



Perceptions of Fluency

Author's Name: Cecilia Dore

University of Reading

British Council ELT Master's Dissertation Awards: Winner

Abstract

It is clear from the extensive literature on fluency in second language speech that the construct is very difficult to define, and yet given that it is one of the main components assessed in speaking tests, consensus over what it involves is important. This exploratory study investigates what native-speaker English language teachers based in Italy and the UK perceive to be the key factors that influence them when they rate learners for fluency. The aim was to determine to what extent they agree and whether opinions on the question are related to training and teaching experience.

A quantitative and qualitative online questionnaire was distributed and forty-eight responses were received. Analysis of the data revealed that the participants were largely in agreement over a cluster of variables that indicate cognitive processing e.g. effortlessness and automaticity, and the surface evidence of this underlying machinery e.g. pauses, hesitations, and reformulation. Reservations were expressed regarding rate of speech. A striking amount of importance was attributed to coherence, which could also be seen as evidence of cognitive efficiency, and to length of run, referred to as “chunking” by many respondents in the qualitative data. Finally, rhythm and intonation were perceived to be highly influential. The findings have pedagogical implications, especially if verified by further research.

The data revealed that the UK-based participants generally have fewer years of teaching experience but have had more specific training than their Italy-based colleagues. The UK group were seen to be more in line with official speaking test descriptors, while the Italy-based participants showed a broader interpretation of the term fluency.

Table of Contents

List of tables and figures	v
List of appendices	vi
List of abbreviations	vii
1. Introduction	1
2. Literature Review	
2.1. Introduction	3
2.2. Definitions of fluency: broad versus narrow	3
2.3. A cognitive perspective: Segalowitz's three-way distinction	5
2.3.1. Cognitive fluency and Levelt's "blue-print" of speech production	6
2.3.2. Utterance fluency	8
2.3.3. Perceived fluency	10
2.4. Linguistic and phonological variables	12
2.4.1. Naturalness: communicative competence, formulaic sequences, fillers	12
2.4.2. Intonation as an indicator of fluency	13
2.5. Speaking tests and their descriptors	14
2.6. Conclusion and research questions	15
3. Methodology	
3.1. Introduction	16
3.2. Learner participants	16
3.3. The speaking task	17
3.4. Sampling strategy – the teacher participants	18
3.5. TP profile	19
3.6. Instruments	21

3.7. Questionnaire design	22
3.8. Grouping the variables	23
3.9. Data analysis	26
4. Analysis	
4.1. Introduction	27
4.2. Global rating of the three speakers	27
4.3. Research question 1: What aspects of fluency do teachers pay to when rating non-native speaker fluency?	28
4.3.1. Quantitative data	28
4.3.2. Qualitative data	30
4.3.2.1. Fluency Variables	31
4.3.2.1.1. Utterance fluency	31
4.3.2.1.2. Cognitive fluency	32
4.3.2.1.3. Perceived fluency	32
4.3.2.2. Complexity Variables	33
4.3.2.3. Phonology Variables	33
4.3.2.4. Global Aspects	34
4.3.3. The causes of disfluency	34
4.3.3.1. Cognitive processing	35
4.3.3.2. Linguistic complexity	36
4.3.3.3. Personality and affective factors	36
4.3.4. The subjective element	36
4.3.4.1. Language specialist versus lay person	37

4.3.4.2. Listener preferences	37
4.4. Research Question 2: Are there differences between the two groups of teachers in terms of perceptions?	38
4.5. Research Question 3: Are teachers' perceptions correlated with training and years of experience?	40
4.6. Conclusion	42
5. Discussion	
5.1. Introduction	43
5.2. Key aspects of fluency	43
5.3. Differences between the two groups of TPs: perceptions, training and experience	45
5.4. Towards a definition	46
5.5. Implications for pedagogy	48
5.6. Limitations of the study	49
6. Conclusion	50
7. References	52
8. Appendices	56

List of tables and figures

List of figures

Figure 1: Mean age, teaching experience and years in the same institution	20
Figure 2: Highest academic and teaching qualifications	21

List of tables

Table 1: Grouping of the 20 variables	24
Table 2: Source of data and method of analysis to answer the research questions	26
Table 3: Best speaker according to TPs in each country	27
Table 4: Global ratings for the speakers in each country	28
Table 5: Descriptive statistic for the 20 variables	29
Table 6: Percentage of open-ended answers that identify variables in each group	31
Table 7: Mann-Whitney U-tests to measure differences in perceptions in Italy and UK	39
Table 8: Number of mentions in qualitative data of significantly different variables according to country	40
Table 9: Significant correlations between Highest Academic Qualification, Highest Teaching Qualification, Years of Experience, and variables	41

List of appendices

Appendix 1: Levelt's 1999 blueprint of speech production	56
Appendix 2: Kormos' (2006) overview of measures of fluency	57
Appendix 3: IELTS and CEFR speaking test descriptors	58
Appendix 4: Picture story	59
Appendix 5: Information sheet for learners	
Appendix 6: Consent form for learners	
Appendix 7: Further excerpts from qualitative data	60
Appendix 8: Printed version of online questionnaire	64

CD: 3 recordings of monologic learner speech

List of abbreviations

CAF: Complexity, Accuracy and Fluency

CELTA: Certificate in English Language Teaching to Adults

CEFR: Common European Framework of References for Languages

DELTA: Diploma in English Language Teaching to Adults

GA: Global aspects

IELTS: International English Language Testing System

L1: First language

L2: Second language

SLA: Second language acquisition

TP: Teacher participant

SPSS: Statistical Package for the Social Sciences

Perceptions of Fluency

1. Introduction

Fluency in a second language is something all learners aspire to (Kormos & Dénes, 2004), but what exactly does it mean to be fluent? This study explores some of the many definitions of this elusive term when applied to second language (L2) speech and investigates which components of speech teachers consider when rating learner fluency. It is a topic of considerable interest and importance to language learners, language teachers, language performance evaluators and second language acquisition (SLA) researchers alike, and it has produced a wealth of fascinating literature. Studies have approached fluency from various angles – linguistic, psycholinguistic and sociolinguistic – but, disappointingly, the main problem still lies in producing a clear and universally accepted definition of the construct (Chambers, 1997).

Any informal survey of L2 learners and language teachers' opinions on what fluency is will demonstrate how widely ideas on the topic differ. This first came to the author's attention when she attended a conference talk on L2 speech fluency and was surprised to hear the construct defined exclusively in terms of temporal variables such as rate of speech and number and length of pauses. The immediate reaction was to wonder how many people might be considered disfluent in their first language (L1), let alone when speaking a second or third language. This alarming thought was followed by the concern that she and her colleagues, teachers of English as an L2 in higher education, might have been assessing learner fluency wrongly over the past two decades. The realisation that professional teachers at university level all have an intuitive notion of what they are assessing but that these notions differ led to an interest in exploring what SLA researchers have identified as the components of fluency, and comparing these with the ideas held by the practitioners in the classroom.

Between these two levels of professionals there are the examiners of language tests, who follow the guidelines of written descriptors when they assess learners. These descriptors or benchmarks, informed by the more theoretical considerations of researchers, describe the test taker's performance at different levels (Fulcher, 2003). It

is logical that they in turn should inform teachers who prepare their learners for tests and thus learners themselves. Ideas about what constitutes any construct, whether it be fluency, accuracy or lexical range, should be consistent between these three categories of professionals.

The theoretical grounds for this study, therefore, are that only when there is a clear definition of what fluency consists of will it be possible to measure and operationalise it and advise pedagogy as to how fluency in second language learners can be enhanced. On the practical level, clear criteria will help learners, teachers and test raters. Learners need to know how to assess their own progress as part of the process of becoming autonomous (Préfontaine, 2013). Teachers need to be able to assess their learners and know how they can refine their methodology to help them. Test raters need to know which criteria to include and which to exclude when assessing test takers, as their judgments affect people's futures. According to Kormos and Dénes (2004), few studies on perceptions of fluency have investigated the role played by a combination of linguistic, temporal and phonological variables. The present study attempts to address this gap.

In order to establish what native speaker university teachers of English take into consideration when rating L2 speech fluency, a quantitative and qualitative questionnaire was adopted. The category of university teachers was chosen for the respondents as members of this profession frequently have to assess learners for fluency (Kormos and Dénes, 2004). The survey was carried out in two countries, Italy and the UK, to explore the extent to which notions are local or more widespread. The aim of this exploratory study, therefore, is to identify the most commonly perceived components of fluency according to university teachers in two countries, and to propose what should be included in a definition of the term on the basis of the findings. The second chapter will review the literature that helped shape the study and its research questions. Particular emphasis will be given to Segalowitz's (2010) cognitive approach. A detailed description of how the study was conducted will follow in the methodology chapter. The fourth chapter will analyse the findings, and the fifth will discuss them and their implications for language learning, teaching and assessment. In the conclusion, recommendations will be made for future research.

2. Literature review

2.1. Introduction

Speaking a language fluently is the ultimate goal for most learners (Kormos & Dénes, 2004). However, there is an ongoing lack of consensus regarding what the term *fluency* means (Chambers, 1997), and consequently it is difficult for teachers and learners to identify how to enhance it. There seems to be the implicit belief that it cannot be taught but that it will develop naturally (Chambers, 1997). Most of the research that has been carried out has focused on temporal features such as speech rate and the number and length of pauses (Lennon, 1990; Tavakoli, 2011; Bosker, De Jong, Pinget, Quené & Sanders, 2012), the advantage of these being that they can be measured and so provide empirical evidence. And yet intuition suggests that this is not the whole story, and studies that investigate raters' perceptions of fluency (Freed, 2000; Kormos & Dénes, 2004), have shown this to be the case.

The theoretical framework underlying the present study is Levelt's (1999) model of speech production and Segalowitz's (2010) cognitive perspective on fluency, in particular on perceived fluency. This chapter will provide some key definitions of fluency before describing Segalowitz's contribution to the debate and Levelt's model of speech production. Some of the many studies that have aimed to define the construct will be referred to, and there will be a section dedicated to specific linguistic and phonological variables that have been identified as indicators of fluency. Finally, there will be a brief discussion of how speaking test descriptors interpret fluency, and the research questions of the present study will be introduced.

2.2. Definitions of fluency: broad versus narrow

A logical first step in attempting a definition of *fluency* is to examine its etymology. The words *fluency* and *fluent* come from the Latin word *fluens*, and have been used as English words for approximately four centuries (Koponen & Riggensbach, 2000). Samuel Johnson's nineteenth century definitions of *fluent* as *liquid*, *flowing*, and *fluency* as *the quality of flowing*, *smoothness*, shed light on what the words mean when applied

to speech. If someone is said to speak a language fluently, it is understood that they speak it with a smooth and easy flow. As Kaponen and Riggenbach (2000) point out, other languages have equivalent words to describe spoken language. In German *flüssig* and *fliessend*, in French *courent*, and *correntemente* in Italian, to give some of the many examples, all mean *flowingly*. In the lay sense, fluency in a language has come to mean global proficiency or mastery.

In one of the first studies investigating fluency, Fillmore (1979) wrote a seminal paper on the language abilities attributed to fluency in L1 speech. He identifies four main dimensions: firstly, the ability to talk at length without pauses and hesitations, so the ability to keep the ball rolling; secondly, the ability to speak coherently and in a “semantically dense manner”, so here there is a focus on quality rather than quantity; thirdly, the ability to use language appropriate to the context, here the concern being related to pragmatic skills; and fourthly, the ability to use language creatively and imaginatively. In the same paper, Fillmore points out that individual L1 speakers vary considerably in their speaking style, and this will be discussed later in section 2.3.2.

Fluency in native speaker speech, therefore, is somewhat complex, and second language learners are usually only too aware of the deficiencies in their second language speech. However, Fillmore’s four-point description of fluent speech reflects the view of fluency as overall competence, or global proficiency. Lennon (1990) makes a distinction between what he terms a “broad” sense and a “narrow” sense. The “broad” sense or non-technical meaning is what he refers to as a layman concept, and may be defined as “the highest point on a scale that measures the spoken command of a foreign language” (p.389). The “narrow” sense of fluency, on the other hand, is more the realm of the psycholinguist or the language teacher, that is, the specialist. In Lennon’s words, fluent delivery in performance is purely a performance phenomenon, and “directs listener attention away from deficiencies in other areas: phonological, grammatical, syntactic, discursive, lexical.” He argues that “fluency is an impression on the listener’s part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently” (1990, p.391). His analysis of twelve variables in a study of German speakers in England investigates a reason for this “impression” and identifies as objective indicators of fluency faster speech rate, fewer filled pauses per t-unit, and fewer t-units followed by a pause.

Numerous studies since Lennon's have continued this investigation of fluency in its narrow sense, that is, as one component of oral proficiency, and the search for its most reliable indicators (see section 2.3.2.). Further defined as "speedy and smooth delivery of speech without (filled) pauses, repetitions, and repairs" (De Jong, Groenhout, Hulstijn & Schoonen, 2015, p.4), and used only to describe the speech of L2 speakers, fluency is a construct, and constructs must be defined in a way that they can be observed and measured and distinguished from other constructs in order for them to be "operational" (Fulcher, 2003). In language tests e.g. the Common European Framework of References for Languages (CEFR), fluency is one of several component constructs in the evaluation of speaking. An accurate understanding of what it consists of is therefore crucial.

