CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

This chapter reviews relevant literature pertaining to the study. It is subdivided into two sections, consisting of Section 2.1 on Syllable and Section 2.2 on Stress. Subsumed under Section 2.1 are discussions on the definition of the syllable, its constituents (Section 2.1.1), how a word is syllabified (Section 2.1.2) and the features superimposed on the syllable, known as suprasegmentals (Section 2.1.3). Section 2.2 begins with the definition of stress, followed by a discussion on lexical stress (Section 2.2.1), sentence stress (Section 2.2.2) and their functions. This is followed by a discussion on measuring stress (Section 2.2.3).

2.1 Syllable

The syllable does not have an agreed phonetic definition among phoneticians as the theories are not entirely adequate (Ladefoged, 2006). Two of such theories, firstly, defines a syllable in terms of properties of sounds, such as sonority or prominence, and secondly, posits that a syllable is the smallest unit possible in the organisation of the sounds of an utterance.

The first theory on the definition of the syllable is based on the assumption that "peaks of syllabicity coincide with peaks of sonority" where sonority is the loudness of a sound compared with other sounds with the same pitch, stress and length (Ladefoged, 2006, p.

240). According to Ladefoged (2006, p. 239), the "sonority of a sound can be estimated from measurements of the acoustic intensity of a group of sounds that have been spoken on comparable pitches and with comparable degrees of length and stress". He estimates that low vowels have greater sonority than high vowels and that nasals have slightly less sonority than high vowels, but greater sonority than a voiced fricative. Voiced stops and all voiceless sounds have very little sonority. This estimation follows the sonority scale of least to most sonorous: Obstruents – Nasals – Liquids – Glides – Vowels (high – mid – low) (Gussenhoven and Jacobs, 1998). This sonority scale is known as the Sonority Sequencing Generalisation, that is, sonority increases towards the syllable peak and decreases towards the syllable margins (Selkirk, 1982). According to Gussenhoven and Jacobs (1998, p. 152), in a number of syllable 'laws', there is a regularity, and this is captured by the sonority profile, where "the sonority of a syllable increases from the beginning of the syllable onwards, and decreases from the beginning of the peak onwards". For example, in a word like *blade*, the sonority increases at /bl/, reaches its peak at /eɪ/ and decreases at /d/.

However, the sonority theory is "inadequate" (Ladefoged, 2006, p. 41) for a number of reasons. Firstly, the acoustic intensity of different sounds may vary among different speakers, for example, one speaker may pronounce /i/ with a sonority that is greater than /l/, whereas another may not; these two sounds have about the same sonority (Ladefoged, 2006, p. 239). Secondly, one syllable could have two peaks of sonority, for example, in the word *spa*, the peaks are in the first and last segments (Ladefoged, 2006). The syllable peak is the syllable which is highest in the sonority hierarchy. There is also an inclination

for the sonority of a successive segment to be higher than the preceding one, for instance, in the word *cry*, /k/, being an obstruent, has a lower sonority than /r/, which is a liquid. Therefore, clear peaks of sonority may not determine the number of syllables in every case. In view of this, Ladefoged (2006, p. 241) suggests that it is easier to say that "syllables are marked not by peaks in sonority but by peaks in prominence".

A sound is prominent when it is louder than other sounds around it. Ladefoged (2006, p. 241) explains that "a sound is prominent because it forms the peak of a syllable; it is syllabic because it is prominent". In a similar vein, Ashby and Maidment (2005, p. 145) explain, a syllable can be seen as "one pulse of speech", always containing "one loud or prominent part... and optionally have consonant sounds preceding or following the vowel". Ladefoged (2006, p. 241) further explains that the "relative prominence of two sounds depends in part on what their relative sonority would have been if they had the same length, stress and pitch, but it also depends in part on their actual stress, length and pitch". However, he adds, this theory is also inadequate, as "one cannot state a procedure for combining sonority, length, stress and pitch so as to form prominence" so the prominence of a sounds cannot be measured (Ladefoged, 2006, p. 241). In other words, a syllable cannot be defined by a peak of prominence as the prominence of a sound cannot be measured.

