December 2018

Non-Native Attitudes to /θ/ and /ð/: A European Case Study

Rias Van den Doel
Utrecht University, the Netherlands, w.z.vandendoel@uu.nl

Anne-France C. H. Pinget
Utrecht University, the Netherlands, a.c.h.pinget@uu.nl

Hugo Quené
Utrecht University, the Netherlands, h.quene@uu.nl

NON-NATIVE ATTITUDES TO /θ/ AND /ð/: A EUROPEAN CASE STUDY*

RIAS VAN DEN DOEL
Utrecht University, the Netherlands
w.z.vandendoel@uu.nl

ANNE-FRANCE C.H. PINGET
Utrecht University, the Netherlands
a.c.h.pinget@uu.nl

HUGO QUENÉ
Utrecht University, the Netherlands
h.quene@uu.nl

Abstract
This paper investigates the evaluation of the English sounds /θ/ and /ð/ as produced by European non-native speakers. Using the data from a larger web survey, we compared the error judgements by different native and non-native users of English. This was done to establish whether there is any normative convergence among European non-native speakers, or if this was counteracted by other patterns, such as the presence or absence of these sounds in their L1s. Our analysis shows that while European non-native judges do not differ consistently from native-speakers in their judgements, there are also subtle differences between different groups of non-native speakers, implying that we should be careful not to generalise across groups about non-native attitudes to these sounds.

Keywords: attitudes, dental fricatives, endonormative, Euro-English, exonormative, non-native speakers

1. Introduction

Due to increasing globalisation, more communities are exposed to linguistic variation, putting them in a position to evaluate the many different language varieties and features they encounter. According to Blommaert (2001:126), this "has opened a new and wider space for measuring diversity *as aberrance* from newly reinforced or reinvented standards, customs and benchmarks. . . . Rejection and exclusion on the grounds of features of one's linguistic resources are rife, although such reactions need not always be categorical." The example Blommaert (2001: 81) provides of "the production of local, deviant normativity" by non-

---

native speakers (NNSs) of English in a South African township can illustrate how different, localised normative practices are emerging globally. Interestingly, Blommaert's description of a locally differentiated normativity is conceptually quite distinct from the idea of ever-converging norms induced by the shared interests and practices of larger communities (such as NNSs of English) as proposed by some other scholars. The notion of normative convergence is, for instance, implicit in early conceptualisations of a pan-European English by English as Lingua Franca (ELF) researchers such as Seidlhofer (2001). The latter has argued that Europe is "witnessing the emergence of an endonormative model of lingua franca English which will increasingly derive its norms of correctness and appropriacy from its own usage rather than that of the UK or the US" (2001:15). Another ELF researcher, Jenkins, has also addressed the issue of normative convergence within this context:

> For example, only two mainland European languages, Greek and Spanish, include the physiologically difficult sounds /θ/ and /ð/ in their pronunciation repertoires. . . . When they speak English, the majority of Europeans substitute these sounds with either /t/ and /d/ or /s/ and /z/. It is therefore unlikely that /θ/and /ð/ will be features of "Euro-English" accents. What is not clear at this early stage is whether the former – as used by many Italian and Scandinavian speakers of English, or the latter – as used by many French and German speakers of English, will ultimately become *the accepted norm*, or whether there will be scope for regional variation in this respect within "Euro-English" (2001: 17, our italics).

Even though Jenkins (2017: 343) has distanced herself from some of the views expressed in this article, it may still be interesting to explore European NNSs' normative attitudes. The rationale for this would not be merely to establish whether ELF researchers are justified in disowning any previous claims about European English. If any patterns are found in the assessment of specific features, this would help to support or reject categorical claims about the evaluative behaviour of users of English in terms of convergence or differentiation. We may, for instance, find that specific groups of NNSs do indeed agree on the relative insignificance of particular features of non-native speech. Alternatively, we may be able to uncover distinctly local patterns in NNSs' "production" of normativity – ranging from a preference for a nativised pronunciation model to a closer alignment with NS norms (idealised or otherwise), or even to a more truncated repertoire in accent evaluation (cf. Blommaert 2011: 213). If such differentiation is indeed attested, it would reflect the complex responses found in various groups to language features to which they are exposed through globalisation, in addition to any local variation in exposure to English and in levels of proficiency.[1]

    The English fricatives /θ/ and /ð/, labelled as dental in a British and as interdental in a North American context (Ladefoged and Maddieson 1997: 143–

---

[1]   Since the differential effects of proficiency on normative convergence do not feature largely in ELF conceptualisations of a pan-European English, we have not addressed the issue here.

144; but see Collins and Mees 1996: 141–2), are often described as difficult to master for NNSs. Despite their occurrence in a wide range of different languages, including varieties of Spanish, Greek, Arabic, Burmese and Swahili, these sounds are widely viewed as exotic or "marked" (Wells 2010), and this is seen as a factor in inhibiting their acquisition (Derwing and Munro 2015: 70). Seen as indexical of NS accents, many NNSs do indeed consider /θ/ and /ð/ (hereafter also referred to as dental fricatives, or DFs) to be salient potential stumbling-blocks – an added difficulty for some being the confusion arising from the representation of both sounds as the same digraph <th> (Collins and Mees 1996: 142).

DFs are often reported to be absent from many inner and outer circle varieties of English. Often, however, the substitution of either /θ/ or /ð/ by other sounds is dependent on phonological, semantic, pragmatic and sociophonetic contexts (cf. Schneider 2004: 1123). For instance, while /θ/ may be realised in Southern Irish English as a dental or an alveolar stop, some speakers may use [θ] word-finally "in careful speech or in reading style" (Hickey 2004: 330). Similarly, even though neither sound is used in Jamaican Creole, they both feature in more acrolectal Jamaican English (Devonish and Harry 2004: 477). This suggests that categorical claims about the absence of DFs from specific varieties (such as those in Walker 2010: 29) should not be taken at face value. This is not a trivial point, since the assumption that /θ/ or /ð/ do not occur in many varieties of English has contributed to the perception of their reduced relevance to learners (cf. Crystal 2001: 57). For instance, Kirkpatrick states:

> I feel sorry for poor learners of English who spend hours of classroom time trying to master the R[eceived] P[ronunciation] sounds of /θ/ and /ð/, as these are difficult sounds to learn if they do not exist in your language and, it turns out, they are not used in many varieties of English anyway (2007: 17).

