

Foreign accent strength and intelligibility at the segmental level

Rubén Pérez-Ramón^{a,b}, María Luisa García Lecumberri^{a,*}, Martin Cooke^c

^a Language and Speech Laboratory, Universidad del País Vasco, 01006 Vitoria, Spain

^b Waseda University, Tokyo, Japan

^c Ikerbasque (Basque Science Foundation), Spain

ARTICLE INFO

Keywords:

Intelligibility
Foreign accent
Consonants
Interlanguage

ABSTRACT

The relationship between strength of foreign accent and intelligibility is not straightforward. This relationship resists a simple characterisation due in part to the multiplicity of cues that carry accent in the word- and sentence-level materials typically used in the study of accent. One of the principal conveyors of accent is the phonetic segment. The current study attempts to isolate this segmental contribution to foreign accent and consequently measure the relationship between segmental accent and intelligibility for listeners with differing linguistic correspondence to the target and accented language. English, Spanish and Czech listeners identified English words in which the initial consonant was either intact, or had been replaced by a Spanish-accented counterpart; in a second task, they rated the accent strength of the same tokens. All speech material was produced by an English-Spanish bilingual talker. Overall, Spanish listeners displayed a smaller loss of intelligibility due to the accented segment than native English listeners, while the Czech cohort experienced the largest intelligibility loss. However, the relationship between accent strength and intelligibility loss was not linear, varying with phoneme identity and its role in a listener's first language. These findings suggest that how accented and intelligible a sound is depends strongly on the interactions between the phonological systems of speakers and listeners.

1. Introduction

The speech of non-native talkers frequently presents phonetic deviations from native speakers' productions, giving rise to a foreign accent (FA). FA has been widely studied with regards to both its strength and its effects on communication, and in particular how it affects intelligibility. Contrary to what is frequently believed, several studies (e. g., Munro and Derwing, 1995a,b; Derwing and Munro, 1997; Thomson, 2018; Zahro, 2019; Munro and Derwing, 2020) have shown that degree of foreign accent (DFA) and intelligibility are not necessarily highly-correlated. For example, Derwing and Munro (1997) demonstrated that strongly-accented speech can be quite intelligible. Munro and Derwing (1995a) speculated as to possible reasons for the non-dependence of these two dimensions in their utterance-based study, suggesting that intelligibility could have been promoted by the utterance context while listeners may have focused on segmental and suprasegmental deviations when making DFA judgments. Similarly, Varonis and Gass (1982) pointed out that FA assessment might have been influenced by grammatical errors.

The complexity of interpreting the relationship between DFA and

intelligibility partly stems from the 'holistic' nature of most studies of FA, in which judgements are based on words or utterances where FA may be present throughout the token. The holistic approach makes it both difficult to associate accent with individual segments, and challenging to disassociate potentially confounding effects on accent judgements arising from suprasegmental factors such as intonation (Kang et al., 2010), duration (Tajima et al., 1997) or speech rate (Munro and Derwing, 2001), or from higher-order linguistic factors such as syntax and pragmatics. Therefore, the question remains whether and how individual segments may affect the relationship between perceived degree of nativeness and intelligibility. Thus, one of the aims of the current study is to determine if such a close relationship exists, despite the fact that these two dimensions have been found to be only partially correlated in the past.

The segment-level perspective on FA adopted in the current study permits us to address a second question concerning the effect of FA and intelligibility on listeners with different first languages (L1). It is generally accepted that foreign accents are less intelligible than native accents, a belief that has been widely discussed in the literature (Munro and Derwing, 1995a; Munro, 1998; van Wijngaarden et al., 2002; Bent

* Corresponding author.

E-mail addresses: rperez.ram@gmail.com (R. Pérez-Ramón), garcia.lecumberri@ehu.es (M.L. García Lecumberri).

<https://doi.org/10.1016/j.specom.2022.01.005>

Received 23 February 2021; Received in revised form 10 October 2021; Accepted 13 January 2022

Available online 20 January 2022

0167-6393/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

and Bradlow, 2003; Sereno et al., 2016; Thomson, 2018). Nonetheless, the relationship between the intelligibility of a foreign accent and a listener's L1 has been explored (Major et al., 2002; Stibbard and Lee, 2006). Bent and Bradlow (2003) found that an L1 match between non-native talkers and non-native listeners facilitates speech intelligibility: Chinese listeners identified more English words spoken by Chinese speakers than spoken by English speakers. This effect has also been found for other L1 pairs, such as Polish-accented Danish (Korpala and Sobkowiak, 2020). Bent and Bradlow (2003)'s results also suggest that non-native listeners suffer a smaller or equal intelligibility deficit relative to native speech when listening to non-native speech of speakers with L1s different from their own: Korean listeners identified as many English words spoken by Chinese speakers as spoken by English speakers. However, Munro et al. (2006) found only marginal evidence that a shared L1 between speaker and listener results in better understanding. In addition, Stibbard and Lee (2006) observed that groups with differing L1s did not experience an intelligibility benefit, and indeed low-proficiency speakers led to a drop in intelligibility for non-native listeners, an effect also found by Shin (2018).

Rather than exploring the effect of the listener-speaker interaction on intelligibility by considering whether listeners and speakers share an L1, Pinet et al. (2011)'s study analysed the overall similarity of the vowel systems amongst different groups of native and non-native listeners and speakers in order to account for the intelligibility of sentences in background noise. They conclude that a strong predictor of accent intelligibility is the acoustic-phonetic similarity between a speaker's and a listener's accent in the target language (as determined in their study by the distance between their individual vowel systems). A similar conclusion is found in Stringer and Iverson (2019) applied to word recognition accuracy.

