Putting the Accent on Intelligibility: What Constitutes "good" Pronunciation in the Context of English as a Lingua Franca? A case study of learners of different LIS

Thomas Christiansen

Abstract
This paper examines the way in which the Lingua Franca Core (LFC), together with selected features of delivery, can be shown to affect intelligibility and thus constitute what a sample of learners regard as "good" pronunciation (cf. Christiansen 2011b). We do this by making reference to a questionnaire-based survey in which respondents have to rate for intelligibility recorded extracts of NS and NNS. We compare the results of this survey with a detailed phonological analysis of the extracts, in an attempt to ascertain whether the presence of LFC features as opposed to features of NS pronunciation (see Jenner 1997) corresponds to respondents' assessment of intelligibility. It is found that, while the degree of LFC or NS pronunciation in a given extract is a reliable indicator of its general intelligibility to an L2 learner, other observable and measurable aspects of each extract's delivery (e.g. speed, length of pauses and average number of discernible words between pauses) need to be taken into account for a more precise analysis.

Keywords: ELF pronunciation; ELF intelligibility; ELF prosodic features; native-speaker model of pronunciation.

1. Introduction

One of the most interesting recent trends in the teaching of English to Speakers of Other Languages (ESOL) has been the gradual move away from the idea that native speaker (NS) pronunciation represents the only valid model for learners (cf. Walker 2010; Seidlhofer 2011). This has come about both in recognition of the sheer number of different varieties of English that exist around the world and through the realisation that, in any case, much EIL (English as an International Language) communication occurs in English as a Lingua Franca contexts where the communication is mostly between non-native speakers (NNS). In particular, it has been suggested that a model

of pronunciation based on the vague idea of a 'native speaker' (a particularly difficult concept to define in regard to English or Englishes) should be replaced by a so-called core approach (Jenkins 1998, 2000, 2002, 2007; Walker 2010). Central to the latter is the concept of LFC (Lingua Franca Core), which rests on mutual intelligibility for the majority of participants actually engaged in the discourse whether or not they be NS and without privileging any single group, NS or not, as opposed to intelligibility for some hypothetical, and perhaps absent, NS.

In Sections 2 and 3 we discuss a survey whereby 174 learners in Albania, Italy and Germany assessed six short extracts principally as regards pronunciation and intelligibility. In Section 4, we compare the results with a detailed phonological analysis of the extracts, in an attempt to ascertain how far the presence of LFC features as opposed to features of so-called 'standard' NS pronunciation corresponds to respondents' assessment of intelligibility. In Section 5, we find it necessary also to examine how aspects of delivery affect intelligibility.

2. Background to speaker extracts

This study revolves around six short extracts (approx. 76 to 128 seconds in length) of speakers from four different countries (Austria, Germany, Italy, Poland, and the UK – see Table 1) taken from recorded interviews, precisely at the point where respondents were asked to speak for one minute on "what kind of person makes a best friend or companion?".

As part of an online survey, also using computer-generated voices to ascertain how far perception of grammatical accuracy is affected by pronunciation (see Christiansen 2011a), the extracts were assessed for pronunciation (P), fluency (F)[1] and intelligibility (I) by a panel of six experts (all trained and experienced oral examiners working for an international examining board in Italy and, by chance, NS British English). For the assessment of pronunciation, and fluency, the descriptors in the Common European Framework of Reference

[1] Fluency is problematic in that it is largely an intuitive concept that defies a simple objective definition (Luoma 2004: 88).

(Council of Europe, 2001)[2] were employed. There being no scale for intelligibility in the CEFR, one had to be drafted ad hoc. The mean scores of these various levels of assessment are summarized in Table 1[3]:

TABLE 1
Weighted mean scores for grammatical accuracy/fluency/intelligibility/pronunciation: comparison groups

	L1	Control Group (CG)			Total /270	Rank
		F /90	I /90	P /90		
Sp1	German (Germany)	66.67	77.5	57	201.17	5
Sp2	Polish	58.33	82.5	69	209.83	3
Sp3	Italian	71.67	85	78	234.67	2
Sp4	German (Austria)	63.33	72.5	66	201.83	4
Sp5	English (England)	88.33	87.5	90	265.83	1
Sp6	Italian	40	62.5	48	150.5	6

The final two columns give the total score of all the categories. This allows one to rank different speakers. Sp5 (NS) comes first, followed by Sp3, Sp2, Sp4, Sp1 and Sp6. For both comparison groups, both predominantly NS, the NS subject is in effect for each area a model of accurate linguistic performance, especially P, for which he is given full marks by the speaking-test experts without exception.