Jenkins (2000: 138) has also argued that it is unnecessary for NNSs to acquire sounds such as the DFs that do not exist in all NS varieties. In her view, it is "unreasonable to have 'higher' expectations of non-native as opposed to native speakers" (p. 139). In addition, Jenkins (2000: 137) has claimed to have empirical evidence that DFs are irrelevant to intelligibility in lingua franca English. This is why Jenkins did not include these sounds in her well-known list of essential non-native pronunciation targets, the Lingua Franca Core or LFC (2000: 159ff.). While the LFC does not actually proscribe /θ/ and /ð/ in non-native English, their use may well be inappropriate in specific contexts. As Jenkins has pointed out, some native speakers (NSs) "will have to accept that it may already be better, depending on their E[nglish] as an I[nternational] L[anguage] interlocutor, to use substitutions of /θ/ and /ð/" (2000: 228). Anyone insisting that the LFC is exclusively concerned with intelligibility may be struck by the evaluative overtones of such pronouncements, which appear to be concerned with norm-based appropriacy.

Even though ELF researchers have moved away from their earlier interest in features (Jenkins et al. 2011: 288), this does not mean that all the suggestions of the LFC have been abandoned, such as the recommendation to drop DFs from the pronunciation curriculum (e.g. Walker 2010). This suggestion has even been adopted by vociferous critics of the LFC, such as MacKenzie (2014: 126), who also stresses the absence of DFs in many NS and NNS varieties of English. However, MacKenzie also recognises that didactic considerations should not only be based on "mere intelligibility", but also take account of NNSs' own attitudes and aspirations (2014: 132–3) – a point previously made by Van den Doel (2008). This is important, because the LFC's overriding concern with intelligibility may obscure any issues non-native users have with the acceptability of particular non-native pronunciations. If, as Jenkins has posited (2000: 137), non-native realisations of the DFs do not affect intelligibility in lingua franca English, any persistence on the part of NNSs in evaluating these as "errors" suggests that considerations of acceptability also play a part in their normative behaviour. It may even imply that such NNSs have become more responsive to the stigmatisation which which different groups of NSs tend to regard NNS realisations of /θ/ and /ð/ (Van den Doel 2006: 290; Jenkins 2000: 138), and in some cases actively contribute to such stigmatisation.

While ELF researchers tend to be more concerned with intelligibility than acceptability, and may to some extent have abandoned their earlier interest in endonormative convergence in European English, the issue of local European norms and varieties has continued to interest a few scholars. Apart from general studies on Euro-English (Mollin 2006), attention has been paid to lexico-grammar (Breiteneder 2009; Forche 2012), pragmatics (Klimczak-Pawlak 2014), and to describing specific European varieties of English (Bushfeld 2011; Salakhyan 2012; Kautzsch 2014; Edwards 2016). Apart from Jenkins's interesting suggestions about the pronunciation features of European English (2001:17), phonology has rarely been discussed. Without explicitly referring to the notion of European English, Beinhoff (2008, 2013) has investigated Greek and German listeners' attitudes to Greek and German accents in English. In addition, Van den Doel and Quené (2013) have investigated claims of emerging phonological norms in European non-native speakers (Eu-NNSs), but without considering individual sounds.

Recently, Modiano has drawn renewed attention to the notion of Euro-English, speculating that Brexit "will clear the sociolinguistic space for the emergence of an authentic European English" (2017: 313). The mixed reactions to his article suggest that the issue continues to be controversial, and could benefit from fresh scholarly attention. Of course, we may well want to be content with Schneider's claim that "empirical, realistic linguists . . . have consistently failed to identify such a variety", and that there is "no evidence for a homogenizing tendency likely to produce a single, reasonably coherent variety in the long run" (2017: 353). Admittedly, it may be difficult to produce evidence for the notion of European English as a "reasonably coherent variety." But an examination of Europeans'

attitudes to non-native English may still reveal homogenizing tendencies pointing to some kind of normative convergence. They may also attest to the distinctly local patterns of differentiation proposed by Blommaert (2011).

For instance, it would be interesting to know how European non-native speakers of English (Eu-NNSs) identified, prioritised and evaluated non-native realisations of two sounds which, although often considered to be indexical of NS accents, have been reported to be irrelevant to intelligibility. If Eu-NNSs as a group are indeed more accepting of any substitutions of DFs by other Europeans than their NS (or possibly non-European non-native) counterparts, such a more accommodating orientation to the NNS accents of their fellow Continentals may be interpreted in different ways. It could be variously seen as evidence of endonormative convergence, however limited, among Eu-NNS, or perhaps even of a pragmatic indifference, shared with other NNSs, to the preservation of non-essential phoneme contrasts. However, if such relative leniency is not attested, this would suggest that Eu-NNSs' judgements of NNS realisations of /θ/ and /ð/ are affected by factors other than "mere intelligibility" – such as acceptability. If, for instance, some NNSs have become susceptible to NS stigmatisations of NNS realisations of DFs, this would be of interest to those wishing to review any ELT pronunciation training curricula designed to serve such users' interests.

Since some Eu-NNSs actually have /θ/ and /ð/ in their phonological inventories (e.g. speakers of Greek and Castilian Spanish), it would be interesting to explore if speakers of these languages identify, prioritise and evaluate DF-substitutions any differently from judges whose languages do not feature any DFs at all (e.g. Dutch, Polish and Finnish).[2] This would reveal whether, in the case of these sounds, local norms override any pan-European normative convergence. It has already been shown that NS judgements of NNS realisations of /θ/ and /ð/ are affected by the way these sounds were realised in the accents the judges were familiar with (Van den Doel 2006). If such "accent parallelism" between linguistic background and non-acrolectal realisations also impacts NNSs' judgements of the speech of other NNSs, this would suggest that transferred L1 norms, rather than any considerations of the non-nativeness of the speech judged, play a part in lingua franca communication. This has not been investigated systematically with regard to DFs, but a precursor study by Beinhoff (2008) revealed that German and Greek listeners evaluated non-native realisations of /θ/ equally severely.

Given the discussion about NNSs' responses to NNS realisations of /θ/ and /ð/ in especially a European context (e.g. Jenkins 2001), we have attempted to determine if, and possibly to what extent, Eu-NNS listeners demonstrate any evaluative convergence in their identification, prioritisation and evaluation of these sounds as pronounced by fellow NNSs. In order to investigate this, we analysed additional data from the Internet survey described by Van den Doel and

---

[2]   For reasons of space, we have omitted any references to the allophonic and sociolinguistic variation of DFs in Spanish. For a concise overview of allophonic variation, see MacKenzie n.d.).

Quené (2013), which was designed to enable the assessment of speech samples produced by different NNSs of English from the European continent.