2.3. A cognitive perspective: Segalowitz's three-way distinction

In his search for a framework for discussing L2 fluency, Segalowitz (2010) proposes a cognitive perspective on the question. He casts the following anchor question:

What features of L2 oral performance serve as reliable indicators of how efficiently the speaker is able to mobilise and temporally integrate, in a nearly simultaneous way, the underlying processes of planning and assembling an utterance in order to perform a communicatively acceptable speech act?"
(p.47)

The question refers both to the underlying cognitive processing producing the utterance and to the characteristics or features of the utterance itself, and leads Segalowitz to distinguish between three different senses of the term fluency. In brief:

- *Cognitive fluency* refers to the underlying cognitive machinery that transforms the intended message into spoken language. In terms of the original metaphor of flow, if this machinery is efficient, the process will be smooth and effortless.
- *Utterance fluency* refers to the surface properties of speech e.g. rate of speech, pauses, hesitations, repetition, reformulation. They provide evidence of the underlying cognitive fluency.

- *Perceived fluency* focuses on the listeners' impressions and the inferences they make about a speaker's cognitive fluency based on their perceptions of utterance fluency.

This three-way distinction and how the three are interrelated is the theoretical basis of the present study, and will now be examined in more detail.

2.3.1. Cognitive fluency and Levelt's "blue-print" of speech production

In a more recent study, Lennon (2000) synthesised earlier definitions of fluency into the following working definition: "the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing" (p. 26). This is an exceptionally complicated mechanism. In Kormos' (2006) words: "One of the most complex automatic human activities is linguistically encoding what one wants to say in his or her mother tongue" (p. 38).

Levelt's model of speech production, first proposed in 1989, describes the cognitive processes involved in monolingual speech. It is the model most frequently cited in studies of speech production (Kormos, 2006; Segalowitz, 2010). In simple terms, there are three stages: first speech is conceptualised, then it is formulated or encoded, and finally it is articulated. The speaker draws on three different stores of knowledge: knowledge of the external and internal world, the *mental lexicon*, which contains lemmas and morpho-phonological codes, and the *syllabary* containing phonological information.

Conceptualisation is a pre-verbal stage where both macro-planning and micro-planning take place, macro-planning being the elaboration of the communicative intention, and micro-planning involving the selection of elements that convey perspective, thematic roles and new and old information. Formulation, the second stage, requires the retrieval of information from the mental lexicon. Lemma selection activates the lemma's syntax, and phrases are generated. The result of this is the *surface structure*, which is then processed further as morpho-phonological encoding takes place. Phonetic encoding

draws from the syllabary, and finally the resulting *articulatory score* is converted into overt speech by the articulator (Dörnyei & Kormos, 1998).

These three stages of speech production are mirrored in the process of speech perception, whether the speech is another speaker's utterances or one's own (self-monitoring). First it is perceived, then it is decoded in the speech comprehension system, or *parser*, and finally it is interpreted by the conceptualiser. Self-monitoring also occurs at three stages, according to Levelt's model: during conceptualisation and formulation, and after articulation, and any errors perceived trigger off signals that may lead to repair. Repair is typical of spontaneous spoken language and includes repetition, reformulation, and self-correcting. When speech involves interaction, it is further complicated (Thornbury, 2005), as each utterance is a reaction to the interlocutor's previous one. Moreover, the articulation of one utterance and the planning of the next may overlap. For speech to be fluent, considerable automaticity is required to carry out these parallel processes.

De Bot (1992) adapted Levelt's 1989 model in his discussion of L2 speech. He illustrated the reasons why this is often slower and more hesitant than L1 speech. Whereas processing by the formulator and the articulator is automatic and simultaneous in the speaker's L1, in L2 speech the speaker's attention is required for grammatical and phonological encoding and therefore part of the output can only be processed serially (Dörnyei & Kormos, 1998). With demands made on the L2 speaker's attentional resources, the working memory is not free to focus on conceptualising and formulating. Moreover, L2 speakers often lack the language they require to express their intended message, and this interrupts the cognitive flow. L1-L2 interference, an additional complication, is due to both languages sharing the same lexicon and a common syllabary, according to de Bot. In 1999 Levelt revised his "blue-print" to accommodate de Bot's theories (Appendix 1).

Schmidt (1992) claimed that fluency in speech production is "an automatic procedural skill" (p. 358), the word "skill" emphasising that it is a performance phenomenon rather than a question of knowledge. While grammar and lexis represent elements of linguistic knowledge, fluency is the speaker's ability to use his/her interlanguage to communicate in real time. As when learning any skill, first rules have to be learnt, and initially their application is conscious and requires attention. With practice, however, they are

consolidated and eventually produced automatically. One of the most frequently cited theories regarding how automatization in L2 speech develops, is Anderson's (1996) adaptive control of thought (ACT*). According to this, automaticity occurs when *declarative knowledge*, or explicit knowledge of the language system, is converted into *procedural knowledge*, or automatic processes which operate without conscious effort.

The view of fluency as a cognitive process would ring true for most people. Most of us are familiar with the feeling that our fluency, whether in our L1 or L2, is affected by our ability to access the appropriate language in real time, and that innumerable situational and affective factors can affect our cognitive processing. However, it is difficult to measure and therefore operationalise cognitive fluency. Researchers have devised ways of measuring *reaction time* (RT) in lexical decision tasks, and test whether comprehending word meaning is unstoppable, or ballistic, in reading tasks (Segalowitz, 2007), and the findings may predict speech fluency, but this is of limited help to raters and teachers on a practical level. Retrospective interviews, or stimulated recall, where learners recount their thoughts as they spoke after performing a speaking task, also provide interesting insights into cognitive fluency (Gass & Mackey, 2000; Kahng, 2014), and real-time magnetic resonance imaging (rtMRI), shows brain activity (Kormos, 2006). Nonetheless, cognitive fluency is not directly accessible to the naked ear, and this explains the appeal of considering surface features of speech the best indicators of fluency.

2.3.2. Utterance fluency

Segalowitz's second definition of fluency, utterance fluency, refers to the temporal properties of speech and repair. These, according to Segalowitz's vision, depend on the speaker's cognitive fluency. Tavakoli and Skehan (2005) differentiate between three categories of utterance indicators: speed fluency, that is, the density and speed at which speech is delivered; breakdown fluency, that is, the extent to which speech is interrupted by pauses; and repair fluency, for example the number of repetitions and corrections present in speech.

The advantage of utterance fluency is that the variables can be measured, for example by using the software Praat, and therefore ratings are more objective. A large number of different quantitative measures are proposed in the numerous studies that investigate utterance fluency, with different researchers advocating different aspects of production as the best indicators of fluency. Kormos (2006) gives an overview of the ten main measures (Appendix 2), which include mean length of runs, the mean length of pauses, the number of filled pauses per minute, phonation-time ratio (the amount of time spent speaking as a percentage of the total speech sample time), and the number of stressed words per total words. She states that most studies (Freed, 2000; Lennon, 1990) conclude that the best predictors are speech rate, that is, the number of syllables articulated per minute, and the mean length of runs, that is, the average number of syllables produced between pauses. However, there are many exceptions, and Tavakoli, Campbell and McCormack (2015) point out that there is a fair amount of overlap between the measures, some being “composite” e.g. speech rate, which combines pausing and speed aspects of fluency, others non-composite e.g. articulation rate, which considers speed only.

Kormos and Dénes (2004) conclude in their study on perceptions of fluency that pace, that is, the number of stressed words per minute, is one of the best indicators of fluency and has the added advantage of being easy to calculate. They report that findings on the frequency of filled and unfilled pauses and repair phenomena are generally contradictory and demonstrate in their own study that they do not influence perceptions of fluency. The relative unimportance of repair measures is confirmed by Bosker et al. (2012). In short, each study measures a different cocktail of variables using a variety of tools and different sample sizes, and this makes it difficult to draw clear conclusions.

Moreover, De Jong et al. (2015) question the extent to which such fluency measures are valid indicators of L2 proficiency. They echo Fillmore (1979) when they argue that fluency is partly determined by individual speaking style and personality, and may also be culture-based. They investigate fluency measures for Turkish and English learners of Dutch in both their L1 and their L2 and conclude that for most measures, especially syllable duration and filled pauses, the speakers’ performances correspond in both languages. They argue, therefore, that L1 fluency measures need to be taken into consideration when considering L2 fluency. Derwing, Munro, Thomson and Rossiter

(2009) also found a significant correlation between the L1 and L2 fluency behaviour of Slavic and Mandarin speakers of English.

This raises the important question of whether it is reasonable to judge L2 speakers as disfluent if it reflects their L1 speech. Regarding pausing, Tavakoli (2011) helps clarify the question by pointing out that it is where speakers pause that is significant, not the number or length of pauses. Whereas native speakers pause at the end of clauses, L2 speakers tend to pause mid-clause as a result of processing difficulties. A further consideration is whether pausing is due to linguistic planning or content planning (Fulcher, 2003). L1 speakers frequently pause to think about content when speaking spontaneously, and therefore this should not be interpreted as an indication of L2 disfluency. False starts and reformulation are also typical of many L1 speakers, as Conversation Analysis has shown (Schegloff, Jefferson & Sacks, 1977). Deciding which surface features of speech constitute fluency and disfluency, therefore, is not a straightforward matter.

2.3.3. Perceived fluency

Segalowitz's third view of fluency, perceived fluency, is the most complete of the three as it combines cognitive fluency and utterance fluency and adds the listener to the equation. This seems logical, given the difficulties in measuring fluency objectively. As previously illustrated, cognitive fluency is difficult to quantify and temporal and repair phenomena are not always reliable indicators and therefore it is inevitably the listener's impressions that count. Lennon (1990) claimed that "fluency reflects the speaker's ability to focus the listener's attention on his/her message by presenting a finished product, rather than inviting the listener to focus on the working of the production mechanisms" (pp. 391-392). In other words, if the speaker communicates his/her message effectively without pausing or hesitating unduly, the listener will infer that the speaker possesses cognitive fluency.

In light of the centrality of the listener and the subjective element in Lennon's definition of fluency, it seems appropriate to consider the identity of the listeners in past studies. Interestingly, according to Bosker et al. (2012), who use untrained raters in their

investigation into what influences perceptions of fluency, fluency judgements from untrained native-speaker raters are equivalent to those of expert raters. They refer to other studies (Derwing et al., 2009; Pinget, Bosker, Quené, & De Jong, 2014), where the same claims are made. Freed's (2000) six judges are also "educated native speakers" of the learners' L2 with no training or experience in language teaching, and their ratings correlated well with the expert raters' test scores. In fact, to the researcher's knowledge, very few studies on perceptions of fluency have used professional teachers as raters. Kormos and Dénes (2004), Wennerstrom (2000), and Préfontaine (2013) are exceptions, but Préfontaine's study is concerned with exploring learner self-assessment rather than investigating the components of fluency, and Wennerstrom's focuses exclusively on intonation. The use of experienced teachers as participants in the current study addresses this shortfall.

Past research into perceptions of fluency has considered a variety of variables. Whereas Bosker et al. (2012) restrict their raters to features of breakdown, speed and repair, Freed (2000) asks her judges to describe the basis on which they evaluate the speakers in their own words, and then select from a list the components they consider important in creating an impression of fluency. The components range from temporal, breakdown and repair variables to considerations such as idiomatic language, vocabulary, accent, grammar and interaction. Freed's findings confirm that the "popular notion of fluency" (p. 262) is subjective, that it includes elements of the narrow sense of the construct such as hesitation and repair phenomena, but is actually much broader. Four out of the six judges specified "better grammar and vocabulary" as major factors that influenced them, and half of them selected "better accent" as important. In short, the raters' perceptions of fluency in Freed's study were influenced by an eclectic mix of variables.

Kormos and Dénes (2004) state that there is a lack of studies that investigate the role of a combination of temporal, linguistic, phonological and interactional variables in perceptions of fluency, and set out to address this gap. They find that their professional teacher participants also consider fluency to be more than a question of temporal phenomena, with accuracy and lexical diversity emerging as important. Moreover, one of their main findings is that pace is a reliable indicator of fluency. The number of stressed words one can say in a minute is found to be more important in perceived fluency than a higher number of unstressed words. Stress, therefore, a component of

phonology, is seen to play a significant role. They conclude that a definition of fluency needs to include speed, pace, smoothness and accuracy.

2.4. Linguistic and phonological variables

This section will address further linguistic and phonological variables that have been identified as relevant to perceptions of fluency in past studies.

2.4.1. Naturalness: communicative competence, formulaic sequences, fillers

The fluency descriptor for the top band of CEFR states that the speaker can speak “with a natural colloquial flow.” This is reflected in Segalowitz’s (2010) section on “naturalness” (p.113), and the importance of *communicative competence*. First advocated by Hymes in 1967, this sociolinguistic approach to language involves knowing the appropriate register considering the social norms of the situation and the status of the participants (Segalowitz, 2010). Knowledge of discourse models is stored in Levelt’s *knowledge of the external and internal world* (Appendix 1) and is accessed during the pre-verbal macro and micro-planning stages of speech. Without sociolinguistic competence in the L2, Segalowitz claims, speakers will struggle as they search for the appropriate language, and this will cause dysfluency.

Naturalness in speech is also enhanced by knowledge of *formulaic sequences*. Ubiquitous in native-speaker speech, (Nattinger & DeCarrico, 1992; Wray, 2002), these are ready-made chunks of language that are either completely fixed, for example *last but not least*, or contain open slots which allow variation, for example *take (someone) an amount of time (to do something)*. In their groundbreaking paper on nativelike naturalness in speech, Pawley and Syder wrote that:

fluent and idiomatic control of a language rests to a considerable extent on knowledge of a body of “sentence stems” which are “institutionalized” or “lexicalized”.