An alternative approach is to consider the similarity of listeners' and speakers' phonological systems. Indeed, the phonological mismatch hypothesis introduced in Imai et al. (2005) proposes that phonological representations are critical, and that, as a consequence, intelligibility will be better when non-native listeners' phonological representations match those of the target speech.

In the current paper, our aim is to isolate segmental FA and examine its effects by considering the phonological systems of the listeners. By focusing on modifications at the segmental level while controlling for speaker differences, the goal is to be able to derive a reliable association between individual segments and foreign accent effects. This association can then be used to explore the role of similarities between listeners' L1 phonological systems on accent effects.

In summary, the current study addresses two questions. First, does the relationship between intelligibility and degree of foreign accent depend on the identity of the segment? DFA and intelligibility have been found to be only moderately-correlated in holistic studies, a finding that is explicable if their relationship differs across segments. The second question concerns whether the relationship between foreign accent and intelligibility at the segmental level depends on the phonological system of a listener's L1. Holistic studies reported varying results depending on the matching between the target language, the accented language and the L1 of the listener, suggesting that the native phonological system has an impact on foreign accent perception.

To answer these questions, we compared three listener groups with differing L1s by examining their responses to English words in which the initial consonant was either intact (and therefore produced with a native pronunciation) or replaced by a Spanish-accented segment. The three groups differed in the relationship of their L1 to the languages of the target speech and the foreign accent in the following way: (1) the native English cohort shares the L1 with that of the target speech; (2) the native Spanish cohort has an L1 that matches that of the accented segments; and (3) the native Czech cohort differs in L1 to both the language of the target words and the accented segments. Native English listeners provide an intelligibility benchmark against which to measure the intelligibility displayed by NN listeners. The Spanish listener group will reveal

whether a matched language benefit is manifest at a segmental level, or whether it is a higher-level effect. The Czech listener group permits an examination of the effects of accented segmental realisations that are either alien or overlap with those of their L1.

To avoid problems present in holistic approaches to FA, we use a methodology in which single segments are replaced by their most common Spanish-accented counterparts (e.g., the initial segment [h] is replaced with [x] in the word *house*) while leaving the rest of the token unaltered. The methodology, piloted in García Lecumberri et al. (2014), is based on speech material produced by a balanced bilingual talker, side-stepping any confounds due to speaker differences. That study compared three techniques for segment replacement – code-switching, where the speaker is asked to pronounce only one segment with a foreign accent; splicing, a technique in which a native segment is taken out of the signal and a foreign segment is spliced in its place; and bilingual speech synthesis, generated with HMM technology – and concluded that splicing is currently the most reliable and flexible approach, as it does not require the high level of control of in-word code-switching while maintaining a higher degree of naturalness than speech synthesis. The splicing technique was also found suitable as the base for the FA gradation technique developed in Pérez-Ramón et al. (2020). The current study makes use of the spliced dataset introduced in García Lecumberri et al. (2014).

2. Methods

Listeners from each of the three L1 backgrounds (English, Spanish, Czech) first identified words which either contained or did not contain a single foreign-accented consonant segment in word-initial position, and in a subsequent task the same participants rated these items for degree of foreign accent. The dataset for the native English baseline group and the native Spanish cohort comes from a study which compared methods for introducing foreign accent at the segmental level (García Lecumberri et al., 2014); this dataset has been augmented for the current study with responses from a group of participants with Czech as their L1.

2.1. Participants

All Spanish and Czech participants were in their second year of an English Studies degree at the University of the Basque Country (Vitoria-Gasteiz, Spain) and Charles University (Prague, Czech Republic) respectively. They all had passed the necessary exams to take part in English courses at a B2 or C1 level according to the European Framework of Reference for Languages (Council of Europe, 2001). The native English group were Erasmus students from Southern England who were studying at the University of the Basque Country.

Recruitment of the two non-native groups was based on the following criteria: (i) participants were not at an advanced level in any language other than the required level of English; (ii) no participants self-reported hearing impairment. All listeners were paid for their participation.

Cohort size and composition was as follows: English ($n = 9$, 3 males, mean age: 22.3, std. 2.6); Spanish ($n = 21$, 3 males, mean age: 20.4, std. 1.8); Czech ($n = 33$, 9 males, mean age: 21.2, std. 3.8). A smaller sample of listeners matched with the target language was deemed sufficient because a higher degree of homogeneity in phonetic competence can be assumed for them, as the task was performed in their L1.

2.2. Stimuli

The corpus for this experiment consists of 88 words (Table 1) for which native and foreign-accented versions were generated through the splicing technique introduced in García Lecumberri et al. (2014), which is summarised below. Zipf frequency (Van Heuven et al., 2014; Zipf, 2016) of the words ranged from 2.65 to 6.82 (mean = 4.65; std = 0.82) on a 0–8 scale. Note that the term 'foreign-accented' in this context

Table 1

Native English segments (N), corresponding foreign-accented Spanish segments (FA), and words into which accented segments replaced native segments in initial position. Underlined segments in the FA column correspond to the Spanish realisations that overlap with the Czech phonological inventory.

| N | FA | Words |
|----------------|----|-------------------------------------------------------------------------------|
| h | x | half, hammer, happen, happy, heaven, hen, holiday, house, hundred, hungry |
| ɹ | ɾ | race, rain, rainbow, reach, reason, rice, rich, rider, run, rush |
| t ^h | t̄ | tea, teach, teacher, tiger, time, tiny, tom, touch, towel, type |
| k ^h | k̄ | cabbage, cash, catch, collar, cottage, cousin, keeper, king, kiss, kitchen |
| ð | d | theirs, them, then, there, they, this |
| v | b | valley, van, vanish, veal, veil, vein, village, violin, voice, voter, voucher |
| w | gw | waiter, wall, wane, wash, washing, weather, while, wife, win, winter, woman |
| j | ɟ̞ | yawning, years, yen, yes, yesterday, yours, use, useful, useless, user |
| ɟ̞ | j | gender, ginger, gipsy, jam, jeans, jelly, journey, jump, jane, jim |

refers to those tokens in which a waveform segment corresponding to an accented phoneme has been inserted, although in reality the foreign accent is confined to the initial phoneme.