Although the prominence of a sound cannot be measured, it is clear that syllables in an utterance have different degrees of prominence. A syllable is more prominent when it is louder and more salient than other syllables. According to Bolinger (1958), there are at

least five different levels of prominence in English syllables and these are displayed in Figure 2.1:

```
                          syllable
                         /        \
              full (strong)       reduced (weak)
               /        \
         stressed      unstressed
          /      \
    accented    unaccented
     /     \
nuclear   pre-nuclear
```
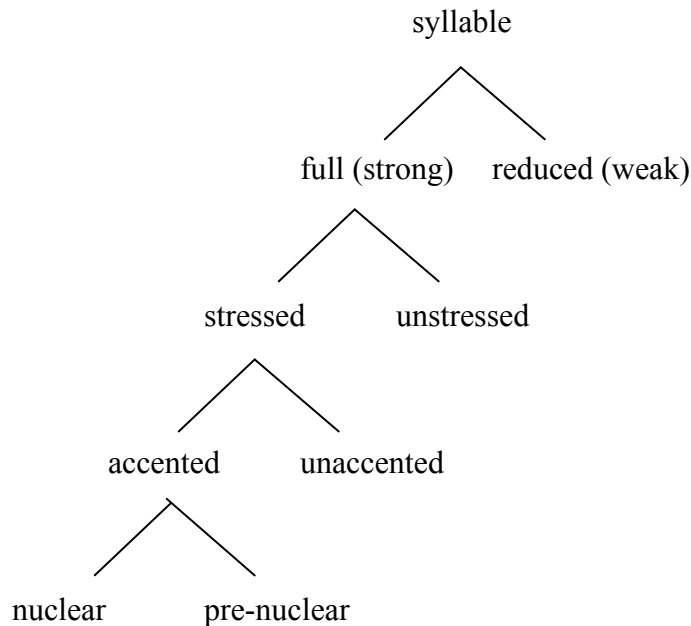
Figure 2.1: Prominence Hierarchy for English Syllables (adapted from Bolinger, 1958)

In Figure 2.1, a syllable can be full or reduced depending on its segments or the sounds it contains. A syllable is reduced if it has a weak vowel /i/, /u/ or /ə/ or a syllabic consonant, /l/, /m/, /n/ or /r/. Reduced vowels are less prominent than full vowels. A syllable is full if it contains any vowel other than a weak vowel. More prominent than others are tense vowels and diphthongs. In this case, it may be stressed or unstressed depending on whether it carries additional marks of prominence, such as timing and loudness. Thus, a stressed syllable would be pronounced longer, louder or with a noticeable difference in pitch, for instance, the syllable *trans-* in *transfer* is longer than the syllable *-fer* in the noun, but shorter in the verb (Knowles, 1987).

However, Knowles (1987, p. 104) contends that not all accented syllables are lengthened as "the duration of the accented syllable depends on its weight, and on the number of trailing syllables". Not all syllables are long as some may be short and clipped, for example, in the word *happen* (Knowles, 1987, p. 104). In *happen*, the first vowel, /æ/, moving on to the trailing syllable is almost immediate, which is typical of a light syllable. According to Knowles (1987, p. 104), a light syllable is "typically short, and has one of the vowels /i, e, a, o, ʌ, u/ immediately before the segment at the syllable boundary". In *transfer*, the first vowel, /a:/, is a heavy vowel, as it can be lingered on and then on /n/ before moving on to the trailing syllable, typical of a heavy syllable. According to Knowles (1987, p. 104), a heavy syllable is "typically long [in duration], and either contains a heavy vowel or… has an approximant in the closing sequence before the segment at the syllable boundary". Thus, in the case of *transfer*, /a:/ is a heavy syllable. The duration of a heavy syllable is longer before a voiced consonant than before a voiceless one, for example, the /n/ varies in the words *renting* and *rending* (Knowles, 1987, p. 104). Thus, the first syllable can be expected to be longer in *rending*, compared to *renting*.

In the last part of Figure 2.1, it was shown that syllabic segments are more prominent than non-syllabic segments. Accented vowels aligned with tonal target are more prominent, while the last accent of a phrase is more prominent than earlier ones, also known as the nuclear accent or the tonic accent. For example, the word *photographic* /ˌfəʊtə'græfɪk/ have prominences on the first and third syllables, but only the third will

The second point of view on defining a syllable is that syllables are the smallest unit of speech. Ladefoged (2001, p. 14) explains that in speech, phonemes exist as independent units of vowels and consonants. However, in continuous speech, they rarely exist independently. Speech sounds are sometimes used alone, e.g. in exclamation *Oh!*, but speech is not the succession of independent sounds. For example, to say *pie*, the lips are closed for the /p/, and the tongue is lowering for the beginning of /aɪ/. While the lips are opening to release the burst, the tongue is fronting and elevating for the off-glide of the diphthong. This coarticulation produces a unit called the syllable. At the same time, there is a delay in voicing after the release of /p/ resulting in an aspirated /p/.