(1983, p. 191)

Fluency, therefore, depends on the speaker having a store of these culturally determined fixed expressions. Pawley and Syder (1983) claim that they present a processing advantage over sequences of words generated creatively (non-formulaic sequences). According to this theory, they are committed to the long-term memory as single units and stored in the mental lexicon, from where they can be retrieved as “unanalysed wholes”, or with the minimum of encoding. This leaves the working memory free to focus on comprehending and conceptualising rather than on formulating.

Research shows that learners who incorporate lexical phrases or formulaic sequences into their speech are perceived as more fluent (Wood, 2009; Boers, Demecheleer, Eyckmans, Kappel & Stengers, 2006), not only because they are demonstrating a nativelike naturalness in their knowledge of these expressions, but also because they are pronounced as phonological units, without hesitations or pauses in the middle. Tavakoli (2011) demonstrates that L2 learners hardly ever pause in the middle of formulaic sequences.

Included in taxonomies of formulaic sequences are phatic expressions, conversational routines, idiomatic expressions, colloquialisms, and fillers. Fillers are words and expressions such as *sort of*, *you know*, *I mean*, which contribute little in content as they are effectively lexicalised pauses, but abound in L1 speech and can serve the purpose of combatting the pressures of online planning. Dörnyei and Kormos (1998) propose teaching learners to adopt them in order to maintain the flow of speech and Segalowitz (2010) and Chambers (1997) also refer to them as problem-solving strategies to avoid the emergence of disfluencies.

2.4.2. Intonation as an indicator of fluency

Wennerstrom (2000) argues that intonation makes an important contribution to fluent speech in L2 speakers, and should be considered as a variable of fluency. In her view, the use of pitch to differentiate items in the information structure of the discourse and to indicate utterance boundaries are essential components in the perception of fluency. If a speaker fails to assign appropriate emphasis to focus words in an utterance, instead giving equal pitch to every word, the relationships between words and phrases are lost,

and speech sounds disjointed. Moreover, the failure to use appropriate intonation at utterance boundaries makes turn-taking in interaction difficult. She concludes by saying that it is not longer utterances or fewer, shorter pauses that lead to a perception of fluent speech, but “the ability to speak phrasally rather than word-by word” (p. 125) and to use intonation to communicate the main idea of each utterance coherently.

2.5. Speaking tests and their descriptors

At this point it would be interesting to see how speaking tests interpret the construct. This is relevant because tests are informed by research, and many teachers, especially those who have to rate their learners and/or prepare them for recognised tests will be influenced by the test benchmarks.

Two of the best known testing systems in Europe are CEFR (Council of Europe, 2001) and the International English Language Testing System (IELTS). As far as the speaking components are concerned, there are four in the nine-band IELTS: Fluency and Coherence, Lexical Resource, Grammatical Range and Accuracy, and Pronunciation. According to IELTS, therefore, fluency and coherence are closely related. The six-band CEFR, on the other hand, divides the components of oral performance into Range, Accuracy, Fluency, Interaction, and Coherence. Fluency and coherence are seen as separate, and pronunciation is absent altogether. It should be noted that in both tests, range of lexis and accuracy are distinguished as separate components from fluency, and it is logical that they should only be assessed once. Given this, one would expect teachers who are also language testers or who prepare their students for official tests to draw a distinction between the three. Kormos and Dénes’ (2004) university teacher participants, however, consider accuracy and lexical diversity important when rating fluency.

The table in Appendix 3 groups a sample of the fluency descriptors for IELTS and CEFR according to Segalowitz’s three-way distinction to show how cognitive fluency, utterance fluency and perceived fluency are represented. It is interesting that both tests draw an explicit distinction between content-related and language-related hesitations. In IELTS one descriptor states *any hesitation is content-related rather than to find words or grammar*. An equivalent in CEFR is *Only a conceptually difficult subject can hinder*

a natural, smooth flow of language. However, this raises the question of whether the reason why a speaker hesitates can always be perceived.

2.6. Conclusion

Lennon suggested that "it might be possible to identify some variables that function as core fluency components and some that are peripheral" (Lennon, 1990: p. 413). As has been seen, a multitude of studies have attempted to identify the best predictors of perceived fluency, and most have identified a cluster of temporal variables as central to the question. Exactly which are the most influential, however, is not clear as findings are contradictory. It seems that twenty-five years on, we are still no closer to solving the problem.

The focus of the present study is what professional teachers in higher education consider to be the important components of fluency. It compares the responses of two different groups of university teachers: those based in Florence, Italy, with those working for universities in the UK. The salient difference between the two is that the former have a considerable amount of experience but relatively little training in EFL/TESOL as this is not required by the Italian university system, while in the UK training and teaching qualifications are part of the selection procedure. The three research questions, therefore, are as follows:

Research questions:

1. What aspects of fluency do teachers pay attention to when rating non-native speaker fluency?
2. Are there differences between the two groups of teachers in terms of perceptions?
3. Are teachers' perceptions correlated with training and years of experience?

3. Methodology

3.1. Introduction

In this chapter the procedure adopted to address the three research questions presented at the end of section 2.6, will be described in detail. In brief, three recordings of learner speech and an online questionnaire were produced and sent to university teachers in Italy and the UK. The teacher participants (TP) were asked to rate the recordings for fluency, and then answer questions of both a quantitative and qualitative nature. These questions were designed to stimulate their views on what fluency involves. The choice of learner participants and speaking task, a description of the TPs and the procedure followed in conducting the survey, the rationale behind the writing of the quantitative and qualitative questionnaire, and the tools used to analyse the data obtained will be discussed. Both general and specific instruments of research design will be referred to, and the care taken to ensure that ethical principles were observed will also be illustrated.

3.2 Learner participants

Three learner participants were selected to perform a recorded speaking task. It was decided to include women only, as voice could influence fluency ratings and therefore become a confounding variable. All three learners were non-native speakers, had taken IELTS in August 2014, and had been studying at the University of Reading for nine months previous to making the recording. They had obtained different scores for IELTS speaking and were of two different nationalities. Speaker 1, Kazakh, aged 24, had scored 6.5; Speaker 2, Chinese, aged 18, had achieved 5.5; Speaker 3, aged 25, Chinese, had scored 7.0. The three presented an interesting variety of strengths and weaknesses as regards speaking e.g. temporal aspects of fluency, use of repair, range of lexis and grammatical structures, and command of phonological features, and it was felt that this would stimulate different views on fluency.

It is recognised that the IELTS scores achieved by the learner participants were not very recent, and that their spoken performance could have improved since. However, the

recordings made were designed to stimulate reflection on the components of fluency rather than test the degree of consensus with official test ratings. The TPs were asked to give the three learners an impressionistic score for fluency to check the level of interrater reliability (section 4.2.), and this is important to give the study internal validity, but the main focus will be on what factors influenced their scores. It is also recognised that the specific speaking characteristics of the speakers selected may influence the responses and therefore skew the findings. For example, a speaker who frequently resorts to hesitations and filled pauses is likely to generate comments regarding the role of those variables in fluency/disfluency, while a speaker with a monotonous voice will stimulate observations on the importance of intonation. However, it is believed that experienced teachers will see beyond the specific examples when discussing the topic.

The learners, or speakers, were Masters or undergraduate students at the University of Reading. All three declared that they felt relaxed about being recorded, and showed no sign of embarrassment towards the recording device. This is important as anxiety can affect speakers' performance (Fulcher, 2003). It was made clear that the recordings would be confidential and would remain anonymous, and that they would be destroyed as soon as possible. Apart from age, L1 background and most recent IELTS score, no personal information was sought. Each participant read the information sheet and signed the consent form first (Appendices 6 and 7).

3.3. The speaking task

For the purposes of this study a task was required that would inspire each speaker to produce an audio recording of approximately three minutes of uninterrupted speech. After careful consideration, a picture story was chosen as the best task type to provide these samples for four reasons. First, a picture story does not make demands on the short-term memory or imagination for content, because the learners have the pictures in front of them as they speak (Fulcher, 2003). Secondly, it allows the tester to control the task. With methodical piloting, the tester can ensure that the pictures do not include any culturally remote images or require descriptive vocabulary beyond the learners' range. Thirdly, the sample produced will be uninterrupted monologue of a reasonable length,

and this facilitates scoring as no other factors, such as comprehension of the interlocutor, are involved. Finally, the use of pictures means that each speaker has the same guided task, but it allows some scope for creativeness. This helps raters make comparisons between the speakers (Fulcher, 2003).

Two picture stories were piloted with the help of three colleagues, and one was rejected because it was deemed too complex. The picture story chosen (Appendix 4) comes from *ELSA Handbook* 2003, and was found in *Testing second language speaking* (Fulcher, 2003, p. 69). It consists of six pictures given in the correct order and a relatively straightforward, linear storyline. The clear time sequence of events in the story lightens the processing burden for the speaker, and therefore the cognitive complexity, and this facilitates fluency and accuracy (Tavakoli & Foster, 2008).

Research shows that planning time increases fluency (Tavakoli & Skehan, 2005), and therefore the speakers were given one minute to think through their stories. They were encouraged to tell the story in the past tense and to make it as interesting as possible. They were then recorded using a digital voice recorder. In all, five speakers were recorded, and three were selected. One was rejected because the story is about a man winning the lottery and it emerged that the speaker, a Saudi Arabian, was unfamiliar with lotteries. In line with Skehan's (1998) discussion of cognitive familiarity, this lack of cultural knowledge affected her fluency. Another speaker was rejected because she was considerably more fluent than the others, and it was felt that a sample with a greater number of fluency problems would generate more discussion. Speaker 1 spoke for 2.16 minutes, Speaker 2 for 3 minutes, Speaker 3 for 2.53 minutes. The three recordings are on the CD.

3.4. Sampling strategy – the teacher participants

The TPs are all native-speaker English language teachers working at university level in Italy and the UK. Two countries were involved in order to see whether notions regarding fluency are local or more widespread. It seemed logical to consult professionals in the field for this exploratory study as language teachers frequently have to assess learner fluency (Kormos and Dénes, 2004). University teachers were chosen

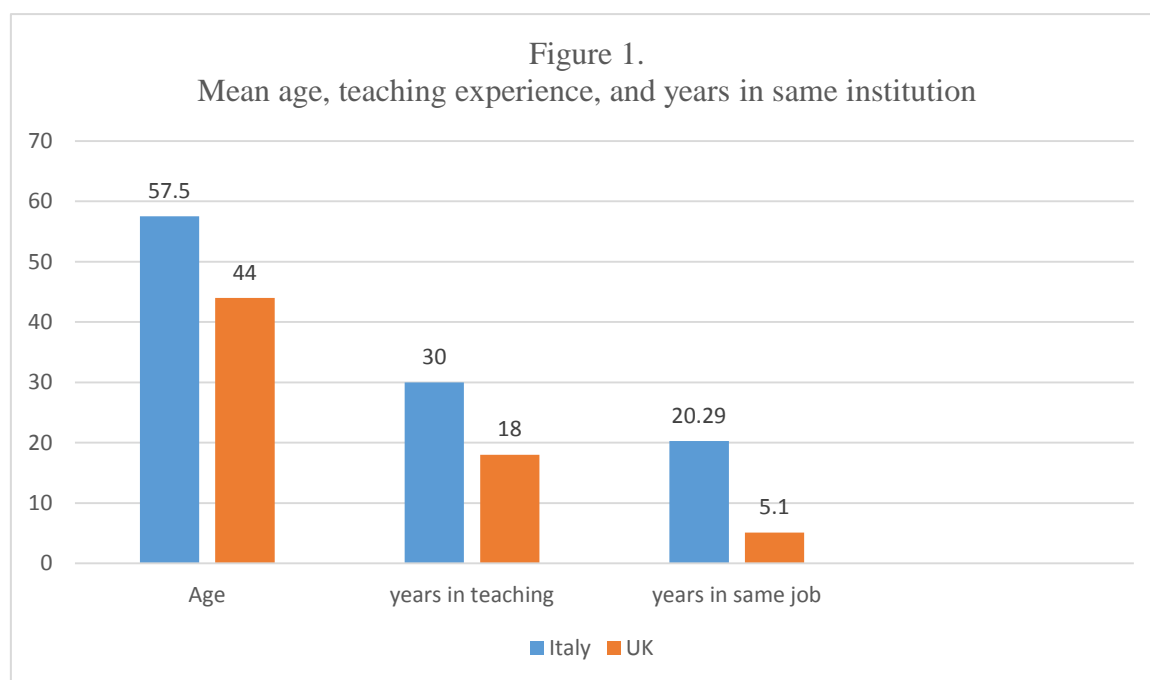
because higher education is the researcher's field of work and area of interest. As in Kormos and Dénes' study, therefore, the participant sample is an example of *purposeful sampling* (Dörnyei, 2007), where the participants possess certain key characteristics related to the purpose of the investigation, but also of *convenience sampling*, given that the sample is made up of willing participants from the institutions that the researcher has been associated with. Having decided that the questionnaire would be best distributed online, given the listening element and the different destinations, the challenge was to obtain a sufficient number of responses. Whereas if participants are approached face-to-face there is a higher likelihood that the survey will be completed, online communications are notoriously easy to ignore or forget. Email addresses of English language teachers at the Universities of Reading, Southampton and Bedfordshire in the UK, and Florence in Italy, were found using university websites, and a covering email was sent containing the link to the questionnaire and the recordings in MP3 format in attachments.