In order to generate foreign-accented words via the splicing technique, a highly-proficient Southern British English/Spanish bilingual speaker was recruited. This speaker was judged to be accent-free in both languages by native speakers of Spanish and English (the second and third authors, respectively). As the goal of this rating was not to assess degree of accentedness but to decide if the speaker could be considered accent-free in both languages, only one rater per variety was deemed necessary. The speaker produced two sets of items. The first set consists of the English words in Table 1 pronounced with a native accent. The second set contains non-words produced with Spanish pronunciation. The purpose of the second set of tokens was to provide accented segments to replace native target segments. It is well known that in foreign language production one of the preferred speaker strategies when faced with a difficult target language sound is to use the most similar one in their L1 (Major, 2001; Safa, 2018), which is why in this study we have chosen the L1 sounds typically used by Spanish learners of English for the target language segments analysed. To achieve this goal, segments in the second list were produced in an identical acoustic context as the target segment in order to minimise artifacts arising from coarticulatory differences after splicing. For example, in order to process the English word *house* [haus], the Spanish non-word *jasa* [xasa] was recorded, so both the target native segment [h] and the inserted segment [x] were in a similar acoustic context. Once both lists were recorded, the target segment of the English word was replaced with the segment of the Spanish non-word via splicing. In order to avoid any bias related to the splicing technique as opposed to the accentedness of the segment, native word tokens were also generated following the same procedure by replacing the target native segment with the same segment extracted from a different repetition of the same word. An overlap of between 10 and 50 ms was applied between the spliced segment and the rest of the word in order to smooth the transition from one to another. A detailed explanation of the splicing technique is provided in García Lecumberri et al. (2014).

Recordings took place in the acoustic booth of the Phonetics Laboratory in the Faculty of Letters at the University of the Basque Country via a Fireface external sound card, controlled with Audacity software.

2.3. Choice of consonant inventory

The 9 English target consonants listed in the first column of Table 1 were chosen to explore differences between the phonological systems of English, Spanish and Czech. English targets were replaced with instances of their most common Spanish-accented realisations, as shown in the second column.

In the Czech phonological system, voiceless plosives are not aspirated, so [k] and [t] are closer to the Spanish system than to English.

Czech rhotics are also closer to the Spanish versions, with a trill phoneme /r/ which is usually realised as [r] but can also be pronounced as [r̄]. /r/ and /r̄/ are different phonemes in Spanish, while in English the rhotic phoneme is pronounced [ɹ]. The English voiceless glottal fricative [h] does not exist either in Czech or in Spanish, but there is a phonemic contrast in Czech between the voiceless velar fricative [x] (a sound which is phonemic in the Spanish system) and the voiced glottal fricative [ɦ]. Assimilation in Czech results in devoicing at the end of the word before a pause, so these two sounds contrast in cases like *báh*, pronounced [bu:x] but [boɦa] in genitive (Dankovičová, 1997). In English, sounds [j] and [ɟ̞] are distinguished at the phonemic level, but in Spanish these two sounds are possible realisations of one phoneme, /j/. In Czech [ɟ̞] did not use to be part of the phonological inventory, but it has recently started to be considered as a new phoneme, used mainly in loan-words (Duběda, 2020). For this reason [j] and [ɟ̞] are distinguished in Czech at the phonemic level, as in English. Similarly, [b] and [v] are different phonemes in Czech and English, but not in Spanish. In Czech, [w] is historically related to [v] (Romportl, 1973), and has been developing into a voiced labiodental fricative for several centuries. Finally, the English phoneme /ð/ is not present in the Spanish or Czech phonological systems, but is an allophone of /d/ in Spanish.

2.4. Procedure

Listeners heard a total of 176 stimuli (88 words in both native and accented forms). Native and foreign-accented tokens were mixed, and stimuli were presented in a different randomised order for each participant.

The experiment consisted of two tasks, carried out in two separate sessions that took place within 48 h of each other. In the first session listeners performed a word identification task, in which participants typed the word they heard into a text box in an on-screen interface. In the second session, listeners were asked to rate each token by selecting a number on a 7-point scale labelled “strength of foreign accent”. Following Munro et al. (2006), the endpoints of this scale were labelled as “(1) Native-like” and “(7) Very strong”. During this task, participants were able to see the word orthographically on the screen so that they were aware of the intended word. Both tasks were carried out using custom Matlab programs. On average, the tasks required 9.2 and 9.5 min respectively.

3. Results

3.1. Word identification

Since the main goal of this study was to examine FA intelligibility at a segmental level, in the word identification task a token was considered to be identified correctly if the target segment was correct. For example, for the word pair [t]old–[t^h]old, a response of *tool* would be considered as correct, since the target segment (i.e. [t^h]) was correctly identified. Correct answers were detected and scored by comparing the phonemic transcriptions of the target word with transcriptions of the participant’s response, using the British English Pronunciation Dictionary (Beep, 1997).

An analysis of responses revealed that listeners attempted to make lexical sense of each accented token. It was also evident that even though some words generated minimal pairs such as *van-ban* in the native and foreign-accented forms, listeners did not necessarily restrict their response to one of these forms. Table 2 displays the most frequently reported incorrect responses for each segment.

Fig. 1 presents absolute word identification results (left panel) as well as the change in accuracy from native-accented stimuli to foreign-accented stimuli (right panel). This relative measure is used in subsequent analyses since the absolute word identification rates naturally vary across and within cohorts due to potential differences in listeners’ English lexical knowledge.