3. Questionnaire survey

To further examine the issue of which models of English are more intelligible to NNS learners, we drew up a questionnaire (Christiansen 2011b) to be administered by teachers to classes of students of B1-B2 level of the CEFR. In 3.1 and 3.2, we discuss the results.

[2] Grammatical accuracy (CEFR 2001: 114) and Phonological Control (*ibidem*: 117). In fact, in the latter there is some explicit bias against ELF e.g. band 3 (B1): "Speaker's pronunciation is clearly intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur".

[3] Weighting was necessary because the scales were different. Ninety serves as a convenient common base: the marks for P multiplied by 18, those for F by 10, GA by 11.25, and those for I by 15.

3.1. Methodology

In the questionnaire, respondents (174 in total) had to rate five statements against a simple five-point Likert scale: 1) the grammar of this speaker is generally good; 2) the vocabulary of this speaker is generally good; 3) the pronunciation of this speaker is generally good; 4) I can easily follow the general meaning of what this speaker is saying; 5) this speaker's accent is easy for me to understand. Finally they had to choose an alternative to answer the question: what is the speaker's native language? a) I can't say b) English c) French d) German e) Italian f) Polish g) Portuguese h) Spanish i) Swedish j) Turkish.

The key questions for our research are 3 and 5. Question 4 was included to disambiguate between the two general meanings of intelligibility; it refers to how the discourse is organized, while Question 5 focuses on the phonological features of the discourse (one of the key questions of this research)[4].

3.2. General results

Table 2, below, gives the mean for the various responses for questions 3 (pronunciation) and 5 (intelligibility) on the questionnaire (see 3.1), comparing the equivalent scores given by speaking-test experts[5].

TABLE 2
Summary of mean scores of questionnaire

	Pronunciation			Intelligibility		
	QR*	CG*	Diff*	QR*	CG*	Diff*
Sp1	62.84	57	+5.84	59.72	77.5	-17.78
Sp2	70.01	78	-7.99	68.68	85	-16.32
Sp3	58.86	69	-10.14	60.28	82.5	-22.22
Sp4	60.87	66	-5.13	62.63	72.5	-9.87
Sp5	48.41	90	-41.59	48.49	87.5	-39.01
Sp6	41.79	48	-6.21	55.78	62.5	-6.72
MEAN			-10.87			-18.65

*Key: QR = questionnaire respondents (174 language learners); CG = Control Group; Diff = Difference between QR and CG.

[4] The latter has been referred to as comprehensibility by Munro and Derwing (1995).
[5] As in Section 2, the scores are weighted.

The figures for the various means (bottom row) show that learners rate the extracts lower for both criteria than did speaking-test experts, with the exception of pronunciation, Sp1. The former's comparative lack of expertise and perhaps a lower tolerance of performance which they consider to be outside the norm may account for this. There is a striking difference in perceptions of intelligibility, which may in part be explicable by the fact that learners, such as NNS at B1-B2 level, would generally find discourse in English harder to follow than would NS speaking-test experts.

It is also interesting to look at the differences in marks awarded to individual speakers. Of these, the most striking are those for the NS (Sp5). Respondents disagree markedly with speaking-test experts, awarding this speaker considerably lower marks in both pronunciation and intelligibility. Conversely, in both pronunciation and intelligibility, respondents rate Sp2 most highly. This point is examined in greater depth in Christiansen 2011b, where it is shown that there is a general tendency on the part of respondents to award higher marks for NS pronunciation and intelligibility, perceived rather than real. This is not just interesting because of the questions it raises about the basis on which NNSs recognise NSs; it also leads one to question the practicality of the NS prototype that traditional methodologies see as central to language acquisition and consequently to assessment of language performance exists at all on a practical level. It is arguable that NNS learners, perhaps subconsciously, have acquired another model which has little directly to do with the way that NS of English speak but is rather based on the more universal and practical criteria of intelligibility.