The questionnaire (Appendix 8) was created using the freely available software Google Forms (<https://www.google.com/forms/>), which allows for the creation of various types of questions. The software automatically and anonymously collates all responses submitted in an online spreadsheet. The anonymity of the questionnaire means that the ethical procedure is simple: participants give their consent by completing and submitting the questionnaire, and this is made explicit in its introduction. Reminders were sent at weekly intervals for three weeks, and 48 responses were received, 24 from each country.

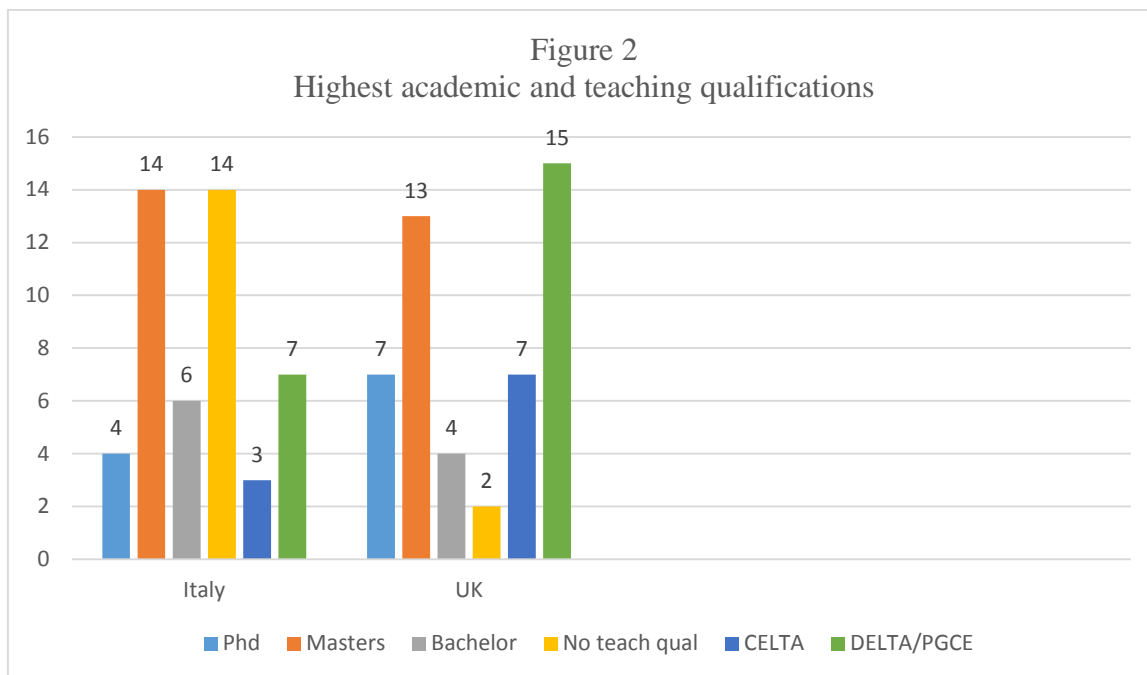
3.5. TP profile

A series of descriptive analyses conducted using IBM Statistical Package for the Social Sciences (SPSS) version 21 reveal some interesting data about the TPs and how they compare in the two countries. As regards gender, five of the 24 Italy-based TPs are male, 19 female, while there is a more even spread in the UK, with 10 males and 14 females. The statistics show considerable differences in age between the Italy-based TPs ($M = 57.5$, $SD = 6.52$) and the UK-based TPs ($M = 44$, $SD = 10.02$); and in the number of years in teaching in Italy ($M = 30$, $SD = 5.46$), and in the UK ($M = 18$, $SD = 9.86$).

The most sizeable difference, however, is in the number of years working in the current institution, with a mean of 20.29 years in Italy ($SD = 7.5$) compared to a mean of 5.1 years in the UK ($SD=4.01$). In Italy, therefore, teachers tend to stay in the same workplace for much longer periods (Figure 1).



As regards qualifications, the TPs were asked to state their highest academic qualification, that is, degrees at Bachelor, Masters or Doctorate level, including for subjects unrelated to language teaching. It was found that although in the UK marginally more of the TPs have PhDs or DPhils, on the whole the number and level of academic qualifications are similar in the two countries (Figure 2). The TPs were also asked what teaching qualifications they hold, and here the two groups of TPs differ considerably. It should be pointed out that MAs in subjects such as Applied Linguistics and English Language Teaching are not counted in this category as they are considered academic qualifications rather than practical teaching qualifications. DELTAs and PGCEs are grouped as equivalent, both being one-year full-time courses with a strong practical element, and in the results only the highest qualification was recorded i.e. if the TP has both a CELTA and a DELTA, only the DELTA was entered. Figure 2 shows that 14 (over 58%) of the 24 Italy-based respondents do not hold any teaching qualifications, although one of them has a related Masters degree. In the UK, two TPs have no practical teaching qualifications, but have related Masters degrees.



To sum up, the Italy-based respondents are generally older, have been working for the same institution for considerably longer, are equally well-qualified in terms of academic qualifications, but have fewer qualifications specifically for teaching. In fact, teaching qualifications are not required to work in Italian universities and are generally not recognised by commissions that select applicants for teaching posts.

3.6. Instruments

The decision to use a questionnaire with a quantitative and qualitative framework was informed by the researcher's reading on research methods in Applied Linguistics and by studies that helped shape the present study. Dörnyei (2007) discusses the strengths and weaknesses of quantitative and qualitative research and explains how quantitative inquiry can be more "objective" than qualitative research, being based on data analyses conducted using statistical computer software rather than the researcher's interpretation of answers. An important advantage of the quantitative approach is that if the study is designed well, the findings should be generalisable to other contexts. However, it can be "overly simplistic, decontextualized and reductionist" in its generalisations (Dörnyei, 2007, p. 35). Qualitative methods, on the other hand, are suited to smaller, exploratory

studies, and are usually more concerned with the subjective opinions of individuals. They give an “insider perspective”, the insiders in the present study being the teacher participants, and can produce richer data. The interpretation of this data, however, is left to the researcher, and is therefore subject to personal prejudices and idiosyncrasies. By combining the two approaches, i.e. by using mixed methods, the best of both paradigms can be exploited.

It was decided, therefore, to include a combination of closed-ended semantic differential scale questions (quantitative) and open-ended questions (qualitative) in the questionnaire. According to Dörnyei (2007), triangulation is a form of validity check. Kormos and Dénes (2004) and Freed (2000), two influential studies in the design of this research, also combined the two methods, although very little space is dedicated to the comments of the assessors in the former.

3.7. Questionnaire design

Following Dörnyei’s (2010) recommendations, the questions regarding biographical information were put at the end of the questionnaire. The rationale behind this is that participants may be put off by direct questions about their age, careers and qualifications at the beginning, and there is the risk that they will not continue.

It was decided that rating the three learners for fluency would be a dynamic lead-in to the topic. The TPs are asked to indicate a score on a seven-point semantic differential scale: 1= not very fluent, 7= extremely fluent. As Dörnyei (2010) points out, the greater the range of points, the greater the scope for sensitivity, but if the range is too wide, the choice becomes more difficult. A seven-point scale was chosen because Freed (2000), an inspiration for many aspects of the present study, opted for this range, and because an odd number allows for a mid point, should raters wish to be non-committal. It was also seen as a compromise to satisfy teachers familiar with IELTS scoring, and therefore 9-point scales, and those used to Cambridge English exams, which mark out of 5.

Having rated the speakers, the respondents are asked four open-ended questions, and these allow them to provide richer data than quantitative questions allow for. Care was taken to avoid influencing respondents in the wording of the questions (Dörnyei, 2010).

Question 4 asks the TPs to justify their choice of the most fluent speaker, while question 5 widens the topic by asking them to list the key characteristics of fluent speech in general. Questions 6 and 7 are designed to approach the topic from different angles, as Dörnyei (2010) suggests. Question 6 (*What do you think are the underlying causes that make speech disfluent?*), taps into views on cognitive processing and conditions that can affect fluency, while Question 7 (*To what extent do you think speech fluency is in the ear of the listener?*) focuses on perceived fluency and, indirectly, whether fluency can be rated objectively. The four open-ended questions answer Research Question 1 (see Table 2, p. 26).

In the quantitative questions that follow, respondents are asked to rate 20 variables for importance in assessing fluency on a 5-point semantic differential scale (1= not important, 5= most important). This section of the questionnaire was also inspired by Freed (2010), although Freed gave only 8 variables and her judges ticked the ones they considered important. It was felt that a scale for each individual variable was more precise than Freed's approach. The respondents were given the opportunity to add any further thoughts in a separate open-ended question (question 28).

The questionnaire was piloted by sending it to four colleagues, two retired ex-colleagues in Italy, and two Masters colleagues. All four gave either written or face-to-face feedback, and minor changes were subsequently made e.g. the labelling of variables to avoid jargon and potential confusion.

3.8. Grouping the variables

The 20 variables were chosen on the basis of the researcher's reading of previous fluency studies, books on the topic, and speaking test descriptors (CEFR and IELTS). It was decided to group the variables as this increases the likelihood of patterns emerging. Influenced in particular by the literature on Complexity, Accuracy and Fluency (CAF) (Housen & Kuiken, 2009), the researcher categorised the variables in four groups: Complexity, Fluency, Phonology and Global Aspects (Table 1).

The origins of the CAF triad lie in the SLA literature of the 1980s, when a distinction was made between accuracy and fluency (Brumfit, 1984). In communicative language

teaching, classroom activities are often classified as “fluency-based” or “accuracy-based” according to the language learning they are designed to enhance. Moreover, learners are often seen to be more “norm-oriented” or “communicative-oriented”, the former being learners who focus on learning the grammatical rules and accuracy, possibly at the expense of fluency, while the latter aim to communicate fluently but pay less attention to accuracy (Fulcher, 2003). Complexity was added to the triad in the 1990s as a result of advances in cognitive psychology and psycholinguistics (Levelt, 1989; Anderson, 1996), and research into the trade-off between complexity, accuracy and fluency as they compete for attentional resources (Skehan, 1998; Tavakoli et al., 2015). Housen and Kuiken (2009) propose a definition of each of the three constructs. Accuracy, the easiest to define, is “the ability to produce error-free speech” (p. 461); fluency is “the ability to process the L2 with native-like rapidity” (Lennon, 1990, p. 390) or “the extent to which language produced in performing a task manifests pausing, hesitation, or reformulation” (Ellis, 2003, p. 342); complexity is defined as “the extent to which the language produced in performing a task is elaborate and varied” (Ellis, 2003 p. 340).

Table 1: *Grouping of the 20 variables*

Fluency	number and length of pauses, number of hesitations, length of runs, rate of speech, repetitions, reformulations, appropriate use of fillers, automaticity, effortlessness
Complexity	grammatical accuracy, variety of vocabulary, complexity of topic content, complexity of structures
Phonology	accent, native-like rhythm
Global Aspects	global proficiency, communicative competence, coherence, formulaic sequences, <u>colloquialness</u>

The Fluency group, therefore, includes variables that reflect Segalowitz’s cognitive fluency, or smooth efficient language processing, and the resulting surface or utterance fluency, which involves variables of speed, breakdown and repair. The Complexity variables comprise *variety of lexis*, *complexity of structures* and *complexity of topic content*, the rationale for including the last being that fluency may be considered “false

fluency” if the content spoken language is too trivial and incoherent (Lennon, 2000, p.40). For grouping purposes, *accuracy* was also included here on the premise that accuracy, like complexity, is linked to the current state of the learner’s interlanguage knowledge, that is, the extent to which the L2 has been internalised (Housen & Kuiken, 2009). The Phonology group is made up of two variables, *accent* and *native-like rhythm*, while Global Aspects (GA) includes variables that refer to *global proficiency*, such as *communicative competence* (phrased in the questionnaire as “the ability to speak in different social situations” for the sake of clarity), *coherence*, and the appropriate use of *formulaic sequences*, (phrased in the questionnaire as “ability to include ready-made chunks/expressions and idiomatic language” to avoid technical language). *Colloquial* language, often formulaic in nature, related to Segalowitz’s “naturalness,” and mentioned in the top band of CEFR fluency descriptors, is also included in this group.

Grouping the variables was not straightforward, and some of them could arguably be categorised differently. *Formulaic sequences*, for example, could be placed in the Complexity grouping along with *variety of lexis* instead of in GA, given that they are multi-word lexical units (Wray, 2002). They could also be grouped under Fluency, as length of run is frequently associated with prefabricated language units (Kahng, 2014). However, they were grouped in GA because of their functional diversity. *Native-like rhythm* falls naturally into Phonology, but it is also related to pace, which is a temporal variable. Pace is the number of stressed syllables per minute (Kormos, 2006), and it is stress that determines rhythm. Rhythm also results from grouping words into runs. There is an argument, therefore, for grouping *rhythm* under Fluency. In short, these groups are open to debate, but grouping is opportune as a point of departure.

Cronbach’s alpha coefficient tests were run to check the internal reliability of the groups of variables, that is, to see whether they are all measuring the same underlying construct. Ideally, Cronbach’s alpha coefficient should be above .7 (Pallant, 2010). However, with scales with fewer than ten items it is common to find low Cronbach values (e.g. .5). The Cronbach alpha coefficient was acceptable for three of the groups (Fluency group = .830, Complexity= .821, Global Aspects = .747), while Phonology, with only two items, achieved a value of .652.

3.9. Data analysis

Once 48 TPs had submitted their responses, the quantitative data were transferred from the Google spreadsheet to SPSS and tests were run to answer the research questions.