Table 2

Most frequently reported incorrect responses for accented items in the word identification task by each group, for each of the 9 consonants in which the native sound (N) was replaced by its foreign-accented (FA) counterpart. Each target word/confusion pair is followed by the total number of occurrences in parentheses.

| N-FA pair | English | Spanish | Czech |
|-----------|---------------------|-----------------------|--------------------|
| x – h | hen/hand (1) | half/have (11) | hammer/hummer (3) |
| ʃ – r | rider/writer (2) | reach/rich (14) | rice/rise (12) |
| tʰ – t | towel/owl (3) | tom/tongue (4) | tom/tongue (4) |
| kʰ – k | kiss/this (2) | kitchen/teaching (13) | keeper/deeper (13) |
| ð – d | theirs/there's (2) | there/bear (8) | there/bear (15) |
| v – b | valley/ballet (2) | view/bail (8) | valley/ballet (8) |
| w – gw | wane/rain (4) | winter/gunther (7) | wane/vain (8) |
| j – dʒ | yawning/journey (3) | yawning/journey (4) | yes/jazz (10) |
| dʒ – j | jim/limb (3) | jam/young (5) | jam/yawn (4) |

Compared to the native English baseline of 13.4 points intelligibility loss due to accent, the Czech cohort suffered a bigger drop in intelligibility (18.3 points) due to the insertion of a foreign-accented segment, while Spanish listeners showed a smaller loss (8.2 points). These results suggest that sharing an L1 with the speaker is beneficial for the listener, while not sharing the L1 of the speaker nor the target language is detrimental, at least in the case of the language pairings tested here.

A linear mixed-effects model was used to compare differences in natively-spoken and accented words for each group. Individual participant responses were treated as random effects. Statistical analysis was carried out using R (R Core Team, 2017) using the *lmer* function from the *lme4* package (Bates et al., 2015). This model indicated a significant difference in the *cohort* factor at the $p < .001$ level (Table 3). A Student's t-test for independent means suggested that the difference in intelligibility loss between the English and Spanish cohorts was statistically significant [$p < .01, t = -3.06$], as were the differences between the Czech listeners and the English cohort [$p < .001, t = 3.36$] and between the Spanish and Czech cohort [$p < .001, t = 9.14$].

3.2. Degree of foreign accent

The other dimension measured in the current experiment was the perceived degree of foreign accent, which was assessed using a 7-point Likert scale. Fig. 2 shows the mean number of words reported by each cohort for every given DFA option. These are aggregated results from all the segments analysed in the experiment.

As expected, all cohorts perceived less foreign accent in the non-accented tokens, with ratings of 1 or 2 being the more frequently

chosen levels for the non-accented tokens compared to 6 or 7 for their accented counterparts. The distribution of ratings was more uniform for the accented tokens, suggesting that listeners were using the entire range, while in the case of non-accented tokens listeners were more often able to identify an unaccented token as lacking accent.

The native cohort were most decisive in accent ratings for non-accented tokens, providing the largest proportion of '1' responses, while Spanish listeners tended to use similar proportions of '1' and '2' responses to these tokens. Although mean ratings are not necessarily meaningful for ordinal scales such as that used here, and are not used in the statistical model, their ranking (English: 1.76, Spanish: 2.16, Czech:

Table 3

Mixed model results for main effect of cohort for the intelligibility loss results.

| Model: intelligibility loss ~ cohort + (1 participant) | | | | |
|----------------------------------------------------------|---------|-------|-------|-------|
| Fixed effects | β | Error | t | p |
| Intercept (cohort English) | 12.12 | 2.51 | 4.83 | <.001 |
| cohort Spanish | -6.58 | 2.99 | -2.19 | <.05 |
| cohort Czech | 7.97 | 2.82 | 2.83 | <.01 |
| Random effects | SD | | | |
| Intercept (participant) | 4.03 | | | |
| Residual | 19.06 | | | |

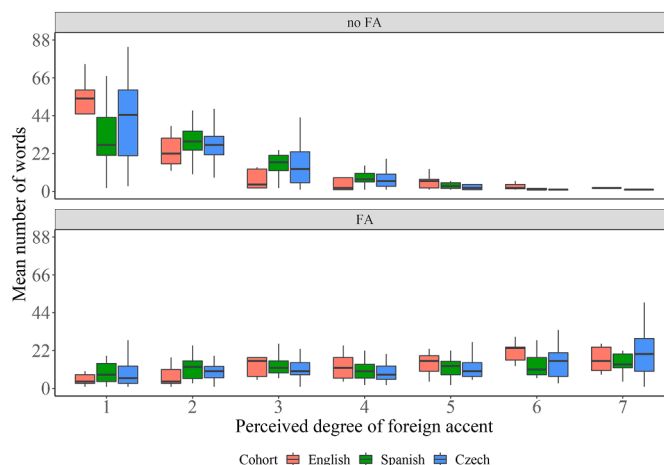


Fig. 2. Perceived degree of foreign accent, where 1 indicates 'Native-like' and 7 stands for 'Very strong' FA, for native (no FA) and foreign-accented (FA) words. Boxes signal the upper and lower quartiles, the horizontal bar represents the median, and the whiskers span to the lower and upper extremes of the reported data.