2. Methodology

The data used for our analysis are derived from an existing Internet survey (Van den Doel and Quené, n.d.), the details of which will only be discussed here where relevant to the present study. For an exhaustive description, see Van den Doel and Quené (2013). The survey was set up so as to allow European NNS speech samples to be judged by three different groups of respondents: (1) Eu-NNSs, drawn from Croatia, Greece, Poland, Spain and the Netherlands; (2) NSs of English from different Inner Circle countries; (3) non-European non-native speakers (NEu-NNSs) from the People's Republic of China. This made it possible to compare the judgements of the different groups, and interpret clear internal consistency within Group (1), or any significant divergences in evaluation between Group (1) and either Groups (2) or (3), as indications of any emerging Continental European norms for accent evaluation. Unlike in the precursor study, we will only be concerned with the three groups' identification, prioritisation and assessment of non-native realisations of DFs – where necessary in relation to the evaluation of other features.

As in Van den Doel (2006), the Internet survey employed was designed to be an accessible platform for both solicited and unsolicited respondents in different locations to download and listen to sound files of different NNSs, and to facilitate the selection and evaluation of any pronunciation features in the stimuli which respondents choose to identify as errors. The survey was kept deliberately short and simple in order to attract respondents with diverging interests and educational backgrounds, featuring simple instructions on pronunciation assessment and no more than 30 audio stimuli. These consisted of three different sentences read by two speakers (male and female) of five Continental European languages, only two of which, Greek and Castilian Spanish, included DFs in their phoneme inventories. The sentences used as audio stimuli had been taken from a larger reading passage entitled "Please Call Stella", as recorded by both NSs and NNSs at the Speech Accent Archive (Weinberger 2011), and had been selected on the basis of their potential for generating salient non-native accent features – DF substitutions in particular. The sentences included in the survey, which contained six tokens eligible for DF-substitution (in bold), were:

(1) And maybe a snack for her bro**th**er Bob.
(2) Ask her to bring **th**ese **th**ings wi**th** her from **th**e store.
(3) Five **th**ick slabs of blue cheese.

For most speakers of British Received Pronunciation, these sentences would feature two instances of initial /θ/ (*things, thick*), two of initial /ð/ (*these, the*), one

of medial /ð/ (*brother*), and one of final /ð/ *(with)*. Note that a minority of speakers of educated British English and some other varieties would use /θ/ in *with*, as would a majority of speakers of General American and Scottish English (Wells 2008: 904). Of course, it would have been preferable to use a word that consistently has final /ð/ in most reference varieties of English (such as *breathe*), but this was unfortunately not included in the relevant reading passage. Similarly, tokens involving medial and final /θ/ were also missing.

The non-native English accents included in the stimuli had been selected to represent major European language groups (Castilian Spanish for Romance, Dutch for Germanic, Finnish for Finno-Ugric, Greek for Hellenic, and Polish for Slavic) and to showcase diversity with the regard to the inclusion of /θ/ and /ð/. Auditory analysis by an expert NS phonetician showed that all Greek and Spanish speakers included in the study consistently pronounced all DFs, whereas this was not the case for the Dutch, Polish and Finnish speakers (whose L1s do not feature these sounds). It would have been ideal if the Dutch, Polish and Finnish speakers had also demonstrated more varied patterns of DF-substitution, but virtually no patterns other than **th**-stopping were attested. Of course, additional speaker variation in DF-substitution could have made the analysis possibly less reliable. Even so, there was already considerable variation between speakers, since perfectly matched guises are not easily produced, or indeed available from the Speech Accent Archive. Partly because respondents objected to the duration of an earlier pilot, we decided only to use a limited number of verbal guises with roughly comparable levels of proficiency. We considered this to be appropriate, since the focus of the experiment is on listeners' possibly converging attitudes to specific features found in the same speakers, rather than on the speakers' performance itself.

We also decided, in line with previous experiments of this kind, to ask respondents to identify any non-native realisations as "errors". While this may predispose respondents to judge the stimuli from an overly prescriptive or NS perspective, the use of an error-based framework is likely to make the experiment more accessible to non-linguists. If the experiment had been preceded by a discussion about the relative arbitrariness of errors within the context of English as a European lingua franca, this could have confused and biased any potential respondents as well. Consequently, we introduced the concept of European Englishes on the experiment's welcome page in fairly neutral terms, with a focus on intelligibility. In addition, since it took extra effort to identify any errors, the set-up of the experiment implicitly encouraged respondents not to designate any deviations from NS norms as erroneous. Thus, it could be argued that the number of errors identified by respondents is a direct reflection of their commitment to participating in the experiment. It could even be an indication that they were especially concerned with factors such as acceptability and stigmatisation.

As is described in Van den Doel and Quené (2013), the Speech Accent Archive recordings were edited, downsampled and each presented, in a random order, on separate web pages of a specially designed web survey. Apart from being asked

to provide a global assessment of the way each individual sentence had been pronounced, respondents were given the opportunity to identify zero to three errors in each of the utterances, by clicking on orthographic representations of *all* the phonemes which could conceivably be assessed as incorrectly pronounced, and to indicate the gravity of such errors, before being allowed to go on to the next web page (see Figure 1). Thus, for each stimulus, it was possible to record (1) participants' global evaluation of the utterance (not considered in the present article); (2) the number and nature of the errors they believed they heard; (3) the severity assigned to each of these errors. Global evaluations (1) were obligatory, but error identification (2) and severity rating (3) were not.

Figure 1. A sample page from the Internet survey held by Van den Doel and Quené (n.d.)

The survey, available without a password at let.uu.nl/~Rias.vandenDoel/personal/ wwstim/eureng/html/, was not specifically targeted at any groups, but since respondents had been approached primarily through the authors' academic network, an educational bias may have been likely (see Van den Doel and Quené 2013: 83). Some respondents, however, had been approached through social media, and all had been offered the opportunity to take part in a lottery for a small prize as an incentive. A more detailed description of the experiment may be found in Van den Doel and Quené (2013: 80–83).

Between 2 April 2009 and 25 April 2010, the responses of 373 participants were collected (see Van den Doel and Quené 2013). The analysis provided in this paper is only based on the responses of a subset of the whole respondent pool (n=350), because 23 participants did not identify any phonemes in the error identification task. Respondents in this subset consisted of (1) 279 self-identified NSs of Croatian (n=22), Dutch (n=121), Greek (n=28), Polish (n=89) and Spanish (n=19); (2) 40 self-identified NSs of English; and (3) 31 self-identified NSs of Chinese, i.e. NEu-NNSs of English. No additional information was logged about the language backgrounds of the NSs of English and of Chinese, making it impossible to compare and contrast the judgements of British and other NSs of

English (as in Van den Doel 2006), or to investigate any differences between Chinese languages or dialects.

All selected responses were subjected to multi-level modelling (Kreft and De Leeuw 1998; Luke 2004; Quené and Van den Bergh 2008). Each model took into account the variances between judges, between speakers, and between items. This implies that the resulting regression coefficients are "corrected" for random variation between and within judges. The multi-level analyses in this study were all performed with the R programming environment. Computations and evaluations were carried out with functions from the packages *lme4* (Bates et al. 2015) for R (R Development Core Team 2016).