To analyse the qualitative data, the procedures proposed by Dörnyei (2007) were followed, and Tavakoli's (2009) qualitative study was used as an example. The respondents' answers were coded according to the variables they mention, the TP country of origin, and the themes addressed. As discussed in section 3.6, qualitative data analysis requires some interpretation on the researcher's part, but validity checks were carried out by comparing the qualitative with the quantitative data to identify correspondences and contradictions, and by counting the references to variables and themes.

Table 2 shows how the data obtained was used to answer each research question.

Table 2: *breakdown of how the data obtained was used to answer each research question*

Research questions	Source of data	Analysis
RQ1: What aspects of fluency do teachers pay attention to when rating non-native speaker fluency?	Questionnaire questions 4,5,6,7,28 (qualitative)	Group open-ended questions thematically
	Questions 1-3, 8-27 (quantitative)	Descriptive statistics using SPSS
RQ2: Are there differences between the two groups of teachers in terms of perceptions?	Questions 8-27 (quantitative)	Non-parametric Mann-Whitney U Test Descriptive statistics according to two countries
RQ3: Are teachers' perceptions correlated with training and years of experience?	Questions 32, 33 and 35	Non-parametric Spearman's Rank Order Correlation

4. Analysis

4.1. Introduction

In this chapter the data collected will be analysed to answer the three research questions presented at the end of Chapter 1. All three questions involved quantitative methods, and the statistical analyses were conducted using SPSS. For the first research question the quantitative findings will be compared with the qualitative data collected from the four open-ended questions of the questionnaire. Two of these focus on the variables the respondents consider important for rating fluency, and the answers will be organised according to the four groups described in the Methodology chapter (Table 1). The answers to the other two open-ended questions will be grouped thematically. The second research question also includes some references to the open-ended answers.

4.2. Global rating of the three speakers

Before answering the research questions, the respondents' global ratings of the three speakers will be compared. It will be remembered that they were required to give a score for fluency from a low 1 to a high 7. Overall, 89.6% of the respondents agreed that the third speaker was the best, while the remaining 10.4% (5 respondents) awarded the highest score to the first speaker (Table 3). Three of these five respondents are based in Italy, two in the UK. It is interesting that the IELTS speaking scores (Speaker 1: 6.5, Speaker 2: 5.5, Speaker 3: 7) are still reflected in these ratings, despite being nine months old.

Table 3: *Best speaker according to TPs in each country*

Country		Frequency	Percent
Italy <i>n</i> = 24	speaker 1	3	12.5
	speaker 3	21	87.5
UK <i>n</i> = 24	speaker 1	2	8.3
	speaker 3	22	91.7

Table 4 shows that the mean ratings for each speaker were similar if not identical in the two countries, and the ranges in the scores are consistent. The standard deviations (*SD*) are indicative of the level of variance, the higher values in Italy indicating greater variance between ratings, the lower values showing more consistency in the UK.

Table 4: *Global ratings for the speakers in each country*

Country	Speaker	Minimum	Maximum	Mean	SD
Italy	Speaker 1	2	5	3.96	1.042
UK		3	5	3.96	.550
Italy	Speaker 2	1	4	2.42	.929
UK		1	4	2.63	.647
Italy	Speaker 3	3	7	5.00	.978
UK		3	6	5.29	.806

4.3. Research question 1: What aspects of fluency do teachers pay attention to when rating non-native speaker fluency?

4.3.1. Quantitative data

Having seen that the global ratings are similar, the factors that influenced TP assessment will be analysed. The TPs were asked to rate twenty variables in order of importance on a 1 to 5 scale. As discussed in the Methodology chapter, these were categorised into four groups: Complexity, Fluency, Phonology and Global Aspects.

Table 5 shows the range, means and variance that each variable achieved. The highest means were achieved for *effortlessness* ($M = 4.08$) in the Fluency category and *coherence* ($M = 4.02$) in the GA category, followed by a cluster of other Fluency variables with a mean of over 3.5 (*automaticity*, *number and length of pauses*, *hesitations*, *rate of speech*, and *length of runs*), and three variables in the GA group: *communicative competence* ($M = 3.69$), *global proficiency* ($M = 3.6$), and *formulaic sequences* ($M = 3.5$). The lowest means were obtained by *accent* ($M = 2.48$), *accuracy* ($M = 2.69$) and *colloquialness* ($M = 2.69$).

Table 5: *Descriptive statistics for the 20 variables*

Variables (according to groups)	Minimum	Maximum	Mean	SD
FLUENCY				
Effortlessness	3	5	4.08	.710
Number & length of pauses	2	5	3.83	.753
Automaticity	2	5	3.79	.898
No. of hesitations	1	5	3.79	.898
Length of Runs	2	5	3.73	.676
Rate of speech	1	5	3.69	.748
Reformulations	2	5	3.48	.850
Appropriate fillers	1	5	3.38	.890
Repetition	1	5	3.02	.934
COMPLEXITY				
Complexity of topic content	1	5	3.08	1.127
Variety of vocabulary	1	5	3.02	1.120
Complexity of structures	1	5	2.79	1.110
Accuracy	1	5	2.69	1.095
PHONOLOGY				
Native-like rhythm	1	5	3.48	1.271
Accent	1	5	2.48	1.052
GLOBAL ASPECTS				
Coherence	1	5	4.02	.956
Different social situations	1	5	3.69	1.133
Global proficiency	1	5	3.60	1.026
Formulaic sequences	1	5	3.50	1.185
Colloquialness	1	5	2.69	1.095

The least variance was for *length of runs* ($SD = .676$), and *effortlessness* ($SD = .710$), this indicating that consensus was mostly consistent for these two high-scoring variables, with ratings clustering around the mean. In fact, a glance at the SD column shows that consensus is consistently higher for the Fluency variables, while all the other

variables show greater variance, albeit coherence to a lesser extent. The least consensus was for *native-like rhythm* ($SD=1.271$) and *formulaic sequences* ($SD=1.185$).

It seems from the descriptive statistics, therefore, that apart from *fillers* and *repetition*, which are rated marginally lower, all the Fluency variables are considered important. The GA variables, with the exception of *colloquialness*, also receive high means, with *coherence* rated remarkably highly.

4.3.2. Qualitative data

Questions 4 and 5

The first two open-ended questions in the questionnaire focus on what the TPs consider to be the key variables of fluent speech, and therefore the responses can be compared with the quantitative data above. Question 4 asks respondents to justify their choice of the most fluent speaker, while question 5 asks them to identify the main characteristics of fluent speech in general. The answers will be grouped according to the four groups of variables, with subsections where subthemes emerge. For reasons of space, only a representative sample of the considerable amount of text received will be reported. Further selected excerpts can be found in Appendix 7.

Table 6 shows the percentage of mentions for each group of variables in the answers. For example, while answering question 5, 78% of the 48 respondents identified one or more Fluency variables as important. The data present a general picture of the emphasis given to each group. It is immediately obvious that for both questions, the greatest importance was given to Fluency variables, with Phonology in second place, Global Aspects in third, and Complexity variables given little importance. This corroborates the quantitative results for Fluency and Complexity, but the importance given to Phonology is surprisingly high. It should be remembered that categorisation of the factors mentioned often required interpretation on the researcher's part. For example, reference to a speaker's use of "expression" was categorised under Phonology, as expression is conveyed mostly through prosodic features such as intonation.

Table 6: *Percentage of open-ended answers that identify variables in each group*

Group of variables	Question 4	Question 5
Fluency	64.9%	78%
Phonology	54%	67.5%
Global Aspects	45.9%	56.7%
Complexity	18.9%	27%

4.3.2.1. Fluency Variables

As can be seen from the percentages above, the answers reflect considerable importance attributed to the Fluency group of variables, with a large number of the respondents mentioning breakdown and repair variables. Excerpts will be organised according to Segalowitz's three-way distinction (section 2.3.). However, subthemes will also be introduced.

4.3.2.1.1. Utterance fluency

Hesitations, pauses and repair are frequently recurring themes, and this confirms the importance awarded in the quantitative results:

TP24: no excessive hesitation or pausing

TP20: Any repair is done unobtrusively

However, as regards temporal features, or speed fluency, eight respondents express reservations about the relevance of speed. For example:

TP23: fluency has little to do with speed

TP35: Speed of speech is also significant, but some people speak slowly whatever language they are speaking

Chunking/runs

As many as fifteen respondents talk about the effectiveness of packaging words into runs, or "chunks." Sometimes these runs are defined as "thought groups", and a link is occasionally made with rhythm:

TP20: packages words into meaningful thought groups, with appropriate pauses and rhythm

TP7: Natural chunking - language comes out as strings rather than discrete words

Flow

Ten TPs use the word “flow” in their answers, which shows recognition of the etymology of the word *fluency*, of utterance fluency, but also of underlying cognitive processes:

TP18: "fluency" means "fluid/flowing" speech, with no gaps

TP3: flow of ideas

4.3.2.1.2. Cognitive fluency

Efficient cognitive processing is a key factor for thirteen TPs, although they use different terminology. Several respondents distinguish between breakdown due to linguistic processing and the search for content:

TP26: minimal pauses/hesitations which show lack of vocabulary/ grammar rather than thinking of another point

TP 37: seems to be able to think and talk at the same time - processing grammar and vocabulary do not get in the way of producing the next utterance

4.3.2.1.3. Perceived fluency

Closely related to cognitive fluency is the perception that the speaker's cognitive machinery is efficient (Segalowitz, 2010). Words such as *control*, *automatic*, *effortless*, and *ease* in the TP answers convey this impression:

TP26: more control over language

TP34: fairly effortlessly

TP15: giving your interlocutor the impression that you can speak freely and easily

4.3.2.2. Complexity Variables

The low percentage of mentions of variables in the Complexity group (18.9 % for question 4, 27% for question 5) indicates that relatively few respondents identify range of lexis and grammar as key components of fluency. This supports the quantitative data, where the four Complexity variables reveal lower means. However, as might be expected, there are some strong advocates of accuracy and a wide vocabulary:

TP24: grammatical accuracy and range are key factors

TP16: a wide vocabulary range is important for fluency

4.3.2.3. Phonology Variables

As indicated by the high percentage of mentions (54% for question 4, 67.5% for question 5), the open-ended answers reveal a remarkably strong tendency to be influenced by prosodic features when rating fluency, particularly intonation and sentence stress or rhythm. This was not so apparent from the quantitative analysis, where only two Phonology variables were listed and *accent* was rated relatively low in importance, while *native-like rhythm* was fairly important, with a high degree of variance for both variables. Unfortunately, there was no question that related specifically to the importance of intonation in the quantitative section of the questionnaire, as triangulation might have consolidated this interesting finding. The following are only three of numerous references to phonological features:

TP8: the third is the most fluent because of her use of intonation

TP35: native-like rhythm of speech helps a lot in giving an impression of fluency

TP29: good intonation, rhythm and sentence stress

Delivery

Moreover, twenty-six respondents referred to aspects of delivery (use of voice, volume, expression, ability to pause effectively, emphatic stress, enunciation) that influence their fluency ratings e.g.

TP11: voice quality, especially pitch, contributes to a listener's perception of fluency

4.3.2.4. Global Aspects

Coherence

The considerable importance given to coherence in the quantitative data is reinforced by the open-ended answers. Eighteen respondents refer specifically to this variable. Many of these associate coherence with the use of cohesive devices such as linking words, conjunctions, and discourse markers:

TP3: ability to express ideas in a coherent, logical manner

TP46: Fluency is intimately connected to coherence and cohesion

Thirteen of the statements categorised under Global Aspects are more general:

TP7: a greater knowledge of language

TP19: successful, unambiguous communication of intended meaning

4.3.3. The causes of disfluency

Question 6 of the questionnaire asks respondents to reflect on the underlying causes that make speech dysfluent. It is designed to approach the topic from a different angle. The responses received will be categorised into three groups: 1. cognitive processing, including dealing with L1 interference 2. linguistic complexity 3. learner personality and affective factors.

4.3.3.1. Cognitive processing

Fourteen respondents refer to the inability to access appropriate language rapidly as a factor in dysfluency:

TP25: speaker is searching for a word or word form perhaps because they are translating or applying grammatical rules

TP28: pauses to think of the next word or plan next chunk

Lack of practice

However, as discussed in section 2.3.1., practice generally increases automaticity. Ten TPs identify lack of practice as a cause of disfluency:

TP15: L2 speakers who do not mix with native English speakers often remain less fluent

TP29: Certainly lack of practice

Cognitive (un)familiarity

The following two obstacles to fluency would be classified by Skehan (1998) under cognitive complexity:

TP43: unfamiliarity with the topic or task

TP19: not understanding or relating to the content of what you are talking about

L1-L2 interference

According to de Bot (1992), L1 interference is also a problem of cognitive processing. Twelve respondents identify this as a cause of disfluency:

TP27: L2 hasn't been internalised - it sounds like a translation

TP16: difficulty in pronouncing sounds which do not exist in L1

4.3.3.2. Linguistic complexity

Insufficient language knowledge, whether lexical, grammatical or command of the phonological system, was cited by fourteen participants as a reason for poor fluency:

TP20: lack of knowledge of the language, particularly vocabulary

TP2: low level of language proficiency

4.3.3.3. Personality and affective factors

Task difficulty, according to Robinson's (2001) three-way framework, can depend on learner-related factors such as motivation, aptitude and personality. Almost half of the 48 respondents in this study identify personality and affective factors as causes of fluency/disfluency, with self-confidence being mentioned the most frequently:

TP7: self-consciousness, insecurity, shyness

TP22: fluency is related to confidence

Amongst the assortment of other factors identified by individuals as causes of disfluency are performance conditions e.g. time pressure, and lack of paraphrasing skills.