Such an interpretation of the situation is in line with Jenkins (1998, 2000, 2002, 2007) who posits the existence of a so-called Lingua Franca Core (LFC), which we shall turn to in the next section, re-examining the data discussed in this section in the light of how it might confirm or contradict Jenkins' hypothesis.

4. Examining data from questionnaire in the light of the Lingua Franca Core (LFC)

To address the issue of intelligibility in NNS to NNS spoken discourse in ELF, or more specifically ELFE (English as a Lingua Franca in Europe), Jenkins gathered data from authentic interaction and drew

up a list of phonological features found as crucial to intelligibility. Jenkins (1998, 2000, 2002, 2007) calls these the *Lingua Franca Core*.

Such a core represents an alternative model to the NS (normally RP or General American), and has obvious implications for language teaching and testing. Most profoundly, it represents a rejection of a specific model whose aim is to sound like a specific type of user (a NS) in favour of functional criteria, with the object of making oneself intelligible to the largest number of speakers possible (see Jenkins 2007: 25).

It should also be noted, however, that it is incorrect to view the LFC and the traditional concept of standard NS as diametrically opposed. Many of the LFC features are also found in the standard, e.g. the features pertaining to consonant clusters and maintenance of contrast between long and short vowels, aspirated initial /p/ /t/ /k/. Other features are absent from the standard: e.g. "L2 regional qualities" in the category of vowel sounds. For our purposes here, we will establish a difference, labelling items as LFC only when they do not also fall within the standard.

In this study, we concentrate on segmental, phonological features, ignoring supra-segmental ones, namely those of pitch and rhythm, mainly because the nature of the recordings did not allow accurate enough transcription of these features, which , as can be seen, are of marginal importance, or relevant only in specific contexts (e.g. stress to signal contrast).

In particular, in the analysis reported below, we retrieved the following phonological phenomena, which we summarize as Table 3:

TABLE 3
Examples of LFC found in extracts

1) Consonant Inventory	2) Consonant Clusters
• Some substitutions of /θ/ and /ð/ acceptable: /d/ for /ð/; /d/ for /θ/; /s/ for /ð/; /s/ for /θ/; /t/ for /ð/; /t/ for /θ/; /v/ for /ð/ • Allophonic variation permissible within phoneme: /ɬ/ for /l/	• Addition acceptable: insert /j/; insert schwa; insert /ʊ/
3) Additional Phonetic Requirements	**4) Vowel sounds**
• Aspiration after word-initial voiceless stops /p/, /t/, /k/: Aspirate initial /k/, /p/, /t/; Aspirate /k/; Aspirate /t/	• L2 regional qualities acceptable if they are consistent, except /ɜ:/.
	5) Production / placement of tonic stress
	• Not applicable

In this section, we will focus on how the results reported in Section 3 compare to the LFC, i.e., whether speakers that respondents found more intelligible showed discernible features of LFC in their pronunciation.

In order to do this, we had detailed phonetic transcripts of the extracts made by two trained phonetic transcribers,[6] who were given no details about either the speakers or the purposes of the research.

After transcripts of each speaker were produced, we analysed each one, noting and counting divergent pronunciations from a 'standard' NS pronunciation, using as a guide the International Phonological Association (IPA) transcripts provided by the *Concise Oxford English Dictionary* (COED).

From such an analysis, it is possible to quantify how many LFC features the extract contains. This exercise inevitably involves an element of subjectivity. The size of the sample of phonemes produced by each speaker however reduces the margin of error. The extracts ranged from an estimated 358 to 940 tokens of phonemes (see Table 4), thus constituting a full selection of phonemes and combinations .