In this way, two dependent variables were modelled:

i.   The *hit rate*, i.e. the probability of a sound being reported as an error by the judges in question. This was not only estimated for all sounds (*overall hit rate*), but also for all DFs only (*DF hit rate*). If any group's DF hit rate approximated the NS baseline value (intercept), this would mean that the number of DF errors identified by the listeners was similar to that selected by the NS. Additionally, if the proportion of DFs actually designated as errors by any group is compared to the proportion of all sounds being reported as erroneous, the resulting *relative hit rate* should indicate the prioritisation given to the selection of DFs as opposed to other potential errors.

ii.   The *error severity estimate*, or the degree (on a 5-point Likert scale) to which judges agreed with the statement "I think this is a serious error" with reference to any sound they had reported as erroneous. This can be calculated for all sounds (*overall error severity estimate*), but also for the DFs (*DF error severity estimate*). If the latter estimate is significantly higher or lower than the NS baseline value (intercept), this would signal a consistent deviation from NS norms in terms of the perceived gravity of the errors selected. As with the hit rates, it is also possible to compare the overall and the DF error severity estimates, in order to determine if any group evaluated DF errors significantly more or less strictly than other identified errors. This *relative error severity estimate* would be another important indicator of the priority given to DFs as opposed to other sounds by specific groups. This estimate, however, refers to the strictness with which the DFs were assessed, rather than to their selection as errors.

We felt the need to distinguish clearly between error identification and assessment, simply because judges sometimes report a great many errors which they consider unimportant, or report fewer errors than do other groups but assess these more critically. For instance, Van den Doel's (2006) investigation of NS evaluation of Dutch-accented pronunciation features revealed how British judges in particular tended to over-report certain errors while simultaneously denying their significance, and how North American respondents would proffer stricter evaluations of the lower number of features they judged to be erroneous. Van den Doel (2006) speculates that such trends may be informed by underlying attitudes

to errors as either "noticeable but not serious" or "serious only where noticeable" (see Van den Doel 2006: 297 for details). Interestingly, the North American tendency to be stricter about fewer errors was also attested for a number of DF-substitutions (p. 297) – a useful reminder that there may be considerable disagreement between different groups of NSs evaluating non-native realisations of DFs.

Because of differences between groups of NSs in the evaluation of DFs, we should not ascribe too much authority to individual NSs' judgements of non-native realisations of these. An estimate of the judgements given by a larger, varied group of NS respondents may be a more reliable indicator of the degree of approximation to acrolectal NS realisations of the DFs, and thus be used as a basis to compare and contrast the potentially different levels of divergence from these responses by both Eu-NNSs and NEu-NNSs. It is for this reason, and not with a view to prioritising NS practices or norms, that we decided to employ a NS baseline or intercept. This is also warranted by our research objectives, which are concerned with establishing significant patterns of variation in error evaluation among groups of judges, rather than with the actual performance of the speakers and the inevitable differences between them.

3. Results

3.1. Identification

In order to calculate the probability of a DF realisation being reported as an error by the different groups of judges, we investigated the hit rate. If a particular judge had marked such a realisation as an error, then we noted this as a "hit". These binary data (hit/miss) were analysed by means of mixed-effects logistic regression (GLMM, Quené and Van den Bergh 2008). The dependent variable in this analysis is based on the odds ratio of observing a hit: if the proportion of hits is $P=0.8$, then the corresponding odds ratio is $(P/(1-P))=4$. For computational purposes, we worked with the natural logarithm of this odds ratio, $\ln(P/(1-P))=1.39$, rather than with the odds ratio itself. The logistic regression model attempts to estimate the "log-odds" or "logit" of a hit (Hosmer and Lemeshow 2000; Pampel 2000), taking language background as a fixed predictor, and the variance between judges, speakers and between items as random factors. As was noted in Section 2, the NSs of English were used as the baseline group. In other words, the regression coefficient reported for the NS group constitutes the intercept or baseline, and the coefficients reported for the other groups constitute deviations (positive or negative) relative to this baseline. The mixed-effects logistic regression analysis of the DF hit rate has been summarised in Table 1.

Table 1. Estimated coefficients for log odds of the DF hit rate (marking a DF as an error), broken down by respondents' language backgrounds. For fixed effects, the regression coefficients are given; for random effects, their standard deviations are given

Fixed effects	Coefficient	(SE)	Z	P
NS (baseline)	−2.598	(0.389)	−6.684	<.001
Eu-NNSs	−0.467	(0.154)	−0.304	.761
NEu-NNSs	−0.413	(0.223)	−1.853	.064
Random effects	Std. Dev.	N		
Judges	0.763	350		
Speakers	0.770	10		
Sentences	0.460	3		
Residuals		21004		

The coefficients reported show that the estimated log odds of a hit for the baseline group of judges, the self-declared NSs of English, was −2.598, corresponding to an observed hit rate of 0.093 (thus 9.3% of the DF realisation marked as an error). The probability of marking a DF realisation as an error for the Eu-NNS judges (estimated log odds −2.598−.467, observed hit rate 0.093) was not significantly different from the NSs. The NEu-NNS (Chinese) judges marked fewer DF realisations as errors than did the NS group (estimated log odds −2.598−0.413, observed hit rate 0.067), which was only marginally significantly different from the NSs. This implies that there are no major consistent differences in DF hit rates between the various groups of respondents, whether native or non-native – a pattern which fails to reinforce the notion of any specifically European convergence with regard to DFs.

Since the accents in the stimuli were derived from various Eu-NNS groups, we have also broken down the DF hit rate according to these various groups, as illustrated in Figure 2, where the speakers' native language (i.e. the accent in English stimulus) is shown on the horizontal axis and the judges' native language is shown with the labels and lines.

Figure 2 shows some divergence between the judgements of different groups of Eu-NNS. Dutch, Croatian and Polish listeners selected a proportion of DFs as errors that did not statistically differ from the proportion selected by the NSs (indicated by a dashed line). However, there was a noticeable tendency for the Greek and Spanish participants to report significantly fewer DF realisations as errors than did the NS judges (i.e. below the dashed line). Not only do such patterns raise interesting questions about the rating behaviour of specific groups, but they also suggest that Eu-NNSs are not necessarily more in agreement about their appreciation of other European Englishes than are other groups of judges.
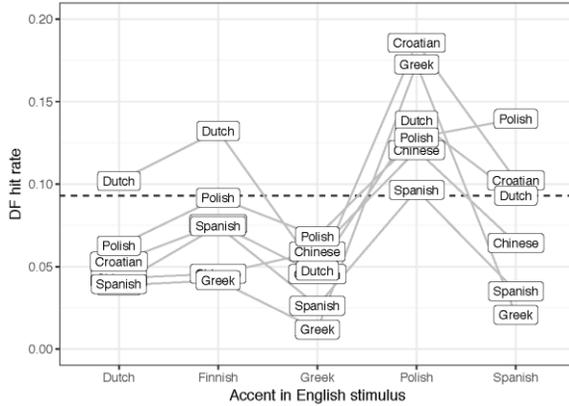
Figure 2. Observed *DF hit rate*, broken down by the speakers' native language (horizontal axis) and the judges' native language (Chinese, Croatian, Greek, Dutch, Polish and Spanish). The observed mean DF hit rate given by native English judges for all speakers is used as a baseline level (indicated by a dashed line).