4.3.4. The subjective element

The fourth open question requires respondents to consider the extent to which perceptions of fluency are “in the ear of the listener”, or subjective. Grouping their answers roughly into three categories: Yes/Partly/No, 60% of the respondents express the view that fluency is subjective, 22% that it is to some extent, and only 10% say that the characteristics of fluency are clear-cut. The arguments of the 60% that feel that fluency is subjective will be categorised into two basic themes:

- it depends on whether the listener is a language specialist or not
- listener preferences and motivation affect assessment

4.3.4.1. Language specialist versus lay person

Ten answers distinguish between language specialist and lay perceptions, thus reflecting Lennon's (1990) distinction between the broad and the narrow sense of fluency:

TP5: Most people use fluency to refer to a person's overall language ability, whereas language teachers use it to refer to a specific aspect of a person's spoken production

TP41: For the 'ordinary' listener, fluency is ascribed to a speaker as a reaction ... which will certainly be individual. For the 'specialist' listener, such as language teachers, fluency is ascribed by matching the speaker's performance against a more defined series of competencies.

4.3.4.2. Listener preferences

Twenty-one answers to question 7 focus on listener motivation and individual preferences and intolerances. Interestingly, phonological factors such as intonation, rhythm and accent, and features of delivery, that is, expression and quality of voice loom large again:

TP3: if you are interested in a topic you may perceive someone to be more/less fluent than someone who is not

TP6: definitely fluency depends on the reaction solicited in the ear of the listener: pleasant or unpleasant to listen to

TP18: different people have different levels of tolerance towards prosodic and grammatical errors

The listener's background

Seven TPs identify the listener's experience of L2 speech as a factor that influences perceptions of fluency:

TP6: The listener's background and exposure to certain foreign accents and pronunciations is definitely a factor

TP34: Familiarity with certain accents or language problems increases your tolerance

In conclusion, four respondents express the opinion that although professional teachers narrow down the meaning of the term, even they never fully escape subjective considerations when rating fluency:

TP19: Although a teacher may wish to react as a professional, following set guidelines and parameters, s/he should always be aware that s/he carries with him/her a whole range of personal preferences which may influence his/her opinion of a student's fluency level

TP22: I think that professionally, teachers and assessors tend to decide what this term means, but, by defining it, they become a standardised 'beholder' (which is, nonetheless, a beholder).

4.4. Research Question 2: Are there differences between the two groups of teachers in terms of perceptions?

In order to establish whether the Italy-based teachers and UK-based teachers differ in their perceptions of the importance of the variables, Mann-Whitney U Tests were run. This is the non-parametric alternative to the t-tests for independent samples and compares medians instead of means. It is appropriate for ordinal variables rather than continuous variables, and it is used when it is assumed that data are not evenly distributed. This is appropriate, given the differences between the profiles of the two TP groups (section 3.5) and the high degree of variance in many of the ratings.

Table 7 shows the data where significant differences between the two countries were revealed, that is, for the variables *accuracy*, *accent*, *colloquialness*, *communicative competence*, *hesitations* and *automaticity*. The effect size for each significant difference

is given in the final column, and can be interpreted according to Cohen's (1988, pp. 79-81) criteria of .1= small effect, .3= medium effect, .5= large effect. By looking at the medians for each country, it can be seen where the variable was rated more highly. In the case of *hesitations*, the medians are identical but the means showed that the variable is considered more important in the UK ($M = 4.08$, $SD = .654$) than in Italy ($M = 3.5$, $SD = 1.022$).

Table 7: Mann-Whitney U-tests to measure differences in perceptions in Italy and UK

	median	U	z	p	Effect size r value
accuracy	Italy: 3 UK: 2	174.500	-2.453	.014	.35
accent	Italy: 3 UK: 2	126.0	-3.498	.000	.504
colloquialness	Italy: 3 UK: 2	178.0	-2.353	.019	.34
communicative competence	Italy: 4 UK: 3	190.500	-2.090	.037	.302
hesitations	Italy: 4 UK: 4	196.000	-2.066	.039	.29
automaticity	Italy: 3 UK: 4	153.000	-2.945	.003	.42

In Italy, therefore, *accuracy*, *accent*, *colloquialness* and *communicative competence* are rated significantly more important in rating fluency, while the two variables from the Fluency group, *automaticity* and *hesitations*, are considered more influential in the UK. The effect sizes are all moderate, apart from for *accent*, where it is large, and *hesitations*, where it is small.

These findings are partly corroborated by the qualitative data. The number of references to the importance of these six variables in the open-ended answers, although often in different words, are compared in Table 8. The numbers for each country reflect the general direction of the Mann-Whitney U Tests. However, it should be remembered that collating the results involved interpretation on the part of the researcher. For example, "use of natural fillers and idiomatic language" was classified as synonymous with

colloquialness. None of the open-ended answers seemed to refer to *communicative competence* i.e. awareness of appropriate language for different social situations.

Table 8: *number of mentions in qualitative data of significantly different variables according to country*

	accuracy	accent	colloquialness	comm. comp.	automaticity	hesitations
Italy	13	10	3	0	2	13
UK	10	5	1	0	5	16

4.5. Research Question 3: Are teachers' perceptions correlated with training and years of experience?

The relationship between teachers' perceptions and their qualifications and years of experience was investigated using the non-parametric Spearman's Rank Order Correlation. This was used for the same reasons as stated in section 4.4. The three variables *highest academic qualification* (including those unrelated to language teaching), *highest teaching qualification* and *the number of years in language teaching* were correlated with the twenty variables in their four groups. For reasons of space, only variables where significant correlations were found are shown in Table 9.

The guidelines established by Cohen (1988) were used for effect size again.

As far as *highest academic qualification* is concerned, the only significant correlations found were with *native-like rhythm*, with a medium negative correlation between the two variables ($\rho = -.459, n = 48, p < 0.01$), and a smaller negative correlation with *accent* ($\rho = -.308, n = 48, p < 0.05$). *Academic qualifications* and *native-like rhythm* share quite a large variance of 21%, while the coefficient of determination between *academic qualifications* and *accent* is only 9.5%. The more highly academically qualified respondents, therefore, seem to perceive less importance in the two Phonology variables.

Table 9: Significant correlations between Highest Academic Qualification, Highest Teaching Qualification, Number of Years of Experience, and variables

Variable Group	Variable	High. Ac.Qual	High.Teach. Qual	Years Teaching
		rho	rho	rho
Fluency	Automaticity	-.184	.464**	-.174
	Rate of speech	.143	.061	-.509**
Phonology	Accent	-.308*	-.097	.352*
	Rhythm	-.459**	-.166	.300*
Complexity	Accuracy	-.214	-.162	.413**
	Variety Lexis	-.039	.008	.467**
	Complex. Struct	-.133	.035	.429**
Global Aspects	Colloquialness	-.265	-.035	.420**

*Correlation is significant at $p < .05$ level

** Correlation is significant at $p < .01$ level

As regards the relationship between *teaching qualifications*, so specific training, and the twenty variables, the results revealed a moderately strong positive correlation with only one variable: *automaticity* ($\rho = .463$, $n = 48$, $p < 0.01$). The more training, therefore, the greater the awareness of the importance of *automaticity* for fluency.

The *number of years teaching*, or teaching experience, revealed some more interesting results, with a strong negative correlation between experience and *rate of speech* ($\rho = .509$, $n = 48$, $p < 0.01$) and a shared variance of 25.9%. It seems that more experienced teachers are unimpressed by rapid L2 speech. Moderate positive correlations, on the other hand, were found between experience and *accuracy* ($\rho = .413$, $n = 48$, $p < 0.01$), *variety of vocabulary* ($\rho = .467$, $n = 48$, $p < 0.01$), and *complexity of structures* ($\rho = .429$, $n = 48$, $p < 0.01$). Three out of the four Complexity variables, therefore, are seen as important by teachers who have been in the profession longer. There were also moderate positive correlations between number of years teaching and *colloquialness* ($\rho = .420$, $n = 48$, $p < 0.01$), suggesting that more experienced teachers rate the use of

colloquial language in the L2 highly, and smaller correlations with *accent* ($\rho = .352$, $n = 48$, $p < 0.01$), and *native-like rhythm* ($\rho = .300$, $n = 48$, $p < 0.01$).

4.6. Conclusion

To briefly review the main results, the quantitative data reveal that the variables in the Fluency group are all considered important, with the repair variables achieving slightly lower means. This is reflected in the open-ended responses, although the repair variables regain importance here, and reservations are expressed about the relationship between fluency and speed. A striking amount of importance is attributed to chunking, or length of run, in the participants' reflections. The variables in the Global Aspects group also achieve high means in the quantitative data, especially coherence, and the importance of coherence is clearly reflected in the qualitative answers. However, the participants do not spontaneously mention the other variables in Global Aspects when allowed a free rein. Instead they attribute a remarkable amount of importance to Phonology, especially intonation and rhythm. The Complexity variables are deemed the least important in both sets of data. These findings, along with the implications of the correlations between qualifications, teaching experience and perceptions of the 20 variables, will be discussed in more detail in the following chapter.

5. Discussion

5.1.Introduction

This chapter will summarise and discuss the findings presented in chapter four in light of the research questions and the literature review. The first section will address the first and main research question, designed to determine which factors influence university teachers when they rate L2 speaker fluency. This will be followed by a brief discussion of the differences between the two groups of TPs in terms of perceptions and how these correlate with training and experience. References will be made to related studies where appropriate. On the basis of the study results, a tentative proposal will be made regarding what a definition of fluency should include, and the implications for pedagogy considered. Finally, the limitations of the current study will be identified.

5.2. Key aspects of fluency

As far as the quantitative findings are concerned, two of the most highly rated variables reflect cognitive efficiency or, given that we cannot see or hear the workings of a speaker's brain, perceived cognitive efficiency. *Effortlessness* and *automaticity* achieve the highest and fourth highest means, and many views on fluency expressed in the qualitative data reflect Lennon's definition that "fluency is an impression on the listener's part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently" (1990, p.391). One of the TPs in the current study echoed this quotation perfectly with "fluency is giving your interlocutor the impression that you can speak freely and easily without constantly interrupting yourself in the search for words or ideas." Segalowitz's view of perceived fluency, therefore, is well represented in the data.

One of the most interesting findings in the analysis, reflected in both the qualitative and the quantitative data, is the importance attributed to *coherence*. The perceived relevance of this variable can also be explained in terms of cognitive machinery. If a speaker is in control of his/her cognitive processing, attentional resources will not be focused on

formulating and articulating but will be free to plan and conceptualise. The resulting speech is more likely to be lucid and coherent. This fits with Lennon's (1990) claim that "fluency reflects the speaker's ability to focus the listener's attention on his/her message by presenting a finished product...." (pp. 391-2). The notion of "presenting a finished product" is neatly expressed in Wright's (2014) recently coined term "performative competence." However, as previously pointed out, the IELTS speaking test assesses Fluency and Coherence as a single component of speech, demonstrating that the notion of coherence as an indicator of cognitive fluency is not new.

In line with previous research (Lennon, 1990; Kormos & Dénes, 2004; Freed, 2000), a cluster of temporal and breakdown variables, components of Segalowitz's utterance fluency, are rated as important in both the quantitative and qualitative data. Reservations were expressed about speed, however, with some respondents arguing that speed alone was not synonymous with fluency and that some speakers speak slowly in their L1. This is in line with De Jong et al. (2015) and Derwing et al. (2009)'s studies, which demonstrate that temporal, breakdown and repair features of L1 speech are carried over into L2 speech. Moreover, the current study found a significant negative correlation between years in teaching and the perception of rapid speech as an indicator of fluency. It would be interesting to carry out further research to verify this as it goes against Kormos' (2006) claim that most studies identify speech rate as one of the two best predictors of fluency.

Kormos and Dénes (2004) and Bosker et al. (2012) report on the relative unimportance of repair measures, and in the current study the variables *reformulation*, *fillers* and *repetition* are rated slightly lower in importance than the other Fluency variables. However, the means obtained are still relatively high, and several TPs mention the importance of fewer self-corrections as an indicator of fluency in the qualitative data.

Kormos' (2006) other most important measure of fluency is length of runs, and the TPs in the current study voice clear agreement over this variable. The high score in the quantitative data is expressed in words in the qualitative data with the many references to the effectiveness of chunking ideas, or packaging words into "thought groups". Chunking also implies an overlap with phonological variables. If a speaker speaks in

runs, or chunks, this will inevitably affect the stress patterns of speech, and therefore intonation. Rhythm and intonation are repeatedly identified as influential in rating fluency in the respondents' open-ended answers. The perceived importance of *native-like rhythm* is substantiated in the relatively high mean achieved in the descriptive data ($M = 3.48$, $SD = 1.271$), particularly in Italy ($M = 3.75$, $SD = .989$), and this is in line with Kormos and Dénes' (2004) conclusion that a definition of fluency should include pace. Stress is an element of intonation, and in phonological terms connected speech is made up of tone units with tonic syllables, heads and tails (Roach, 2009), which combine to produce patterns of rising and falling pitch. Wennerstrom's (2000) study argues in favour of considering intonation a variable of fluency, and demonstrates how L2 speakers who use a broad pitch range and pause appropriately are rated by trained raters as more fluent than speakers who use a limited pitch range. Her concluding statement that fluency is "the ability to speak phrasally rather than word-by word" (p. 125) is similar to the notion of chunking popular in this study. In short, length of run, chunking, stress, rhythm and intonation are found important for fluency and they are interrelated.