Analysis of each extract allows one to compare the proportion of core, standard and remainder in each, making it possible to measure their relative degrees of adherence both to Jenkins' Lingua Franca Core and to the supposed standard NS model. Results in Table 4, show the figure for total number of phonemes exclusive of filled pauses:

TABLE 4
Relative proportions of core, standard and remainder in extracts

	Sp1	**Sp2**	**Sp3**	**Sp4**	**Sp5**	**Sp6**
Core %	7.68	11.70	13.22	10.90	8.10	14.45
Standard %	67.61	69.89	68.39	72.88	75.42	63.98
Remainder %	24.71	18.40	18.39	16.21	16.48	21.56
Total phonemes	599	940	522	697	358	422

It emerges that the speaker who uses the most core features (judged by Jenkins to be positive for intelligibility) is L1 Italian Sp3 and the one

[6] Sonia D'Apolito and Bianca Sisinni of CRIL, Università del Salento, Italy.

who uses most features that are neither core nor standard (remainder), and thus negative for intelligibility, is the L1 German, Sp1.

Unsurprisingly, but confirming the accuracy of the phonetic transcripts (produced with no background information about the speakers), the speaker that uses the most features closest to NS standard norms is Sp5, the NS. The figure for this may seem low (only 75.42%). However, a margin of error has to be tolerated given the imperfect nature of the recordings with which the transcribers had to work. Furthermore, the standard pronunciation given in the *COED* only gives the pronunciation of words in isolation and not in the context of utterances. Consequently, many natural features of NS performance, such as vowel reduction and catenation are classed, in this analysis, as remainder and not standard or core.

To gauge which of these parameters carries most weight in terms of intelligibility, a difficult problem when one considers the variety and complexity of the different aspects of performance involved, one may adopt the simple device of comparing the rankings of the various speakers as regards each phenomenon with the rankings for intelligibility taken from Question 5 of the questionnaire (i.e.: "This speaker's accent is easy for me to understand"). This, we do in Table 5:

TABLE 5
Speaker rankings according to proportions of core, remainder and standard

	Remainder (inverted)			Core		Standard		Intelligibility (QR*)
	Rank	RCS*		Rank	RCS*	Rank	RCS*	
–	Sp4	4	+	Sp6	1	Sp5	0	Sp2
	Sp5	1		Sp3	2	Sp4	3	Sp4
↕	Sp3	3	↕	Sp2	3	Sp2	3	Sp3
	Sp2	2		Sp4	3	Sp3	4	Sp1
+	Sp6	3	–	Sp5	4	Sp1	4	Sp6
	Sp1	4		Sp1	3	Sp6	4	Sp5
	Tot.	17		Tot.	16	Tot.	18	

*Key: RCS = ranking comparison score. QR = questionnaire respondents

To ascertain which of the three features (remainder, core and standard) has most effect on intelligibility, a straight comparison can

be made between each ranking and that for intelligibility; the one that proves most similar being considered most relevant.

The ranks may be compared by assigning a value to each ordering on the basis of how far it matches that of intelligibility. For this purpose, we propose a new analytical tool which we have dubbed a *ranking comparison score* (RCS). This can be calculated by assigning one point for each speaker placed above or below it in the rank, provided that they also do so in the ranking for intelligibility from the questionnaire. For example, Sp6 comes in fifth place for remainder. However, of the four speakers that come above it in intelligibility (Sp1, Sp2, Sp3 and Sp4), only three do in remainder (Sp2, Sp3 and Sp4) and there is no correspondence with those that come below (Sp1 v. Sp5). It is thus assigned three points.

By looking at the totals for the different RCSs (bottom row), it can be ascertained that the, marginally, most relevant is the presence of standard features, and the least, core features, with absence of either standard or core features (remainder) coming in between. However the figures for all three parameters are too close to conclude that Jenkins' LFC hypothesis is incorrect, partly because it must be underlined that we have measured core minus those standard features that Jenkins includes in the LFC.

Indeed, the proximity of values of the various parameters points rather to the fact that, as Jenkins maintains, the adherence to a standard modelled on a NS is not such a determining factor in intelligibility, despite its coming top according to our RCS. This is because if it were true that standard pronunciation constituted the model by itself then the differences in RCS would be much larger than found here.

The fact that the figures for the three parameters are so close (indeed the RCS for core, the lowest of the three, is only two below that of standard, the highest) is still a remarkable result given that the idea of NS pronunciation providing the only model is still engrained in many peoples' attitudes. However, Table 5 does not account completely for the respondents' ranking of the extracts according to intelligibility because the totals for divergence are all relatively low: between 16 and 18 out of a maximum of 30 (i.e. 53.33% to 60% correspondence by our measure). This implies that other factors must be taken into account beyond pure phonetics and the individual phonemes produced by the speaker in question.

