL, April 2002, Salt Lake City, Utah*.

Cheng, Winnie, Chris Greaves, and Martin Warren. 2008. *A Corpus-driven Study of Discourse Intonation: The Hong Kong Corpus of Spoken English (Prosodic)*. Amsterdam: John Benjamins.

Chun, Dorothy. 1998. Signal Analysis Software For Teaching Discourse Intonation. *Language Learning & Technology* 2: 61-77.

Chun, Dorothy. 2002. *Discourse Intonation in L2: From theory and research to practice*. Amsterdam: John Benjamins.

Cruttenden, Alan. 2014. *Gimson's Pronunciation of English* (8th ed.). Routledge.

Deterding, David. 1994. The Intonation of Singapore English. *Journal of the International Phonetic Association*, 24, 61-72.

Flege, James. 1995. Second-language speech learning: findings, and problems. In Strange, Winifred (ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. York, Timonium, Maryland.

Grabe, Esther, Burton S. Rosner, José Eugenio García-Albea, and Xiaolin Zhou. 2003. Perception of English Intonation by English, Spanish, and Chinese Listeners. *Language and Speech*, 46(4), 375–401.

He, Xuliang, Vincent van Heuven, and Carlos Gussenhoven. 2012. The selection of intonation contours by Chinese L2 speakers of Dutch: Orthographic closure vs. prosodic knowledge. Second Language Research, 28, 283–318.

Hudson, Toby, Jane Setter, and Peggy Mok. 2019. Nuclear tones in Hong Kong and British English. Calhoun, Sasha, Paula Escudero, Marija Tabain, and Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019. Australasian Speech Science and Technology Association Inc., Canberra, Australia, pp. 320-323.

Isbell, Daniel. 2016. The perception-production link in L2 phonology. *MSU Working Papers in Second Language Studies* 7 (1).

Jenkins, Jennifer. 2000. *The phonology of English as an international language*. Oxford: Oxford University Press.

Lin, Angel M. K. 2008. "Code-switching in the classroom: research paradigms and approaches", in King, Kendall, and Nancy Hornberger (eds), *Encyclopedia of language and education, research methods in language and education* (Vol .10, 2nd ed. pp. 273-286). New York: Springer.

Matthews, Stephen and Virginia Yip. 2013. *Cantonese: A Comprehensive Grammar*. London: Routledge.

Mok, Peggy, Yanjun Yin, Jane Setter, and Noor Mat Nayan. 2016. Assessing knowledge of English intonation patterns by L2 speakers. *Proceedings of Speech Prosody* 2016, 543-547. Boston.

Mompeán Gonzáles, José Antonio, and Rafael Monroy Casas. 2010. Nuclear Tones in English: The Case of Household Consumer Products TV Commercials. *ReVEL*.

Munro, Murray, and Tracey Derwing. 2015. A prospectus for pronunciation research in the 21st century: A point of view. *Journal of Second Language Pronunciation*. 1. 11-42.

O'Connor, Joseph D., and Gordon Arnold. 1973. Intonation of Colloquial English. 2nd edition. London: Longman.

Puga, Karin, Robert Fuchs, Toby Hudson, Jane Setter, and Peggy Mok. 2018. The perception-production link in intonation: evidence from German learners of English. *Speech Prosody* 2018, 18 - 20 June 2018, Poznan, Poland, pp. 685-689.

Roach, Peter. 2009. *English Phonetics and Phonology. A practical course* (4th ed.) Cambridge: Cambridge University Press.

Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the World* (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press.

Setter, Jane, Cathy Wong, and Brian Chan. 2010. *Hong Kong English*. Edinburgh University Press.

Taylor, David S. 1993. Intonation and accent in English: what teachers need to know. *International Review of Applied Linguistics in Language Teaching* 16 (1), 1-21.

Tench, Paul. 2015. *The Intonation Systems of English* (second ed.). London: Bloomsbury.

Warren, Paul. 2016. *Uptalk: The Phenomenon of Rising Intonation*. Cambridge: Cambridge University Press.

Wells, John C. 2006. *English intonation: An introduction*. Cambridge, UK: Cambridge University Press.