The aspects of fluency discussed in this section are considered important by both groups of TPs, and therefore it is the researcher's view that a definition of fluency should include them. There is less consensus across the two groups for other variables.

5.3. Differences between the two groups of TPs: perceptions, training and experience

As far as the Complexity variables are concerned, *accuracy* receives a low global importance rating in the quantitative data, and the qualitative data second this, although those who do argue in favour of its importance express strong views. *Accuracy*, *variety of lexis* and *complexity of structures* are all seen to be more important for fluency in Italy than in the UK, *accuracy* significantly so, and all three variables are significantly correlated with years of experience. It was shown in section 3.5 that the teachers based in Italy have mostly been in the profession and in the same workplace considerably longer than the UK-based teachers. This raises the question of whether teachers are better informed by experience, or whether their ideas become fossilised over time.

A contributing factor in shaping ideas is likely to be past training. As pointed out in section 2.5, in the IELTS and CEFR speaking tests accuracy, range of lexis and fluency are assessed as three separate components. It seems, therefore, that by attributing little importance to accuracy and lexis when rating fluency, the UK-based participants are more in line with the official test criteria than their Italy-based colleagues. This may be because they are more involved in language testing, and/or it could be because they have more qualifications specifically for teaching, and have been trained to see fluency in Lennon's narrow sense. While the UK-based teachers are significantly more influenced by *automaticity* and *hesitations*, which figure in the official test descriptors, the Italy-based colleagues, besides the Complexity variables already discussed, perceive *accent*, *communicative competence* and *colloquialness* as significantly more important. Of these only the last is included in the fluency descriptors of the two tests considered, with CEFR describing "a natural colloquial flow". It seems, then, that the Italy-based group view fluency more in its broad or lay sense, that is, as denoting global proficiency.

Regarding qualifications, the analyses show a significant correlation between teaching qualifications and *automaticity*, thus suggesting that teacher training increases teacher awareness of cognitive processing. High academic qualifications, on the other hand, reveal curious negative correlations with the two phonological variables *accent* and *native-like rhythm*. This is contrasted by a medium positive correlation between the two Phonology variables and years of experience. It seems that the more academically qualified participants either consider phonology and fluency as unrelated, or they are less interested in phonology in general, while teaching experience seems to increase perceptions of phonology's value as an indicator of fluency.

5.4. Towards a definition

Freed's (2000) study, which used people with no training or experience in language teaching as judges, confirmed that the "popular notion of fluency" (p. 262) includes elements of the narrow sense but also grammar, vocabulary and accent. Although interesting, this reaffirmation of the broad sense does not help us define and operationalise fluency for learning, teaching and testing purposes. Kormos and Dénes'

(2004) study, which used university teachers as raters, produced a narrower definition of fluency, concluding that it is a combination of speed, pace, smoothness and accuracy. Therefore, they too include accuracy in the equation, and as reported in the study, to a lesser degree lexical range. Derwing et al. (2004) raise the question of whether it is possible to ignore grammatical errors, non-nativelike lexical choices and other (non-temporal) variables when rating fluency. However, the raters in their study are again novices in the field, while the focus of the current study is on the ratings of language specialists. The complex interrelatedness of complexity, accuracy and fluency has been a recurrent theme in SLA literature (Housen & Kuiken, 2009), and it has been amply demonstrated how complexity and accuracy interact with fluency (Tavakoli et al., 2015), but if fluency is to be seen as a construct, distinctions between them are necessary.

In line with Freed's (2000) conclusion, the majority of the participants in this study believe that perceptions of fluency are subjective, and certainly if fluency is interpreted in the broad sense, where so many variables play a role, this is highly likely. However, the frequently expressed view that perceptions depend on whether the listeners are language specialists or not is crucial. A central tenet in this study is that, although there will always be a degree of subjectivity when rating speech in real time, this can be limited if clear definitions of the construct are produced and diffused. As Chambers (1997) concluded, "the concept of fluency is confused, multi-layered and therefore needs to be defined specifically. Otherwise the validity of the judgments made by assessors is seriously in question" (p.543).

Taking into account the study findings and the literature on the subject, a definition of fluency would encompass all three of Segalowitz's (2010) interrelated views of fluency and would include:

- adjectives such as *effortless, automatic, smooth, flowing*
- specific utterance variables such as pauses and hesitations
- coherence or "performative competence"
- chunking of language with effective sentence stress and intonation

The existing IELTS and CEFR descriptors are an excellent point of departure and already reflect Segalowitz's three-way definition (Appendix 3). However, it is the

recommendation of this study that they could be amplified to incorporate the elements above.

5.5. Implications for pedagogy

In line with de Bot's (1992) description of L2 speech, the TPs identified poor cognitive processing and lack of linguistic knowledge as causes of L2 disfluency. These are both likely to benefit from increased exposure to the L2, for example by studying in a country where the L2 is spoken (Freed, 2000; Lennon, 1990; DeKeyser, 2001), and from opportunities to practice to improve automatization. However, an approach to processing difficulties is also offered by item-based learning. According to Pawley and Syder (1983), formulaic sequences involve single-step memory retrieval, thus freeing up the working memory, which would otherwise be occupied in formulating and articulating, to focus on conceptualising the message. This diversion of the attentional resources to planning should enhance coherence. Moreover, according to Ellis' (1996) view that the phonological memory remembers language in phonological strings or chunks, the rhythm of speech also benefits. As Tavakoli (2011) demonstrates, learners do not usually pause in the middle of formulaic sequences. Therefore, if learners incorporate formulaic sequences into their speech, length of run, pausing patterns, rhythm, and coherence could all improve. Derwing et al (2004) advocate focusing on formulaic sequences from beginner level.

There is evidence that the introduction of learning strategies such as phrases to buy time (Dörnyei & Kormos, 1998) also helps cognitive processing, and activities that involve pre-task planning are likely to improve coherence. Tavakoli et al.'s (2015) study demonstrates that strategy training and awareness training of pausing, hesitation and repair can lead to improved performance in these areas. Moreover, practice in paraphrasing will limit disfluency when the speaker's linguistic knowledge is inadequate.

As regards the importance of the rhythm of speech for fluency, Derwing et al (2004) also suggest work on appropriate pause placement from early stages of learning may be beneficial. Furthermore, it is the researcher's belief that in-class analysis of good models of sentence stress and intonation, followed by occasional shadow reading improves learner performance. Recording the learners and subsequent self-assessment

may help them gain confidence and self-awareness, the latter being important for learner autonomy (Préfontaine, 2013). Further research to investigate this would be interesting.

These briefly summarised teaching ideas are intended to illustrate how clear notions regarding which variables are involved in fluency and their interrelatedness can help teachers and materials writers develop an approach to help their learners.

5.6. Limitations of the study

The aim of the study was to discover what university language teachers consider important when rating spoken fluency. However, teachers from only one university in Italy were consulted. In order to verify the validity of the findings, a larger sample from a number of different universities in both countries should be found, especially regarding the quantitative data. This would give the researcher greater confidence to generalise.

Regarding the list of variables proposed in the quantitative section of the questionnaire, some improvements could be made. The Phonology group was weak with only two variables, as revealed by the Cronbach's alpha coefficient, and the absence of intonation meant that it was not possible to carry out a validity check to substantiate the considerable importance given to this variable in the qualitative data. *Intonation* could usefully replace *colloquialness*, which was not clearly defined and was already sufficiently covered by *communicative competence* and *ready-made expressions/idiomatic language*.

Finally, the recordings included only samples of monologic speech. Future studies using dialogic samples would generate further and richer reflections on fluency.

6. Conclusion

The aim of this exploratory study was to investigate what native-speaker university teachers based in Italy and the UK believe to be the factors that most contribute to perceptions of fluency. The theoretical framework was provided by Segalowitz's cognitive approach to the construct, which he sees as three interrelated aspects: cognitive, utterance and perceived fluency.

A quantitative and qualitative questionnaire was sent to English language teachers at the University of Florence in Italy and three UK universities. The forty-eight responses were analysed to see the level of consensus concerning the key factors, to identify the prevailing variables, and to investigate the extent to which geographical location, training and experience seem to influence them. The mixed method framework of the study allowed the researcher to check the validity of the quantitative answers against the open-ended qualitative responses.

It was found that the variables most popularly perceived as influential in rating fluency are those that indicate that the cognitive processing is working smoothly (automaticity, effortlessness and coherence) and a cluster of utterance variables that provide surface evidence of good cognitive processing e.g. pausing, hesitations, and repair. Reservations were expressed about speed. Particular emphasis was attributed to length of run, which was identified in the participants' words as "chunking" or "thought groups". They also gave considerable importance to rhythm and intonation, and this was discussed in relation to the perceived importance of chunking utterances. It was found that all three of Segalowitz's views of fluency were represented in the data, and that the main causes of disfluency are inefficient cognitive processing, lack of linguistic knowledge and lack of confidence.

Differences in perceptions were found according to geographic location, with the Italy-based colleagues attributing greater importance to variables related to complexity, global aspects of speech and accent. They revealed a broader interpretation of fluency than their UK-based colleagues, who were more influenced by factors relating to cognitive and utterance fluency. It was tentatively suggested that this might be due to the fact that the UK-based teachers have undergone more training and are therefore

more informed about test benchmarks. Further studies are required to verify this, however, as teachers from only one university in Italy were consulted.

One of the main reasons for the elusiveness of the term fluency is the confusion over what Lennon (1990) called the broad sense of the term, and its narrow sense, and the overlap between the two. It is the interpretation of fluency in its narrow sense that interests SLA researchers and language testers, and given that test scores can influence learners' futures, it is important that there is consensus between them and that teachers and learners are informed. If the variables are clear, the construct can be operationalised and measured. Benchmarks for rating purposes can be discussed and circulated to minimise the subjective element in assessment. A clear definition will also help teachers and materials writers develop a methodology and create materials to improve learner fluency.

Future research is recommended to explore the role of coherence, and how chunking, with its consequences for rhythm and intonation, influences listeners' perceptions of fluency. As regards methodology, materials and activities to improve fluency that focus on the variables discussed could usefully be designed and tested. Greater emphasis to enhancing student awareness of what contributes to fluency should be given, and more attention could be paid to increasing their confidence. Finally, learners should be encouraged to seek out opportunities to interact in the L2. Effective and focussed teaching can guide and inspire, but learner motivation to practise is also fundamental.

(15,318 words)

References

- Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51, 355-365.
- Boers, F., Demecheleer, M., Eyckmans, J., Kappel, J & Stengers, H. (2006). Formulaic sequences and perceived oral proficiency: putting a lexical approach to the test. *Language Teaching Research*, 10: 245–261.
- Bosker, H. R., De Jong, N. H., Pinget, A-F., Quené, H., & Sanders, T. (2012). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30, 159–175.
- Brumfit, C. (1984). Accuracy and fluency: the basic polarity. In H. Riggenbach (Ed.), *Perspectives on Fluency*. (pp. 61-73). Michigan: University of Michigan Press
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- De Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics*, 13, 1-24.
- Chambers, F. (1997) What do we mean by fluency? *System*, 25, 535-544.
- De Jong, N. H., Groenhout, R., Hulstijn, J. H., & Schoonen, R. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behaviour. *Applied Psycholinguistics*, 36, 223–243.
- DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). New York, NY: Cambridge University Press.
- Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31, 533-557.
- Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2004). Second language fluency: judgments on different tasks. *Language Learning*, 54, 655-679.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.

- Dörnyei, Z. (2010). *Questionnaires in second language research: construction, administration, and processing*. London: Routledge.
- Dörnyei, Z. & Kormos, J. (1998). Problem-solving mechanisms in L2 communication. A Psycholinguistic perspective. *Studies in Second Language Acquisition*, 20, 349–385.
- Ellis, N.C. (1996). Sequencing in SLA: phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*. 18, 91-126.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Fillmore, C. (1979). On fluency. In C. Fillmore, D. Kempler, & W. S.-W. Yang (Eds.) *Individual differences in language ability and language behavior*. (pp. 85-101). New York: Academic Press.
- Freed, B. F. (2000). Is fluency, like beauty, in the eyes of the beholder? In H. Riggenbach (Ed.), *Perspectives on Fluency*. (pp. 243-265). Michigan: University of Michigan Press.
- Fulcher, G. (2003). *Testing second language speaking*. Harlow: Pearson Education.
- Gass, S. M. & Mackey, A. (2000) *Stimulated recall methodology in second language research*. Mahwah; London: Lawrence Erlbaum Associates, 2000.
- Housen, A. & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30, 461-473.
- IELTS speaking band descriptors. Retrieved 26th June 2015 from: http://takeielts.britishcouncil.org/sites/default/files/IELTS_Speaking_band_descriptors.pdf
- Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: temporal measures and stimulated recall. *Language Learning*, 64, 809-854.
- Koponen, M., & Riggenbach, H. (2000). Overview: varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on Fluency*. (pp. 102-127). Michigan: University of Michigan Press.
- Kormos, J. & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145-164.