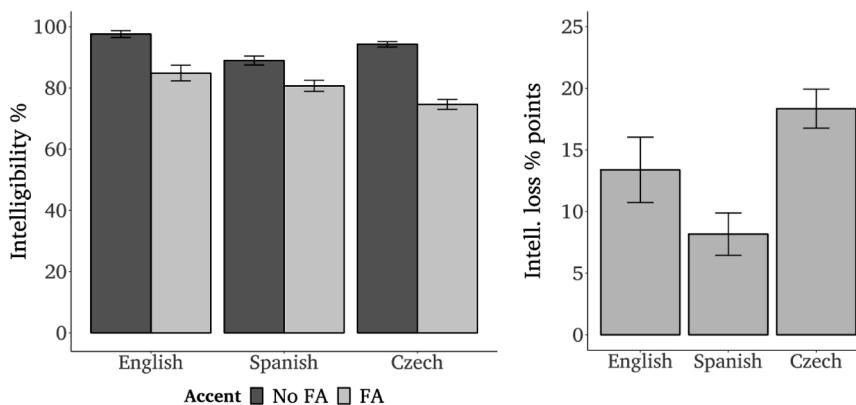


Fig. 1. Left: Intelligibility (defined as the percentage of words where the initial phoneme of the word was identified correctly) for natively-spoken words (No FA) and for words with a foreign-accented initial phoneme (FA). Right: intelligibility loss due to the accented phoneme. In this and subsequent plots error bars represent ± 1 standard error.

1.96) supports the tendencies described above. For the accented tokens, English listeners were more willing to use the higher values of the scale, while the Spanish cohort was least willing (mean ratings English: 4.63, Spanish: 4.35, Czech: 4.57)

Ratings were analysed with ordinal logistic regression via a cumulative link mixed model, using the *clmm* function from the package *ordinal* (Christensen, 2019) in R (R Core Team, 2017). The fixed effects were *cohort* (3 levels) and *accent* (2 levels) and the random effects were *participant* and *segment* (Table 4). The model revealed a significant interaction of *cohort* and *accent* [$p < .001$] and a significant effect for the presence of an accented sound [$p < .001$]. Post-hoc comparisons were carried out using the *joint_tests* function of the package *emmeans* (Lenth, 2021). A pairwise comparison revealed a significant difference between the English and Spanish cohorts for the no FA condition [$p < .05$] in that the latter considered non-accented items as having more FA.

3.3. Relationship between intelligibility and DFA

Individual segments were also analysed separately for DFA and intelligibility. Fig. 3 plots each segment in terms of mean perceived DFA for foreign-accented tokens against intelligibility loss.

A great degree of variability can be observed with regard to the relationship between intelligibility loss and DFA depending on the segment and the listener group.

The foreign accented realisations of [h] and [ʃ] ([x] and [r] respectively) were perceived as highly accented, but did not convey a noticeable loss in intelligibility for any cohort. For the English and Czech groups, a foreign realisation of segments [dʒ] and [j] (their FA realisations being [j] and [dʒ] respectively) led to a strong loss in intelligibility, which was correlated with a strongly perceived DFA, but Spanish listeners did not perceive these segments as particularly accented and did not suffer such an acute loss in intelligibility. Regarding [tʰ], only English listeners perceived its lack of aspiration as mildly-accented. This perceived DFA was accompanied by a moderate loss in intelligibility. A closer look at the results revealed that English listeners perceived the accented counterpart of [tʰ] (i.e. [t]) more frequently as [d] than the other two cohorts (e.g. they confused *touch* with *Dutch*). As for [v] realised as a weak voiced plosive [b], no cohort reported a high DFA, but the loss of intelligibility was quite noticeable, especially amongst Czech listeners. Similarly, Czech listeners also showed a bigger deficit in intelligibility than the other groups for [w] when pronounced as [gw]. Examination of participants' responses indicates that the Czechs were the only group for which a [w]/[v] confusion emerged, e.g. foreign-accented *veal* was perceived as *whale* ($n = 14$). Finally, the foreign accented counterparts of [kʰ] ([k]) and [ð] ([d]) were not rated as particularly accented, and no big loss of intelligibility was reported by any group. The relationship between these different segmental effects and their effects for each listener group will be discussed in what follows.

Table 4

Cumulative link mixed model summary for main effects of cohort and presence/absence of foreign accent for the DFA results. β -values are expressed as log odds (logits).

| Model: DFA ~ cohort × accent + (1 participant) + (1 segment) | | | | |
|------------------------------------------------------------------|---------|--------|---------|------------|
| | β | Error | z | p |
| cohort Spanish | 0.9226 | 0.3508 | 2.63 | $p < .01$ |
| cohort Czech | 0.4866 | 0.3304 | 1.473 | |
| accent yes | 3.8395 | 0.1017 | 37.75 | $p < .001$ |
| cohort Spanish: accent yes | -1.2115 | 0.1144 | -10.594 | $p < .001$ |
| cohort Czech: accent yes | -0.579 | 0.1081 | -5.354 | $p < .001$ |
| Random effects | | SD | | |
| Intercept (participant) | 0.8533 | | | |
| Intercept (sound) | 0.8292 | | | |

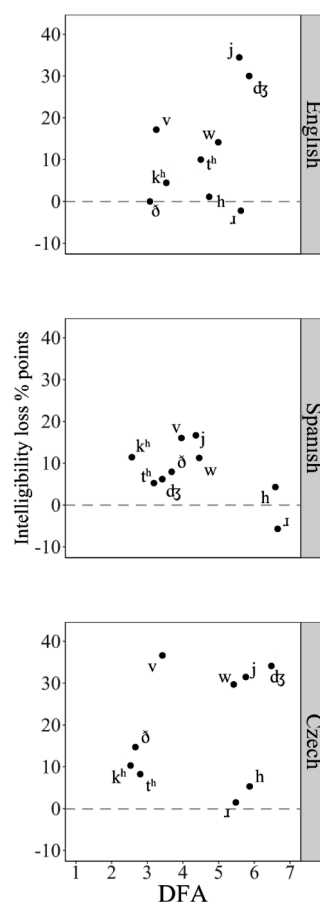


Fig. 3. Reported DFA of accented tokens against loss in intelligibility of each segment across cohorts. The phonetic symbol used for each segment represents the transcription of the target (i.e. the non-accented sound). Negative values represent an increment in intelligibility.