We have also broken down the DF hit rates according to the type of DF (*fortis* /θ/ versus *lenis* /ð/), as illustrated in Figure 3, where the vertical axis still shows the probability of marking a DF as an error, and where the DF type is shown on the horizontal axis. The listeners' language backgrounds are shown in different colours. For this analysis, the words *with* and *the* (in Sentence 2) were left out of consideration for specific reasons. While the pronunciation of the DF in *with* is subject to regional variation, *the* was considered to be a function word with low perceptual salience, which was not selected even once by any of the respondents in our data set.
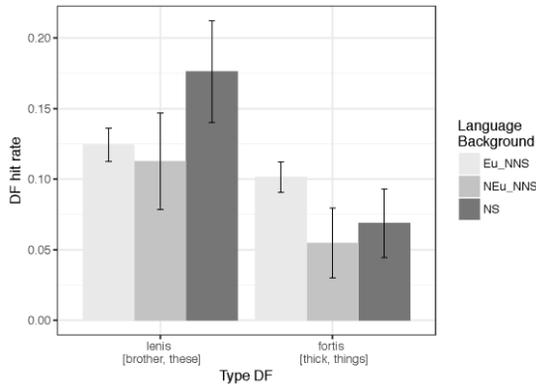

Figure 3. Observed *DF hit rate*, broken down by the type of dental fricatives (fortis or lenis) (horizontal axis) and the judges' language backgrounds. The error bars represent the 95% confidence interval.

Figure 3 shows that overall fortis /θ/ (in the rightmost part) was marked significantly less often as an error than lenis /ð/. Moreover, the NS in our experiment tended to select NNS realisations of /ð/ even more often than did NNSs. This difference in DF hit rate between the NS on the one hand and the Eu-NNS and the NEu-NNS (Chinese) judges on the other is significant. Conversely, the Eu-NNS were inclined to report /θ/ slightly more often as an error than did the NS and NEu-NNS judges.

3.2. Prioritisation

The relative hit rate, which reflects the likelihood of a marked error being a DF, serves as an indication of the prioritisation given to the selection of DFs as opposed to other selected errors. These binary data (DF/other sounds) were also analysed by means of mixed-effects logistic regression (GLMM). The fixed and random factors mentioned in Section 3.1 were also taken into account, and the NSs of English were once again used as the baseline group. The mixed-effects logistic regression analysis of the relative hit rate has been summarised in Table 2.

Table 2. Estimated coefficients for log odds of the *relative hit rate* (i.e. the probability of a marked error being a DF), broken down by respondents' language backgrounds. For fixed effects, regression coefficients are given; for random effects, their standard deviations are given.

Fixed effects	Coefficient	(SE)	Z	P
NS (baseline)	−1.687	(0.557)	−3.031	.002
Eu-NNSs	−0.286	(0.130)	−2.205	.027
NEu-NNSs	−0.385	(0.189)	−2.031	.042
Random effects	Std. Dev.	N		
Judges	0.508	350		
Speakers	0.913	10		
Sentences	0.796	3		
Residuals		11207		

The coefficients reported show that the estimated log-odds for the baseline group of judges, the self-declared NSs of English, was −1.687, corresponding to an observed relative hit rate of 0.213 (thus 21.3% of the marked errors being DFs). The probability of an error being a DF for the Eu-NNS judges (estimated log odds −1.687−0.286, observed relative hit rate 0.166) was slightly lower than for the NSs of English. The NEu-NNS (Chinese) judges selected fewer DFs than other sounds (estimated log odds −1.687−0.385, observed hit rate 0.149), which significantly differed from the NSs. The NSs thus tended to select proportionately more DF errors than other sounds, relative to the Continental European and Chinese judges.

The relative hit rate was also broken down according to the various groups of speakers and judges, as illustrated in Figure 4, where the vertical axis now shows the probability of a DF being marked as an error. The speakers' native language (i.e. the accent in the English stimuli) is shown on the horizontal axis and the judges' native language is represented by means of labels and lines.

Figure 4 reveals that there were some interesting patterns of divergence in the prioritisation of DFs for the different speakers. In particular, the relative hit rate is lower in the Greek and Spanish samples (for all listeners below the dashed line), especially where the judges are NSs of Greek and Spanish themselves. This means that – whenever an error is marked in these samples – it is only rarely a DF. Figure 4 also shows that there was some divergence between the different groups of Eu-NNSs. Unlike the Polish and Croatian judges, Dutch, Greek and Spanish respondents appeared to select DFs significantly less often as errors (as compared to other sounds) than did the NSs of English (indicated by a dashed line). This is evident from their lower relative hit rate. Once again, these patterns suggest that Eu-NNS listeners are not necessarily in agreement about the prioritisation of DFs when judging European Englishes. Some of these tended to agree with NS judges, and others with the Chinese.
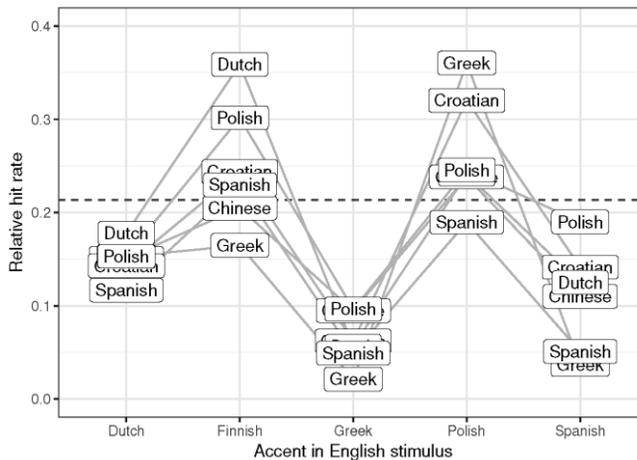

Figure 4. Observed *relative hit rate*, broken down by the speakers' native language (horizontal axis) and the judges' native language (Chinese, Croatian, Greek, Dutch, Polish and Spanish). For each native language, the observed mean relative hit rate given by native English judges for all speakers is used as a baseline level (indicated by a dashed line).