- Kormos, J. (2006). *Speech production and second language acquisition*. London: Lawrence Erlbaum Associates.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387-417.
- Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggensbach (Ed.), *Perspectives on Fluency*. (pp. 102-127). Michigan: University of Michigan Press.
- Levelt W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1999). Producing spoken language: a blueprint of the speaker. In C.M. Brown and P. Hagoort (Eds.). *The neurocognition of language*. Oxford: Oxford University Press, 83-122.
- Nattinger, J. R. & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Pawley, A. & Syder, F.H., (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J.C. Richards & R.W. Schmidt (Eds.), *Language and communication*. (pp. 191-225) London: Longman.
- Pinget, A-F., Bosker, H. R., Quené, H., & De Jong, N. H. (2014). Native speakers' perceptions of fluency and accent in L2 speech. *Language Testing*, 31, 349-365.
- Préfontaine, Y. (2013). Perceptions of French fluency in second language speech production. *The Canadian Modern Language Review*, 69, 324-348.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics*, 22, 27-55.
- Roach, P. (2009). *English phonetics and phonology. A practical course*. Cambridge: Cambridge University Press.
- Segalowitz, H. (2007). Access fluidity, attention control, and the acquisition of fluency in a second language. *TESOL Quarterly*, 41, 181-186.
- Segalowitz, N. (2010). *Cognitive bases for second language fluency*. London: Routledge.
- Schegloff, E.A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361-82.

- Schmidt, R., (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357–385.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Tavakoli, P. (2009). Investigating task difficulty: learners' and teachers' perceptions. *International Journal of Applied Linguistics*, 19, 1-25.
- Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, 65, 71–79.
- Tavakoli, P. & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.). *Planning and task performance in a second language* (pp. 239–277). Amsterdam: Benjamins.
- Tavakoli, P., & Foster, P. (2008). Task design and second language performance: the effect of narrative type on learner output. *Language Learning*, 58, 439-473.
- Tavakoli, P., Campbell, C., & McCormack, J. (2015). Development of speech fluency over a short period of time: effects of pedagogic intervention. *Tesol Quarterly*, 0, pp. 1-25.
- Thornbury, S. (2005). *How to teach speaking*. Essex: Longman.
- Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggensbach (Ed.), *Perspectives on Fluency*. (pp. 102-127). Ann Arbor, MI: University of Michigan Press.
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: a case study. *Canadian Journal of Applied Linguistics*, 12: 39-57.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wright, C. (2014). *Developing academic interactional competence: listening and speaking strategies in EAP settings*. Invited presentation, Swansea University Language Research Centre.

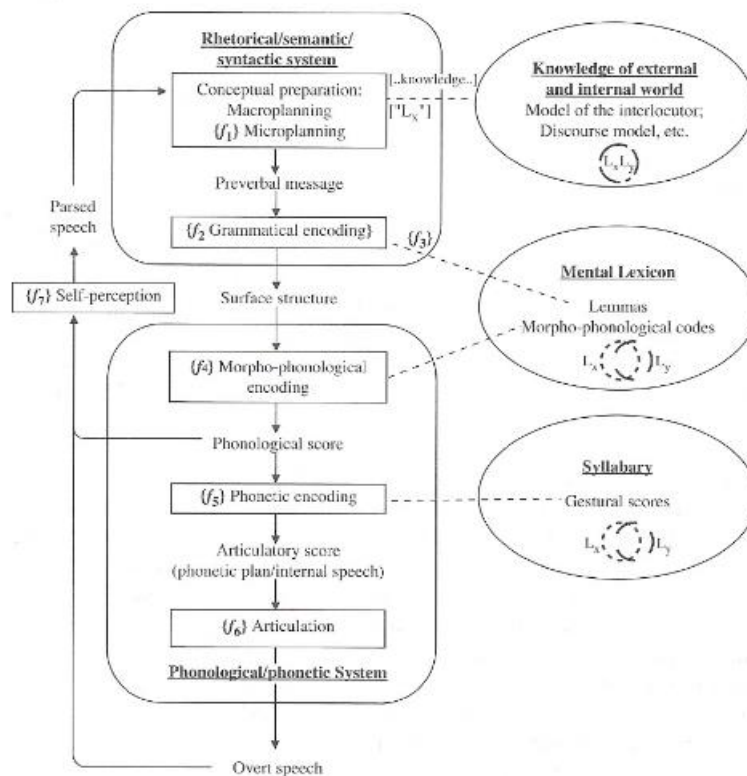
Appendix 1

Segalowitz (2010, p. 9).

Model of the L2 speaker, adapted from Levelt's 1999 blueprint of the monolingual speaker and incorporating De Bot's (1992) points regarding the bilingual speaker.

The dotted and dashed circles show how information pertinent to L1 and L2 is related i.e. partially distinct and undifferentiated systems.

{f} indicates the critical points for underlying processing difficulties related to L2 disfluencies



Appendix 2

Kormos (2006, p. 163). An overview of measures of fluency (retyped).

<i>Measure</i>	<i>Definition</i>
Speech rate	The total number of syllables produced in a given speech sample divided by the amount of total time required to produce the sample (including pause time), expressed in seconds. This figure is then multiplied by sixty to give a figure expressed in syllables per minute.
Articulation rate	The total number of syllables produced in a given speech sample divided by the amount of time taken to produce them in seconds, which is then multiplied by sixty. Unlike in the calculation of speech rate, pause time is excluded. Articulation rate is expressed as the mean number of syllables produced per minute over the total amount of time spent speaking when producing the speech sample.
Phonation-time ratio	The percentage of time spent speaking as a percentage proportion of the time taken to produce the speech sample.
Mean length of runs	An average number of syllables produced in utterances between pauses of 0.25 seconds and above.
The number of silent pauses per minute	The total number of pauses over 0.2 sec divided by the total amount of time spent speaking expressed in seconds and multiplied by 60.
The mean length of pauses	The total length of pauses above 0.2 seconds divided by the total number of pauses above 0.2 seconds.
The number of filled pauses per minute	The total number of filled pauses such as uhm, er, mm divided by the total amount of time expressed in seconds and multiplied by 60.
The number of disfluencies per minute	The total number of disfluencies such as repetitions, restarts and repairs are divided by the total amount of time expressed in seconds and multiplied by 60.
Pace	The number of stressed words per minute.
Space	The proportion of stressed words to the total number of words.

Appendix 3

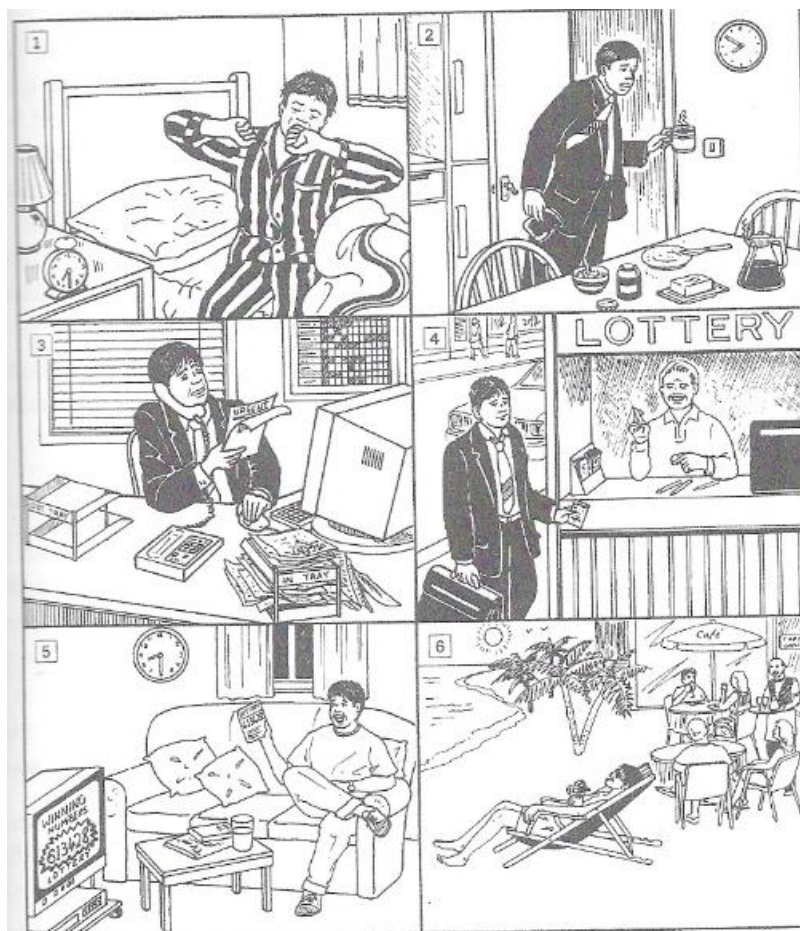
A sample of the fluency descriptors for CEFR and IELTS grouped according to Segalowitz's 3-way distinction of fluency. The third section groups a sample of answers to questions 4 and 5 of the current study's questionnaire in the same categories. The bold has been added by the author to highlight significant words.

	COGNITIVE FLUENCY	UTTERANCE FLUENCY	PERCEIVED FLUENCY
IELTS	<i>any hesitation is content-related rather than to find words or grammar; may demonstrate language-related hesitation at times</i>	<i>only rare repetition or self-correction; occasional repetition, self-correction or hesitation; noticeable pauses and may speak slowly, with frequent repetition and self-correction; pauses lengthily before most words</i>	<i>without noticeable effort</i>
CEFR	<i>Only a conceptually difficult subject can hinder a natural, smooth flow of language; he/she can be hesitant as he or she searches for patterns and expressions; pausing for grammatical and lexical repair is very evident; much pausing to search for expressions, to articulate less familiar words, and to repair communication</i>	<i>...with a natural colloquial flow; ...with a fairly even tempo; ..there are few noticeably long pauses; can be hesitant..; pauses, false starts and reformulation are very evident</i>	<i>...with a natural colloquial flow; avoiding or backtracking around any difficulty so smoothly that the interlocutor is hardly aware of it; Can express himself/herself..... almost effortlessly</i>
CURRENT STUDY	<i>few hesitations caused by speaker searching for lexis or grammar; minor pauses as she searched for words; seems to be able to think and talk at the same time - processing grammar and vocabulary do not get in the way of producing the next utterance</i>	<i>smoother, natural rate; stretches of speech delivered without pauses; flow of ideas without frequent pauses and hesitations; moving ahead instead of correcting</i>	<i>smoother, natural rate; any repair is done unobtrusively; fairly effortlessly</i>

Appendix 4

Picture Story

The six pictures below show a story. Think for a minute about how you are going to tell it, and then tell it in the past tense. Try to make it as interesting as possible.



N.B. Appendices 5 and 6 are the ethics forms and have been removed to preserve the anonymity of the dissertation for the British Council competition.

Appendix 7

Further excerpts from open-ended TP answers to questions 4,5,6 and 7. The organisation follows the sections in chapter 4.

4.3.2.1. Fluency Variables

4.3.2.1.1. Utterance fluency

TP40: pauses seemed to be in appropriate places

TP38: doesn't stop to self-correct when errors are made

Reservations about speed fluency:

TP10: it should not be too fast or too slow

TP40: Speed in itself is not conclusive proof of fluency

Chunking/runs

TP14: provides appropriate chunks by breaking language up into thought groups

TP42: fewer pauses and more chunks of language

Flow

TP12: reasonable flow

TP8: flowing speech without frequent pauses and hesitations

4.3.2.1.2. Cognitive fluency

TP11: more searching for content rather than words or word forms

TP25: few hesitations caused by speaker searching for lexis or grammar

4.3.2.1.3. Perceived fluency

TP21: sounded more automatic

TP12: more ease with the language

4.3.2.2. Complexity Variables

TP13: fairly accurate grammatically and lexically

TP35: fewer grammatical mistakes

TP6: she is held back by her limited lexis

4.3.2.3. Phonology Variables

TP14: her accent and rhythm of speech were better than the others

TP5: use of intonation aids understanding and engages

TP38: intonation also weighed in on my decision

4.3.2.4. Global Aspects

TP42: a good overall communicative performance

TP19: overall comprehensibility is the most important factor

Coherence

TP27: I understood her sequence of ideas

TP20: signalling and narrative coherence

TP16: meaningful, coherent message

Delivery

TP11: Very often listeners react more favourably to speakers with a pleasant voice

TP38: contains natural prosodic features, including (for example) pausing before key information

4.3.3. The causes of disfluency

4.3.3.1. Cognitive processing

TP 13: The tendency to concentrate on accuracy at the cost of losing the thread

TP25: mental processing

TP37: pauses to think of the next word or plan the next chunk

Lack of practice

TP24: immersion and thinking in the 2nd language are key factors in fluency

TP20: lack of oral/aural exposure

L1-L2 interference

TP17: Interference with 1st language

TP7: greater distance between L1 and L2

4.3.3.2. Linguistic complexity

TP38: inadequate lexical range

TP32: the speaker has not mastered the basic grammar structures

4.3.3.3. Personality and affective factors

TP10: fear of making mistakes

TP21: an animated personality can get you a long way

4.3.4. The subjective element of perceptions of fluency

4.3.4.1. Language specialist versus lay person

TP27: There may be some accordance between L2 teachers, but almost none among the general population

TP37: I think that professionally, courses and assessments tend to decide what this term means, but, by defining it, they become a standardised 'beholder' (which is, nonetheless, a beholder).

TP10: Language tutors and teachers might judge fluency as the result of specific factors whereas other people might just associate fluency with the process of getting ideas across efficiently.

4.3.4.2. Listener preferences

TP6: A good accent and rhythm of speech can give an impression of fluency

TP46: voice quality, especially pitch, contributes to a listener's perception of fluency

TP9: the listener's willingness and motivation for understanding is a factor

TP21: I think if you are interested in a topic you may perceive someone to be more/less fluent than someone who is not interested

The listener's background

TP29: Speech fluency depends on listeners' experience of speakers of English as a second language. If they know the variety, they may overlook some problems

TP13: it often depends on how much exposure you have had to that type of pronunciation