4. Discussion

In this study, foreign accent and its effects have been analysed at the segmental level. By means of acoustic manipulations, English word tokens were generated in which the consonant in initial position was replaced with the most common Spanish-accented counterpart. This technique contrasts with the holistic approach of previous FA research (Southwood and Flege, 1999; Munro et al., 2006; Bent and Bradlow, 2003; Stibbard and Lee, 2006) by providing tight control over the generated tokens and limiting the accented component to isolated, selected segments, while avoiding potential supra-segmental confounds. The segmental FA generation technique also prevents uncontrolled variables related to inter-speaker variation, since all the tokens are generated from a single voice.

4.1. Segment-level relationship between DFA and intelligibility

The first aim of the current study was to determine the nature of the relationship between perceived DFA and intelligibility when foreign accent is controlled at the level of individual segments. Across all segments, Figs. 1 and 2 suggest the absence of a strong relationship between DFA and intelligibility: considerable across-cohort differences in intelligibility loss due to FA are not reflected in differences in perceived DFA. However, results for individual segments (Fig. 3) manifest a more complex situation. While some strongly-accented segments result in substantial intelligibility loss for some cohorts (e.g., [j], [dʒ] for native and Czech listeners), others depart from a hypothetical linear relationship between DFA and intelligibility loss. This is especially noticeable in

the case of segments such as [x] (for [h]) and [r] (for [ʃ]), which are ‘bad’ exemplars of the target category, in the terms of Kuhl (1991), but which do not fall within the scope of any other competing phonological category and are therefore intelligible to all listener groups even though they are judged as heavily-accented.

Accordingly, the answer to our first research question is that indeed the relationship between intelligibility and FA is segment-dependent. Our findings suggest that the intuitive idea that “a stronger foreign accent leads to a higher loss in intelligibility” may be true for some segments, even though it does not apply to all segments nor equally for each one.

In the light of these findings, it is possible that the moderate relationship found in holistic studies between these two dimensions (Munro and Derwing, 1995a; Kashiwagi et al., 2006) is due, in part, to this heterogeneous contribution of individual segments to FA and intelligibility. The loss in intelligibility in larger utterances can vary greatly depending on the weight of each native/accented segment in the experimental materials used in a particular study, which would account for divergences between different studies (for example, Munro et al., 2006 study the FA of whole sentences while Bent and Bradlow 2003 focus on key words within sentences).

4.2. Listener-speaker L1 interactions

The second goal of the study was to determine whether the relationship between DFA and intelligibility is affected by listener-speaker L1 interactions at the phonological level. Bent and Bradlow (2003) found that for non-native listeners, a non-native speaker sharing their L1 was as intelligible as a native speaker, and that this benefit also extended to non-native listeners with a different L1 from that of the FA speaker. However, Stibbard and Lee (2006) showed that, in the case of low-proficiency non-native speech, non-native listeners from a different L1 did not find accented speech as intelligible as native speech, questioning therefore what Bent and Bradlow (2003) termed the mismatched interlanguage speech intelligibility benefit.

Methodological differences between our study and that of Bent and Bradlow (2003) do not allow a direct comparison of both sets of results regarding matched interlanguage speech intelligibility benefits. Nevertheless, we observed that overall the loss in intelligibility is less pronounced in listeners whose L1 matches that of the foreign accent, which agrees with their findings. In the current study, Spanish listeners produced similar accent ratings for foreign accented tokens as the English and Czech cohorts, while at the same time exhibiting a substantially smaller intelligibility loss. On the other hand, for the mismatched L1 group (i.e., the Czech cohort) not only was the Spanish FA not as intelligible as the native realisations, but, contrary to what Bent and Bradlow (2003) suggest, these listeners were the ones who suffered the biggest loss in intelligibility due to FA, which agrees with what Stibbard and Lee (2006) found for higher-level structures. It is worth noting that in our study this greater loss in intelligibility by the Czech cohort is not accompanied by a significantly different perception of DFA when compared to that of native English listeners, which shows that the mismatched group was as sensitive to FA as the native group. This finding suggests that the DFA-intelligibility relationship varies not only across segments, as we found for our first research question, but also according to the L1 of the listener. Specifically, Czech listeners are neither matched in terms of the phonological system of the target language (English) or the foreign language (Spanish). Therefore, it is plausible that they are less able than Spanish and English listeners to map the target language-foreign language phonological system relationship, leading to the biggest loss in intelligibility amongst Czech listeners.

The segmental focus of our study allows us to conclude that it is each individual accented segment’s similarity to and role in the listeners’ L1 phonological system which accounts for the relationship between intelligibility and DFA. Some segments that are judged to have a mild

DFA can have an adverse effect on intelligibility because the accented realisation falls within the scope of one or several phonological categories in that listener’s L1 system. This is the case for English [v] pronounced as [b] in Spanish-accented English, which falls into the English /b/ category, and which for Czech listeners also falls within the scope of [w], a trait common to other Slavic languages like Russian (Bondarenko, 2014), Ukrainian (Melezhik, 2014) or Slovak (Jureková, 2015; Metruk, 2018). A similar situation is that of the realisations of [dʒ] (for [j]) and [j] (for [dʒ]), which both English and Czech distinguish at the phonemic level, and therefore cause a very large intelligibility drop. For the Spanish listeners, in whose L1 they are allophones, the detriment is much smaller.

The case of [t^h] also shows that intelligibility reductions are closely related to the role and characteristics of each sound in a listener’s L1. Since (unaspirated) [t] is phonemic in both Spanish and Czech, its foreign-accented realisation in English is nearly as intelligible for the Czech group as for the Spanish listeners and considered moderately foreign-accented for both groups. On the other hand, native listeners find it more accented and less intelligible, since this unaspirated [t] may fall within the English /d/ category (Flege and Eefting, 1988).