Wermke, Kathleen, Johanna Teiser, Eunice Yovsi, Paul Joscha Kohlenberg, Peter Wermke, Michael Robb, Heidi Keller, and Bettina Lamm. 2016. Fundamental frequency variation within neonatal crying: does ambient language matter? *Speech, Language and Hearing*, 19(4), 211-217

Appendix 1: text of the story

Luke had always had a wild imagination. Whenever he read a book, he'd imagine himself as the main character. 'One day I am gonna be a hero,' he often thought to himself, '**#1) just like the people in the *stories.**' (statement: fall 01)

One stormy night, Luke was awoken by the sound of the wind blowing through the open window. Immediately, he recalled a scary movie he'd watched that day with that same situation. Nervously he thought to himself: '**#2) It was just the wind … *wasn't it?**' (open tag: rise)

Just when he thought he'd convinced himself it was nothing, he heard a noise outside his room. Something was gently, rhythmically scratching the floor. Luke was terrified, yet very curious. '**#3) What *is that sound?**' (WH-question: fall) Full of curiosity but still rather scared, he slipped out of bed, reaching for the bedroom door.

The door banged open suddenly. It was his big brother, Jerry! '**#4) Can you *hear that?**' Jerry asked (yes/no question: rise). 'I'm just about to go check it out,' Luke said. 'Then, rather unsure, '**#5) It could just be in our heads … *couldn't it?**' (open tag: rise)

'**#6) Just in our *heads?!**' (echo: rise/fr) Jerry rolled his eyes. 'It's really happening!'

'**#7) Well, in that case, we'll have to go and take a look, *won't we**?' said Luke, taking charge (closed tag: fall).

Jerry swallowed hard. '**#8) Take a *look**,' he repeated (echo: rise/fr), his eyes wider than ever. '**#9)You're actually suggesting we *do that?**' (statement question: rise)

'Oh don't be such a baby, Jerry,' said Luke. '**#10) We'll go and have a *look** … and it'll be fine.' (continuation: level) '**#11) At least, I *hope it will**' he added, under his breath (checking: fall-rise).

They slowly tiptoed towards the dark, empty living room. There, at the other end of the room, right beside the creaky window, was a tiny little green light, floating and shimmering on the floor.

Jerry pushed Luke forward muttering, 'It's probably just your laptop. **#12) Not so scary, *is it?**' (closed tag: fall)

'**#13) So you're making me go in *alone?**' (statement question: rise) Luke was very surprised.

Jerry tried to play things down with humour. '**#14) If anything *happened,**' (continuation: level) he whispered to Luke, trying to appear calm, '**#15) I would write a book about how you'd saved my *life.**' (statement: fall 02)

'**#16) Would you, now?**' Luke said sarcastically. (sarcasm: rise-fall/fall) 'How nice. But you know what? I am going.'

'**#17) Are you *serious?** I was kidding!' (yes/no question: rise) Jerry tried to stop Luke, but Luke had already left for the source of the scratching sound.

Luke carefully approached the green light. He could see a dark shadow, waving its claw in the air. Luke forced himself to look at it…

'**#18) What *is that thing?**' (wh- question: fall) Jerry asked in a shaking voice. '**#19) Is it some kind of a ...*monster?**' (yes/no question: rise). **#20) 'You don't really think it's a monster … *do you?**' (checking tag: rise)

'**#21) A *monster?**' said Luke. (echo: fall-rise/rise) No. **#22) It's just the walking *robot,**' (statement: fall 03) he added, his voice full of disappointment. 'I think **#23) the wind has blown it off the *shelf** and it's landed on its switch.' (continuation: level)

'**#24) It's the *robot?**' (statement question/echo: f-r/rise) Jerry could not believe his ears. '**#25) Well that's just lame, *isn't it?**' (closed tag: fall)

#26) 'Of *course it is' said Luke, defensively and feeling a bit stupid (sarcasm: rise-fall/fall). '**#27) Why don't we go back to bed?**' (WH-question: fall)

'I agree,' said Jerry. 'But you were actually quite brave you know. **#28) Even *though there wasn't anything to be scared of**.' (sarcasm: rise-fall)

Appendix 2: the material used for familiarization (* marks the nucleus of a phrase).

There once was a young golden-retriever called Lucky. '**#1) Today is such a good *day,**' (statement: fall) he thought to himself. '**#2) I'd love to have a walk in the *woods.**' (statement: fall)

Authors

Mr Toby Hudson: Phonetics Laboratory, 41 Wellington Square, Oxford, UK, OX1 2JF (tobyhudson@cantab.net) ORCID: 0000-0003-4871-8372

Prof. Jane Setter: Department of English Language & Applied Linguistics, School of Literature & Languages, Room 210A, Edith Morley Building, University of Reading, Whiteknights, Reading, UK, RG6 6EL (j.e.setter@reading.ac.uk) ORCID: 0000-0001-7334-5702

Prof. Peggy Mok (corresponding author): Department of Linguistics and Modern Languages, Leung Kau Kui Building, The Chinese University of Hong Kong, Shatin, Hong Kong (peggymok@cuhk.edu.hk) ORCID: 0000-0002-9284-6083