3.3. Evaluation

Error severity was measured on a 5-point scale. Subsequently, this measure was recoded in reverse order, as a result of which higher values signify a more severe judgement of the error (1=least severe, 5=most severe). The average error severity

for all other sounds (observed mean=3.41) was used as a baseline to centralise the data and obtain the relative error severity. This made it possible to determine if any groups evaluated DF errors significantly more or less strictly than other sounds. These relative error severity ratings were analysed by means of mixed-effects linear regression (LMM). Once again, the fixed and random factors mentioned in Section 3.1 were also taken into account, and the NSs of English were employed as the baseline group. The mixed-effects regression analysis of the relative error severity ratings has been summarised in Table 3.

Table 3. Estimated coefficients for the relative error severity (in 5-point scale units), broken down by respondents' language backgrounds. For fixed effects, regression coefficients are given; for random effects, their standard deviations are given.

Fixed effects	Estimate	(SE)	t	
NS (baseline)	0.103	(0.166)	0.621	n.s.
Eu-NNSs	−0.135	(0.153)	−0.880	n.s.
NEu-NNSs	−0.153	(0.223)	−0.686	n.s.
Random effects	Std. Dev.	N		
Judges	0.809	329		
Speakers	0.197	10		
Sentences	0.089	3		
Residuals	0.738	1896		

The coefficients reported indicate that the baseline group of judges, the self-declared NSs of English, showed an observed mean DF error severity (of 0.103) that did not significantly differ from zero (the baseline severity rating for all other sound errors). The Eu-NNSs and the NEu-NNS (i.e. Chinese) judges were as strict as the NSs, as is shown by the non-significant differences. Regardless of language background, all judges rated DFs as severely as the other errors.

A closer inspection of the DF severity estimates and the relative severity estimates did not reveal any significant differences between types of DF assessed (fortis vs. lenis) or between European listener groups. Clearly, while there were significant differences in the identification and prioritisation of DFs, this was not in any way reflected in their evaluation.

4. Discussion and Conclusion

Based on our analysis of patterns of identification, prioritisation and evaluation of non-native realisations of DFs, we were unable to verify any normative convergence among our Eu-NNS respondents. Most importantly, there were no significant differences between the Eu-NNSs and the NS judges in terms of DF hit rates or severity ratings, implying that these groups report roughly the same proportion of DFs as errors and that both groups evaluate these errors equally

strictly. Admittedly, the NSs tended to select proportionately more DF errors than other sounds relative to the Eu-NNSs as a group (the relative hit rate), but in this respect they were actually quite comparable to the Polish and Croatian judges. The Spanish, Greek and Dutch judges, however, reported relatively fewer DF errors – although in the Dutch case, this could be ascribed to their *overall* tendency to report more errors than NSs (see Van den Doel and Quené 2013). In short, there was no indication of any overall consistent differentiation between NSs on the one hand and Eu-NNSs on the other. It is especially striking that there should be no significant differences in evaluation between all groups. This may point to a continuing exonormative orientation towards English /θ/ and /ð/ on the European continent, rather than to any specific convergence between Eu-NNSs. This result is in keeping with other studies which failed to establish any clear evidence of emerging European norms for English, such as Mollin (2006) and Van den Doel and Quené (2013).

Secondly, the Chinese respondents showed lower DF hit rates than did either the NSs or the Eu-NNSs (a difference that was marginally significant), and a correspondingly lower relative hit rate than did the NSs. However, the Chinese judges' severity estimates showed no significant deviation from other groups. In other words, while the Chinese respondents appeared to be less inclined to report any non-native realisations of DFs as errors than the NSs of English, the fewer errors they reported were generally considered to be as severe as they would be by the other two groups. In addition, the Chinese tendency to detect fewer DF errors than some other groups was consistent with their overall low patterns of error detection, a tendency shared with the Greek and Spanish judges. This may suggest that, in terms of normative orientation, subtle differences exist between different groups of NNSs – at least when it comes to highly marked sounds such as /θ/ and /ð/. While the NEu-NNSs and Eu-NNSs may be in agreement with the NSs about the severity of DF errors, the Chinese respondents were less inclined, able or willing to report these as such. It would therefore seem unwarranted to generalise about NNSs as a group – at least where perceptions of DFs are concerned. Similarly, Van den Doel's (2006) investigation of NS evaluation of Dutch-accented pronunciation features showed that listener groups agreed on the gravity of specific cases of **th**-stopping, yet show structurally different patterns of detection (297). Van den Doel concluded from this that NS perceptions of DFs in NNS speech are far from uniform, and this may now be extended to NNSs.

Evidently, our data do not show any evidence of convergence among Continental Europeans, or even among groups of NNSs, in their judgements of DF realisations. Nonetheless, some idiosyncratic patterns present themselves in specific groups of NNS listeners. While the Chinese, Greek and Spanish listeners reported fewer DFs as errors, the Polish and Croatian listeners distinguished themselves by selecting relatively more of these than did other NNSs. These different tendencies may be accounted for in a number of ways, ranging from perceptual difficulties on the one hand to truncated repertoires in error detection and pedagogical traditions on the other – as is, for instance, done in Beinhoff

(2014), which showed an effect of proficiency in NNSs' evaluations of Spanish-accented speech. It would, however, be interesting to consider if transferred L1 norms can be invoked to explain some of the variation between the different groups of NNS listeners.

The notion that DFs are much less of a priority to NNSs because of their conspicuous absence in most phoneme inventories suggests a correlation between NNSs' attitudes to non-native realisations of /θ/ and /ð/ and their inclusion in listeners' first languages. However, we did not find any evidence for such a correlation. Even though there were some specific differences between the Greek and Spanish respondents and the NSs in terms of DF hit rate, these can be explained by pointing at overall group tendencies towards error detection and evaluation. Moreover, the Greek tendency to underreport errors of any kind (including DFs) may be difficult to explain on the basis of their L1. In addition, the lack of significant difference between Dutch, Polish and Croatian judges (whose languages do not feature /θ/ and /ð/), and the NS judges (most of whom are likely to be at least familiar with these sounds) on the other, may even be taken to mean that familiarity with DFs does not necessarily affect listeners' judgements – as was also found in Beinhoff (2008) for Greek and German judges.