Finally, there are segments which are judged as very strongly accented and yet they are quite intelligible for all groups ([x] for [h] and [r] for [ʃ]) because they either have no competitors in the listeners’ phonological system (as is the case in English), because they are contrastive in their system (velar and glottal fricatives in Czech) or because these accented realisations match those of the listeners’ L1 (as is the case in Spanish for both sounds and for /r/ in Czech). The fact that they are considered very accented by even the matched-L1 group is probably related to their high acoustic salience.

Our results suggest that FA intelligibility is closely associated with differences between the phonological system of the target language and the listeners’ L1s, in agreement with Imai et al. (2005)’s hypothesis. The current study further shows that this relationship needs to be examined in a per-segment manner. Individual segments with similar representations in the phonological system of the speaker and the listener can produce a smaller intelligibility loss than segments with different or no representation in the L1 of the listener. It is, therefore, crucial to include in studies of foreign accent an analysis of phonetic representations and roles of accented segments in the listeners’ L1 in order to fully understand where the benefit or detriment in FA intelligibility for different listener populations stems from, and to be able to characterise the relationship between degree of FA and intelligibility.

4.3. Limitations

Given that the intelligibility task demanded that listeners could not be presented with the target tokens prior to the main experiment, it was not possible to measure each cohort’s familiarity with this specific set of words. For this reason, we presented both absolute intelligibility scores and relative intelligibility losses in Fig. 1, but analyse only intelligibility loss for accented and non-accented tokens. While it would have been possible to tackle this issue through the use of an independent listener cohort in each native language represented here, this was not feasible in the current study.

5. Conclusions

The lack of a straightforward relationship between DFA and intelligibility has been observed in previous studies. By constraining accent to individual consonantal segments, the current study demonstrates that the study of the segmental level is necessary to fully understand the relationship between DFA and intelligibility. Moreover, the effect of individual segments on these factors is shown to depend on their status and characteristics in the L1 phonological system of the listener and the speaker. Taken together, these findings help explain the segmental basis for foreign accent and its effect on intelligibility in holistic studies of

accent, and elucidate the role of listeners' and speakers' phonological systems in determining whether non-native listeners benefit from non-native speech in a shared or non-shared interlanguage.

CRedit authorship contribution statement

Rubén Pérez-Ramón: Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Maria Luisa García Lecumberri:** Conceptualization, Writing – original draft, Supervision, Project administration, Funding acquisition. **Martin Cooke:** Conceptualization, Formal analysis, Validation, Writing – review & editing, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Spanish MINECO project DIACEX (FFI2012-31579). The authors thank Jan Volin and his team at the Charles University for their help in the recruitment of the Czech participants and providing facilities for running the experiment. Thanks also to Radek Skarnitzl for discussions about the Czech phonological system.