Intonation and delivery is an obvious candidate for a feature having an effect on intelligibility because things such as fluency (i.e. speed of delivery, pausing and hesitations – see Luoma 2004) are well-known to be influential.

5. Delivery and its observable effects on intelligibility

Fluency in particular is a vague concept (see Sections 1 and 2) and, likewise, delivery constitutes a variety of different factors, many of which also associated with fluency. To investigate this area, we choose, as an objective means of measure, three easily observable phenomena: speed of delivery (average words per second – excluding unfilled pauses); average number of discernible words between pauses[7] and percentage of pauses in extract. Taken together, these give an indication not only of the raw speed of the delivery, so to speak, but also how many words are produced between pauses and how long such pauses are. Below in Table 6, we report the figures for each of these criteria for every extract.

TABLE 6
Measurable factors contributing to fluency in extracts

	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6
AWBP*	1.06	2.38	2.33	1.54	3.96	1.89
DS*	1.50	2.36	2.06	1.79	1.65	1.25
Pausing*	5.23	20.22	13.14	26.98	29.37	18.13

*Key: DS = delivery speed (average words per second, excl. filled pauses); AWBP = average number of discernible words between pauses; Pausing = seconds

Table 6 shows how Sp5 (NS) not only produces the longest total pauses but also the highest number of words between pauses. By contrast, pauses make up less of the L1 Polish Sp2's turn but she produces fewer words between pauses. These two performances may thus be characterised as, on the one hand (Sp5), rather erratic in rhythm and, on the other (Sp2), more regular and evenly paced. Such factors would seem intuitively to have obvious implications for intelligibility.

[7] We adopt this criteria to avoid the more problematic and less easily measurable concepts of intonation, tone or breath group etc. – see Cruttenden (1986 35-45).

Such comments aside, the challenge lies in discovering which precise combination of factors (remainder, core, standard, average discernible words between pauses, delivery speed, or total pauses) contributes most to intelligibility. In particular, within these combinations, which features have positive and which have negative effects.

To do this, we first have to devise a means of comparing the figures for the numbers of core, standard and remainder features with those for delivery speed, average number of discernible words between pauses and pausing. Such a thing cannot be done directly as the nature of the data itself is diverse and is measured in different ways (quantities, percentages, times). Our rank comparison score again proves useful here. In Table 7 below, we summarize the data necessary for its calculation.

TABLE 7
Summary of data for core, standard and remainder, AWBP, DS and Pausing in extracts

	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6	
Tot. phonemes*	708	973	549	723	378	458	Tot.
Core	46	110	69	76	29	61	391
Standard	405	657	357	508	270	270	2390
Remainder	148	173	96	113	59	91	680
Core 2*	**11.76**	**28.13**	**17.65**	**19.44**	**7.42**	**15.60**	
Standard 2*	**16.95**	**27.49**	**14.94**	**21.26**	**11.30**	**11.30**	
Remainder 2*	**21.76**	**25.44**	**14.12**	**16.62**	**8.68**	**13.38**	Tot.
AWBP	1.06	2.38	2.33	1.54	3.96	1.89	13.15
DS	1.50	2.36	2.06	1.79	1.65	1.25	10.61
Pausing	5.23	20.22	13.14	26.98	29.37	18.13	113.07
AWBP 2*	**8.03**	**18.11**	**17.70**	**11.68**	**30.10**	**14.39**	
DS 2*	**14.14**	**22.24**	**19.42**	**16.87**	**15.55**	**11.78**	
Pausing 2*	**4.63**	**17.88**	**11.62**	**23.86**	**25.98**	**16.03**	

*Key: Tot. phonemes = total identifiable phonemes including filled pauses; Core, Standard Remainder. 2 = Core / Standard / Remainder figure expressed as a percentage of total number of phonemes of that kind in all six extracts (see total column on right of Core Standard, Remainder); AWBP / DS / Pausing. 2 = AWBP / DS / Pausing figure expressed as a percentage of total number of features of that kind in all six extracts (see total column on right of AWBP, DS, Pausing).