It may be speculated that especially in the context of the present experiment, where both NS and NNS listeners were asked to report on specific errors, judges would be more inclined to resort to some sort of external, prescriptive pronunciation model rather than rely on their own judgements or any transferred L1 norms, but this cannot be verified with our data. Arguably, this could affect NNS judges disproportionately, if it is believed that NNSs by their very nature adopt NS or classroom norms more uncritically than do other groups. This would imply that, as "victims of an ideology that is imposed on them" they have been "brainwashed" by so-called native-speakerism (Jenkins 2007: 59, 187). However, if NNSs are viewed as a diverse group of stakeholders who do not respond to the languages they are exposed to as passive, docile and unthinking recipients (cf. Holiday 2006) and who may be credited with as much agency as NSs, this would argue against such conceptualisations of what should perhaps be termed "*non*-nativespeakerism". But even if we accept that any error bias will play less of a part in more realistic tasks or situations, the fact remains that the experimental conditions were identical for all respondents.

The absence of any effects which may be ascribed unequivocally to transferred L1 norms, and the lack of any significant divergence between Eu-NNSs and NSs, suggest that in this instance, local influences play a subordinate role. In addition, it seems unlikely that the similarity between the two groups can be explained by positing the adoption of more accommodating lingua franca norms for DFs by NSs. This is because non-native realisations of /θ/ and /ð/ are among the most commonly reported sources of "error" for all groups. What can be claimed, however, is that the convergence between Eu-NNSs and NSs cannot simply be extended to include the Chinese listeners, whose lower detection rate suggest either a relative unwillingness, or reduced ability, to report non-native realisations

of DFs as errors. Either could be taken as evidence that Chinese judges resort less to exonormative NS standards for the detection, prioritisation and evaluation of these sounds than do other groups of listeners. If so, it could be argued that a Lingua Franca Core which does not include the teaching of /θ/ and /ð/ to NNSs (as posited in Jenkins 2000) may be more relevant to some NNSs than to others.

Admittedly, our analysis shows that NSs were on the whole more inclined than NNSs to prioritise DFs in their identification of potential errors. A closer inspection of the data reveals, however, that this effect is due to the higher frequency with which the NS judges tended to select NNS realisations of /ð/ than did NNSs. Since this did not affect the evaluation of DF errors by any groups, perceptual differences between groups may have played a part in this. While both NS and NNS listeners may readily confuse /θ/ with other sounds (Hanulíkóva and Weber 2012: 615), it should be pointed out that /ð/ is a notoriously weak sound and may have less perceptual salience to NNSs. In any event, our results suggest clearly that it would be inadvisable to generalise about NSs' and NNSs' attitudes to DF errors without considering whether /θ/ and /ð/ errors may be perceived differently by specific groups. Such differentiation is in keeping with Van den Doel (2006: 239), who found that North American judges evaluated /θ/ and /ð/ errors structurally differently from other NS respondents. If it is true that specific groups of judges, whether NS or NNS, attach more importance to /ð/ than /θ/ errors, this would suggest, pace Jenkins (2000), that categorical claims about the relative insignificance of all DF errors to all groups of NNSs will need to be revisited. At the very least, it would be prudent to refrain from making such premature claims in textbooks aimed at any such groups (such as Walker 2010).

In an experiment of this kind, there will necessarily be a number of limitations. For instance, we were unable to take into consideration factors such as respondents' proficiency, which was not measured objectively, or their general attitudes to evaluation and language learning. In fact, as was also pointed out in Van den Doel and Quené (2013), it would have been interesting to explore the effect of "educational traditions which either favour or disfavour ambitious standards for language learning (and pronunciation training in particular)" (pp. 91–92), and of any local attitudes to "strictness and precision in education in general" (Edgar Schneider, as quoted in Van den Doel and Quené 2013). Since we used the same data set as in Van den Doel and Quené (2013), we have not been able to incorporate this.

Surveys such as EF-EPI (n.d.) and the Eurobarometer (2006) are routinely cited to support stereotypical impressions of proficiency in English in different groups of NNS, but these often either rely on very limited data or on self-reporting. Even if more inclusive or objective criteria are used to differentiate between different groups of NNSs in terms of proficiency, this does not help to account for all the findings of the present experiment, such as lack of significant differences between Dutch, Croatian and Polish respondents. It would therefore indeed be a good idea to investigate the link between proficiency and error assignment in any follow-up studies, as in Beinhoff (2014). This would be one way in which the

differential assessment of DFs by different groups of NNSs can then be more explicitly linked to effects in specific groups. This may be more insightful than attributing this to a possibly artificial, or even spurious, distinction between NSs on the one hand and NNSs on the other. In fact, it could be argued that all differences between groups of NNSs should be investigated separately, and that differences in language attitudes, in educational traditions and in perceptual difficulties should be taken into account, before any conclusions can be drawn about NNSs' overall assessment of DFs as a homogeneous group.

It would also be interesting to consider the effects of different DF-substitutions on NNS users of English. This is an issue not addressed in any detail in the LFC or in the present study, but it would seem highly unlikely that all substitutions will be regarded with the same level of leniency, given the part that some of these substitutions play in the recognition and stigmatisation of specific accents, such as **th**-alveolarisation in French and German English. Unfortunately, we do not have sufficient data to report on this at present. If it turns out, however, that such substitutions are evaluated very differently by NNSs, it would emphasise the importance of factors other than "mere intelligibility" in a lingua franca context.

As it is, there is no indication that non-native substitutions of /θ/ and /ð/ are evaluated more leniently by Eu-NNSs than by their NS counterparts. If we accept Jenkins's (2000) position that NNSs do not perceive such substitutions to be harmful to intelligibility, and at the same time find that they continue to report these as errors, we should entertain the possibility that some judges are swayed by considerations of acceptability. It may be suggested, of course, that such normative behaviour is merely reflective of internalised native-speakerism, and does not merit serious consideration other than as commentary on pervasive language teaching ideologies. However, if, as Jenkins has also suggested, "NNSs should have input into the determining of their pronunciation norms" (2007: 26), the views of different NNS stakeholders, whether exonormative, proficiency-based or the product of local normativity, should be reflected in this. This would include allowing for the possibility that some NNSs are more aware than others of the stigmatisation of specific accent features, and may even actively contribute to this. Jenkins (2000: 160) has claimed that "[t]here really is no justification for doggedly persisting in referring to an item as "an error" if the vast majority of the world's L2 English speakers produce and understand it". However, if we are to take the input of NNSs seriously, it may be argued that precisely such a justification may be found in NNSs' observed persistence in reporting DF-substitutions as errors.

References

Bates, Douglas, Maechler, Martin, Bolker, Ben and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1-48.

Beinhoff, Bettina. 2008. Looking for the 'real' native speaker: the perception of native and non-native English accents by non-native speakers of English. In Ewa Waniek-Klimczak (ed.), *Issues in Accents of English,* 120-139. Newcastle: Cambridge Scholars Publishing.

Beinhoff, Bettina. 2013. *Perceiving identity through accent: attitudes towards non-native speakers and their accents in English.* Oxford: Peter Lang.