References

- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed effects models using lme4. *J. Stat. Softw.* 67, 1–48. <https://doi.org/10.18637/jss.v067.i01/>.
- Beep, 1997. Beep: phonemic transcriptions of over 250,000 English words. (British English pronunciations). URL: <http://svr-www.eng.cam.ac.uk/comp.speech/Section1/Lexical/beep.html/>.
- Bent, T., Bradlow, A.R., 2003. The interlanguage speech intelligibility benefit. *J. Acoust. Soc. Am.* 114, 1600–1610. <https://doi.org/10.1121/1.1603234/>.
- Bondarenko, O., 2014. Does Russian English exist? *Am. J. Educ. Res.* 2 (9), 832–839. <https://doi.org/10.12691/education-2-9-20/>.
- Christensen, R.H.B., 2019. Ordinal - Regression Models for Ordinal Data. R package version 2019.12-10. <https://CRAN.R-project.org/package=ordinal/>.
- Council of Europe, 2001. Common European framework of Reference For languages: Learning, teaching, Assessment. Press Syndicate of the University of Cambridge, Cambridge, U.K.
- Dankovičová, J., 1997. Czech. *J. Int. Phon. Assoc.* 27, 77–80. <https://doi.org/10.1017/s0025100300005442/>.
- Derwing, T.M., Munro, M.J., 1997. Accent, intelligibility, and comprehensibility: evidence from four L1s. *Stud. Second Lang. Acquis.* 19, 1–16.
- Duběda, T., 2020. The phonology of anglicisms in French, German and Czech: a contrastive approach. *J. Lang. Contact* 13 (2), 327–350.
- Flège, J.E., Eefting, W., 1988. Imitation of a VOT continuum by native speakers of English and Spanish: evidence for phonetic category formation. *J. Acoust. Soc. Am.* 83 (2), 729–740.
- García Lecumberri, M.L., Barra-Chicote, R., Pérez-Ramón, R., Yamagishi, J., Cooke, M., 2014. Generating segmental foreign accent. In: 15th Annual Conference of the International Speech Communication Association (Interspeech 2014), pp. 1302–1306.
- Imai, S., Walley, A., Flège, J.E., 2005. Lexical frequency and neighbourhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *J. Acoust. Soc. Am.* 117, 896–907. <https://doi.org/10.1121/1.1823291/>.
- Jureková, P., 2015. The Pronunciation of English in Czech, Slovak and Russian speakers [Unpublished Bachelor's Thesis]. Masaryk University.
- Kang, O., Rubin, D., Pickering, L., 2010. Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Lang. J.* 94, 554–566. <https://doi.org/10.1111/j.15404781.2010.01091.x/>.
- Kashiwagi, A., Snyder, M., Craig, J., 2006. Suprasegmentals vs. segmentals: NNS phonological errors leading to actual miscommunication. *JACET Bull.* 43, 43–57.
- Korpal, P., Sobkowiak, M., 2020. The perception of native vs. non-native Danish speech: bent and Bradlow's matched interlanguage speech intelligibility benefit revisited. *Scand. Philol.* 18 (2), 284–296. [doi:10.21638/11701/spbu21.2020.204/](https://doi.org/10.21638/11701/spbu21.2020.204/).
- Kuhl, P.K., 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Percept. Psychophys.* 50, 93–107.
- Lenth, R.V., 2021. emmeans: Estimated Marginal Means, aka Least-Squares Means. R Package Version 1.5.4. <https://CRAN.R-project.org/package=emmeans/>.
- Major, R.C., 2001. Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology. Routledge.
- Major, R.C., Fitzmaurice, S.F., Bunta, F., Balasubramanian, C., 2002. The effects of nonnative accents on listening comprehension: implications for ESL assessment. *TESOL Quarterly* 36, 173–190. [doi:10.2307/3588329/](https://doi.org/10.2307/3588329/).
- Melezhik, K., 2014. Typological markedness in Ukrainian students' ELF production. *Konińskie Studia Językowe* 2 (2), 191–200.
- Metruk, R., 2018. 'West' or 'Vest'? Pronunciation of English consonants [w] and [v] in the utterances of Slovak EFL speakers. *J. Lang. Educ.* 4 (2), 24–29.
- Munro, M.J., 1998. The effects of noise on the intelligibility of foreign accented speech. *Stud. Second Lang. Acquis.* 20, 139–154.
- Munro, M.J., Derwing, T.M., 1995a. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Lang. Learn.* 45, 73–97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x/>.
- Munro, M.J., Derwing, T.M., 1995b. Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Lang. Speech.* 38, 289–306. [doi:10.1177/002383099503800305/](https://doi.org/10.1177/002383099503800305/).
- Munro, M.J., Derwing, T.M., 2001. Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Stud. Second Lang. Acquis.* 23, 451–468.
- Munro, M.J., Derwing, T.M., Morton, S.L., 2006. The mutual intelligibility of L2 speech. *Stud. Second Lang. Acquis.* 28, 111–131. <https://doi.org/10.1017/s027226310600049/>.
- Munro, M.J., Derwing, T.M., 2020. Foreign accent, comprehensibility and intelligibility, redux. *J. Second Lang. Pronunc.* 6 (3), 283–309.
- Pérez-Ramón, R., Cooke, M., Lecumberri, M.L., 2020. Is segmental foreign accent perceived categorically? *Speech Commun.* 117, 28–37.
- Pinet, M., Iverson, P., Huckvale, M., 2011. Second-language experience and speech-in-noise recognition: effects of talker-listener accent similarity. *J. Acoust. Soc. Am.* 130, 1653–1662. <https://doi.org/10.1121/1.3613698/>.
- R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Romportl, M., 1973. On the Czech system of consonants. *Studies in Phonetics.* De Gruyter Mouton, Berlin, Boston, pp. 105–117. <https://doi.org/10.1515/9783110822021-011/>.
- Safa, A.J., 2018. An overview of cross-linguistic influence in language learning. *J. Appl. Linguist. Lang. Res.* 5 (3), 186–203.
- Sereno, J., Lammers, L., Jongman, A., 2016. The relative contribution of segments and intonation to the perception of foreign-accented speech. *Appl. Psycholinguist.* 37 (2), 303–322.
- Shin, D.J., 2018. Effect of experience to language on speech-in-noise recognition for Korean L2 speakers. *Stud. Engl. Lang. Lit.* 44, 301–319. [doi:10.21559/aellk.2018.44.2.015/](https://doi.org/10.21559/aellk.2018.44.2.015/).
- Southwood, M.H., Flège, J.E., 1999. Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clin. Linguist. Phon.* 13, 335–349.
- Stibbard, R.M., Lee, J.I., 2006. Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis. *J. Acoust. Soc. Am.* 120, 433–442. <https://doi.org/10.1121/1.2203595/>.
- Stringer, L., Iverson, P., 2019. Accent intelligibility differences in noise across native and nonnative accents: effects of talker-listener pairing at acoustic-phonetic and lexical levels. *J. Speech Lang. Hear. Res.* 62 (7), 2213–2226.
- Tajima, K., Port, R., Dalby, J., 1997. Effects of temporal correction on intelligibility of foreign-accented English. *J. Phon.* 25, 1–24. [doi:10.1006/jpho.1996.0031/](https://doi.org/10.1006/jpho.1996.0031/).
- Thomson, R.I., 2018. Measurement of accentedness, intelligibility, and comprehensibility. *Assess. Second Lang. Pronunc.* 11–29.
- Varonis, E.M., Gass, S., 1982. The comprehensibility of non-native speech. *Stud. Second Lang. Acquis.* 4, 114–136.
- Van Heuven, W.J., Mander, P., Keuleers, E., Brysbaert, M., 2014. SUBTLEX-UK: a new and improved word frequency database for British English. *Q. J. Exp. Psychol.* 67 (6), 1176–1190.
- Van Wijngaarden, S.J., Steeneken, H.J.M., Houtgast, T., 2002. Quantifying the intelligibility of speech in noise for non-native listeners. *J. Acoust. Soc. Am.* 111, 1906–1916. <https://doi.org/10.1121/1.1456928/>.
- Zahro, S.K., 2019. Native and non-native listeners perceptual judgement of English accentedness, intelligibility, and acceptability of Indonesian speakers. *Lingua Cultura* 13 (1), 39–44. [doi:10.21512/lc.v13i1.5362/](https://doi.org/10.21512/lc.v13i1.5362/).
- Zipf, G.K., 2016. Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Ravenio Books.