To render the data comparable both horizontally (across the category) and vertically to other features within the same extract, we re-express all raw figures as percentages of the total of all six extracts for that category (in bold). It is the latter figures, which we go on to elaborate in Table 8, where we examine the data for eight combinations of features, for reasons of space, discounting the possibility of all items having a negative effect or of core and standard being negative or of remainder being positive. Once these scores have been calculated, rankings can again be made for each combination of features (Table 9), assigning a positive or negative value to each feature in bold in Table 7, according to whether the feature is considered as beneficial or detrimental to intelligibility (Table 8 below). As on Table 5, the rankings arrived at can be compared with that for intelligibility from the questionnaire (Table 9); the premise being that the highest score (out of 30) for the ranking (right-hand column) will indicate which precise combination accounts best for the original scores awarded by the respondents and thus constitutes the combination of factors which carry most weight as regards intelligibility.

TABLE 8
Scores for extracts according to different combinations (positive and negative)

	Remainder	**Core**	**Standard**	**AWBP**	**DS**	**Pausing**	**Sp1**	**Sp2**	**Sp3**	**Sp4**	**Sp5**	**Sp6**
1	—	+	+	+	+	+	77.3	139.3	95.4	109.7	99	82.5
2	—	+	+	+	+	—	-77.3	-139.3	-95.4	-109.7	-99	-82.5
3	—	+	+	+	—	—	33.7	88.4	67.2	76.5	81.7	55.7
4	—	+	+	—	—	—	24.5	52.7	44	28.8	29.7	23.6
5	—	+	+	—	—	+	-3.8	8.2	5.13	-5	-1.4	0.1
6	—	+	+	—	+	+	-5.7	-5.8	-10.8	-11.5	-46	-16.9
7	—	+	+	—	+	—	-10.6	7.7	-7	19.4	-9.6	3.4
8	—	+	+	+	—	+	17.7	52.2	31.8	53.1	21.5	26.9

TABLE 9
Ranking comparison scores and totals of same for various combinations of features affecting intelligibility

	Ranking ▶						1	2	3	4	5	6	TOTAL RCS ▼
	Intelligibility (QR) ▶						Sp2	Sp4	Sp3	Sp1	Sp6	Sp5	
1	—	+	+	+	+	+	Sp6	Sp1	Sp4	Sp3	Sp2	Sp5	
	RCS ▶						1	1	2	2	1	5	12
2	—	+	+	+	+	—	Sp5	Sp1	Sp2	Sp4	Sp3	Sp 6	
	RCS ▶						0	1	3	3	3	4	14
3	—	+	+	+	—	—	Sp5	Sp1	Sp2	Sp6	Sp4	Sp3	
	RCS ▶						0	1	3	2	2	2	10
4	—	+	+	—	—	—	Sp1	Sp2	Sp3	Sp4	Sp6	Sp5	
	RCS ▶						2	4	3	3	5	5	22
5	—	+	+	—	—	+	Sp6	Sp2	Sp4	Sp1	Sp5	Sp3	
	RCS ▶						1	4	4	3	4	2	18
6	—	+	+	—	+	+	Sp6	Sp2	Sp3	Sp1	Sp5	Sp4	
	RCS ▶						1	4	3	3	4	1	16
7	—	+	+	—	+	—	Sp2	Sp1	Sp3	Sp4	Sp6	Sp5	
	RCS ▶						5	3	3	3	5	5	24
8	—	+	+	+	—	+	Sp6	Sp3	Sp5	Sp2	Sp1	Sp4	
	RCS ▶						1	2	2	2	2	1	10

From Tables 8 and 9, it emerges that the ranking of speakers that most closely matches that given by the respondents is number 7, where high values in core, standard and speed of delivery have a positive effect on intelligibility and where high values in remainder, average number of discernible words between pauses and pausing have a negative one. This combination has a score of 24/30 (or 80%).

The difference between the ranking for the highest combination of features (number 7): other –,core+, standard+, average number of discernible words between pauses –, delivery speed +, pausing –) and that of the questionnaire for intelligibility lies only in inversion of the position of Sp 1 and Sp 4, who, not coincidentally perhaps, share the same L1 (German).