Beinhoff, Bettina. 2014. Perceiving intelligibility and accentedness in non-native speech: a look at proficiency levels. *Concordia Working Papers in Applied Linguistics* 5. 58-72.

Blommaert, Jan. 2010. *The sociolinguistics of globalization.* Cambridge: Cambridge University Press.

Breiteneder, Angelika. 2009. English as a Lingua Franca in Europe: an empirical perspective. *World Englishes* 28. 256-69.

Buschfeld, Sarah. 2011. *The English language in Cyprus: an empirical investigation of variety status.* PhD dissertation University of Cologne, Germany.

Collins, Beverley and Inger M. Mees. 1996. *The PHONETICS of English and Dutch.* Leiden: Brill.

Crystal, David. 2001. The future of Englishes. In Anne Burns and Caroline Coffin (eds.), *Analysing English in a global context,* 53-64. London: Routledge.

Derwing, Tracey and Murray Munro. 2015. *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research.* Amsterdam: John Benjamins.

Devonish, Hubert and Otelemate Harry. 2004. Jamaican Creole and Jamaican English: phonology. In Edgar Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie and Clive Upton (eds.), *A handbook of varieties of English, vol. 1,* 450-480. Berlin: Mouton.

Edwards, Alison. 2016. *English in the Netherlands: functions, forms and attitudes.* Amsterdam: John Benjamins.

Eurobarometer. 2006. Special eurobarometer 64.3: Europeans and their languages. [Online] Available from http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/-ebs_243_en.pdf. [Accessed: 20 July 2018].

EF-EPI. n.d. The world's largest ranking of countries by English skills. [Online] Available from http://www.ef.nl/epi/. [Accessed: 20 July 2018].

Forche, Christian. 2012. On the emergence of Euro-English as a potential European variety of English: attitudes and interpretations. *Jezikoslovlje* 13. 447-478.

Hanulíková, Adriana and Andrea Weber. 2012. Sink positive: Linguistic experience with the substitutions influences nonnative word recognition. *Attention, Perception, and Psychophysics* 74. 613-629.

Holiday, Adrian. 2006. Native-speakerism. *ELT Journal* 60. 385-387.

Hosmer, David W. and Stanley Lemeshow. 2000. *Applied logistic regression.* New York: Wiley.

Hickey, Raymond. 2004. *Irish English: history and present-day forms.* Cambridge: Cambridge University Press.

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

Jenkins, Jennifer. 2001. 'Euro-English' accents. *English Today* 68. 16-9.

Jenkins, Jennifer. 2007. *English as a lingua franca: attitude and identity.* Oxford: Oxford University Press.

Jenkins, Jennifer. 2017. An ELF perspective on English in the post-Brexit EU. *World Englishes* 36. 343-346.

Jenkins, Jennifer, Cogo, Alessia and Martin Dewey. 2011. Review of developments in research into English as a Lingua Franca. *Language Teaching* 44. 281-315.

Kautzsch, Alexander. 2014. English in Germany: Retreating exonormative orientation and incipient nativization. In Sarah Buschfeld, Thomas Hoffmann, Magnus Huber and Alexander Kautzsch (eds.), *The Evolution of Englishes: The dynamic model and beyond,* 203-228. Amsterdam: John Benjamins.

Kirkpatrick, Andy. 2007. *World Englishes: implications for international communication and English language teaching.* Cambridge: Cambridge University Press.

Klimczak-Pawlak, Agata. 2014. *Towards the pragmatic core of English for European communication.* Cham: Springer.

Kreft, Ita and Jan de Leeuw. 1998. *Introducing multi-level modeling.* London: Sage Publications.

Ladefoged, Peter and Ian Maddieson. 1997. *The sounds of the world's languages.* Oxford: Blackwell.

Luke, Douglas. 2004. *Multilevel modeling.* Thousand Oaks, Ca.: Sage.

MacKenzie, Ian. 2014. *English as a lingua franca: theorizing and teaching English.* Abingdon: Routledge.

MacKenzie, Ian. n.d. The allophonic rules of Spanish. [Online] Available from: https://www.staff.ncl.ac.uk/i.e.mackenzie/allophonr.htm. [Accessed: 20th July 2018].

Modiano, Marko. 2017. English in a post-Brexit European Union. *World Englishes* 36. 313-327.

Mollin, Sandra. 2006. *Euro-English: assessing variety status.* Tübingen: Narr.

Pampel, Fred. 2000. *Logistic regression: a primer.* Thousand Oaks, CA: Sage Publications.

Quené, Hugo and Huub van den Bergh (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59. 413-425.

R Development Core Team. 2016. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. [Online] Available from http://www.R-project.org/ [Accessed: 1 December 2016].

Salakhyan, Elena. 2012. The emergence of Eastern European English. *World Englishes* 31: 331-350.

Schneider, Edgar. 2004. Global synopsis: phonetic and phonological variation. In Edgar Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie and Clive Upton (eds.), A *handbook of varieties of English, vol. 1,* 1111-1138. Berlin: Mouton.

Schneider, Edgar. 2017. The linguistic consequences of Brexit? No reason to get excited! *World Englishes* 36. 353-355.

Seidlhofer, Barbara. 2001. Towards making "Euro-English" a linguistic reality. *English Today* 68. 14-16.

Van den Doel, Rias. 2006. *How friendly are the natives? An evaluation of native-speaker judgments of foreign-accented British and American English*. Utrecht: LOT.

Van den Doel, Rias. 2008. The blind spots of Jenkins's Lingua Franca. In Ewa Waniek-Klimczak (ed.), *Issues in Accents of English,* 140-149. Newcastle: Cambridge Scholars Publishing.

Van den Doel, Rias, and Hugo Quené. n.d. The accents of European English: an online experiment. [Online] Available from let.uu.nl/~Rias.vandenDoel/personal/wwstim/eureng/html/. [Accessed: 20 July 2018].

Van den Doel, Rias, and Hugo Quené. 2013. The endonormative standards of European English: emerging or elusive? *English World-Wide* 34. 77-98.

Walker, Robin. 2010. *Teaching the pronunciation of English as a lingua franca.* Oxford: Oxford University Press.

Weinberger, Steven. 2011. Speech accent archive. [Online] Available from http://accent.gmu.edu. [Accessed: 20 July 2018].

Wells, John. 2008. With. *Longman pronunciation dictionary.* 3rd edn. Edinburgh: Pearson, 904.

Wells, John. 2010. Exotic dental fricatives. John Wells' Phonetic Blog. 7 April 2010. [Online] Available from http://phonetic-blog.blogspot.nl/2010/04/exotic-dental-fricatives.html. [Accessed: 20 July 2018].