The spread of RCSs between the various combinations shown on Table 9, from 10 to 24, and the differences in the profile of each combination allows one to see which specific features seem to

carry more weight regarding intelligibility. Combination 4 has the second highest score (22), the two point difference between this and combination 7 being attributable to the fact that, in the latter, delivery speed is classed as a positive, rather than a negative, feature. By contrast, combination 3 with the joint lowest score (10) (with combination 8) differs from combination 7 in only two (average number of discernible words between pauses and delivery speed). It would appear that average number of discernible words between pauses carries greater weight than delivery speed because this time the difference in RCS between combinations 7 and 3 is 14 points. Indeed, by examining Table 9, it can be seen that in all combinations (1, 2, 3 and 8) where average number of discernible words between pauses is classed as a positive feature (i.e. where more words are grouped together between clauses), the scores are notably lower (mean 11.5) than those (mean 20) for the combinations where AWBP is negative (4, 5, 6, and 7).

To further investigate relative weighting of features , we can calculate the mean scores for the various combinations in which a given feature of delivery is classed as either positive or negative and, from these, calculate the differences in those same means for each feature, as shown in Table 10:

TABLE 10
Mean rank comparison scores for combinations in Table 9 classing a given feature as positive or negative

	AWBP		DS		Pausing	
	+	-	+	-	+	-
Mean	11.5	20	16.5	15	14	17.5
Diff	8.5		1.5		3.5	

As can be seen, for both delivery speed (DS) and pausing, the difference in means is relatively small (1.5 and 3.5) while, for average number of discernible words between pauses (AWBP), it is notably larger (8.5). This indicates that the three features exert an unequal amount of influence on intelligibility; the effect on intelligibility of fewer or more words between pauses is greater than the effect of a lesser or greater delivery speed or of more or less pausing.

6. Conclusions

This research has aimed to identify phonetic factors that account for intelligibility in an ELF context. There has not been space to also examine the pragmatics of meaning negotiation which also plays a part, albeit secondary, to the production of phonological segments and supra-segmental features in themselves. It has been shown that intelligibility depends on a syndrome of factors, among which the LFC, together with an abstract NS standard as described in an authoritative reference work, namely the COED. It has also been established that aspects of delivery also play an important part (compare the RCSs in Tables 5 and 9). The exact configuration of these factors necessary for intelligibility seems to be one where presence of core, standard features as well as faster rather than slower speed of delivery are demonstratively positive features and, conversely, the presence of features which are neither standard nor core (remainder), a high average number of discernible words between pauses and a high level of pausing are negative ones. Such a profile of features fits Sp2 best of all, and Sp5, the NS, least well. Such a finding confirms Walker (2010), who in his guide to teaching ELF pronunciation, argues that teaching learners to produce supra-segmental features like weak forms and connected speech, typical of rapid NS speech, is counterproductive in ELF contexts.

Sp2 is the longest in terms of number of words and duration but a distant second to Sp5 (NS) in terms of clustering of words together into groups between pauses; and a close second to Sp5 as regards speed of delivery. Overall, delivery is more measured with a greater balance between production of sounds and pausing. Sp2, therefore, by her general speed of delivery perhaps gives the impression of NS competence but without the irregular delivery. By contrast, Sp5 (NS) is the shortest turn in terms of time and number of identifiable words. But he produces the highest number of discernible words between pauses and displays the highest speed of delivery (see Table 6), when pauses are ignored. It is also the extract where the proportion of pauses is the highest. The overall effect on the NNS respondent is thus perhaps one of an irregular distribution of sound with long pauses punctuated by rapid bursts of sounds in a way reminiscent of a concertina.

Such factors were obviously not taken into account by speaking-test experts. Either some other factor (e.g. tone / pitch) is involved or the

issue is subjective (attitudinal) rather than objective. There is perhaps some L1 bias (see Jenkins 2007: 95-100) in speaking-test experts who are all NS British English and have familiarity with the variety of English used by Sp5 (standard British English with traces of its London variety). Further research is required to clarify this point, especially using varieties of NS that a control group may not be so familiar with.

One notable general finding from the questionnaire is that respondents are not very proficient at identifying the NS used in this experiment, whereas the speaking-test experts gave his phonological control top marks according to the CEFR descriptors, which refer directly to NSs (see note). Such a failure casts doubt on the premise that such learners see a form of NS pronunciation as a model. Rather, respondents tended to mark as NS, a speaker like Sp2, who revealingly, is judged also to be the most intelligible (see Table 3).

One can conclude from this realisation that, in practice, 'standard' NS English does not serve as a model for NNS pronunciation and delivery in English (see Jenkins 2007: 187). It is thus evident that the advent of ELF as a separate variation (or rather subset of variations – see Guido 2008: 21-27) is well underway. It also argues for consideration of some elements of delivery alongside Jenkins' original LFC as features that should be taught or encouraged in contexts of ELF spoken communication.

Nevertheless, the differences in perception of the NS in this study highlight a gap that may still exist between the concerns of educators and those of learners. More ELF-oriented descriptors for use in assessment would no doubt assist in breaching it. Such a thing may involve, at the level of the CEFR and other such documents, distinguishing between languages, primarily English, that are used mainly in lingua franca communication and others used for communication principally with NSs. This is in line with the recommendation of Seidlhofer that there be "[f]irst and foremost, a re-orientation of 'English' away from the fascination with ENL [English as a native language] and towards the cross-cultural role of EIL [English as an International Language]". (Seidlhofer 2003: 22)

The concept of LFC and related concerns of delivery, in particular average number of discernible words between pauses (as identified by Table 10), though in need (as all such concepts are) of continual research, is still a relevant conceptual tool that describes how learners observably perceive intelligibility. As a consequence, and in line with

Jenkins' original motive in identifying it in the first place, it needs more promotion not just among learners but also among educators, and NSs in general, using English in lingua franca contexts.

References

CHRISTIANSEN, THOMAS, 2011a, "Fluency and Pronunciation in the Assessment of Grammatical Accuracy in Spoken Production: An Empirical Study", *Lingue e Linguaggi* 6, pp. 7-20

CHRISTIANSEN, THOMAS, 2011b, "Pronunciation and Intelligibility in ELF: A Case Study", *Lingue e Linguaggi* 6, pp. 21-38

COUNCIL OF EUROPE, 2001, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, C.U.P., Cambridge.

CRUTTENDEN, ALAN, (1986) *Intonation*, C.U.P., Cambridge.

GUIDO, MARIA GRAZIA, 2008, *English as a Lingua Franca in Cross-cultural Immigration Domains*, Peter Lang, Bern.

JENKINS, JENNIFER, 1998, "Which Pronunciation Norms and Models for English as an International Language?" *ELT Journal*, 52/2, pp.119-26.

JENKINS, JENNIFER, 2000, The *Phonology of English as an International Language*, O.U.P., Oxford.

JENKINS, JENNIFER, 2002, "A Sociolinguistically Based, Empirically Researched Pronunciation Syllabus for English as an International Language", *Applied Linguistics* 23/I, pp. 83-103.

JENKINS, JENNIFER, 2007, *English as a Lingua Franca: Attitude and Identity*, O.U.P, Oxford.

JENNER, BRYAN, 1997, "The English Voice", in A. Brown (ed.), *Approaches to Pronunciation Teaching*, Prentice Hall International, Hemel Hempstead, pp. 38-46.

LENNON, PAUL, 1990, "Investigating Fluency in EFL: A qualitative Approach", *Language Learning, 40* (3), pp. 387-417.

LUOMA, SARI, 2004, *Assessing speaking*, C.U.P., Cambridge.

MUNRO, MURRAY J. and DERWING, TRACY M., 1995, "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners.", *Language Learning 45*(1), pp. 73-97.

SEIDLHOFER, BARBARA, 2003, *A Concept of International English and Related Issues: from 'Real English' to 'Realistic English'?* Council of Europe, Strasbourg.

SEIDLHOFER, BARBARA, 2011, *Understanding English as a Lingua Franca: A Complete Introduction to the Theoretical Nature and Practical Implications of English used as a Lingua Franca.* O.U.P., Oxford.

SOANES, CATHERINE and STEVENSON, ANGUS (eds), 2004, *Concise Oxford English Dictionary*, O.U.P., Oxford 11th edition.

WALKER, ROBIN, 2010, *Teaching the Pronunciation of ELF*, O.U.P., Oxford.