Thus, Ashby and Maidment (2005, p. 145) explain that the syllable can also be "a phonological unit – that is, as a grouping of sounds in accordance with certain rules, which differ from language to language". This is also maintained by Pulgram (1970, p. 65) who defines the syllable as "a linguistic unit of the figura type, a segment of the section, which contains one vowel nucleus and whose phonological boundaries, which may be but are not always necessarily signalled phonetically, are determined by a general set of phonological-phonotactic rules of syllabification (see 2.1.1) that are applied to the specific phonotactics of a given language".

In sum, there are two main takes on the notion of a syllable. According to McMahon (2002, p. 105), most phonologists accept the universal syllable template (see Figure 2.2) described in the following subsection.

2.1.1 Constituents of the Syllable

The syllable is not just a combination of vowels and consonants but also of onset and rhyme. Ladefoged (2006, p. 242) explains that the "rhyming part of a syllable consists of the vowel and any consonants that come after it" while the onset of the syllable is formed by "any consonants before the rhyme". Pulgram (1970, p. 40) calls this vowel the nucleus and states that "there are as many syllables in an utterance as there are nuclei". This means that the nucleus is the most important and only compulsory part of the syllable, for example, in words like *eye* and *hour*. The nucleus generally contains a vowel, but if no vowel is available, certain consonants can become nuclear, playing the part of a vowel (McMahon, 2002, p. 105). In rhotic accents in English, /l/, /m/, /n/ and /r/ can play this part. Examples include the words *bottle*, *bottom*, *button* and *butter*, where in each case, the second syllable consists only of nuclear or syllabic /l/, /m/, /n/ and /r/, for example, /bɒt/ + /l/, /bɒt/ + /m/, /bʌt/ + /n/ and /bʌt/ + /r/ (McMahon, 2002, p. 105).

Rhyme can be divided into the nucleus and the coda, whereby the nucleus is the "vocalic part and the coda consists of any final consonants" (Ladefoged, 2006, p. 242) (see Figure 2.2). Ladefoged (2006, p. 242) provides the examples of the words *I* and *owe*, which consist of a one-rhyme-single syllable, which is also the nucleus, neither comprising an onset nor a coda, and *splint* and *stripes*, single syllables containing onsets with three

consonants and codas with two consonants. In onsets, the sonority of each segment tends to be higher than the preceding segment, for example, in /fl/, /br/ and /kw/.
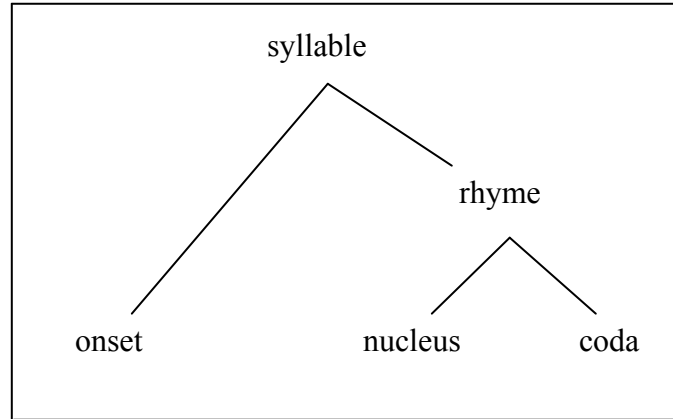

Figure 2.2: The Structure of the Syllable

In explaining the constituents of the syllable, McMahon (2002, p. 106) states that the syllable has its grammar, and the patterns of acceptability differ across languages. For instance, every syllable in Arabic must have an onset, and there are no codas in Hawaiian. The Japanese syllable is always open, that is, ends with a vowel (V), with an option of a single consonant (C). Knowles (1987, p. 68) shows that English, by contrast, tolerates clusters as complex as CCCVCCC as one syllable, as in the word *strengths* /streŋθs/. Although English allow both onsets and codas, and complex clusters as syllables, it has its restrictions, and these are called phonotactic constraints. Examples of these constraints in English are:

1. In a $C_1C_2C_3$ onset, C_1 must be /s/, for example, *strum* /strʌm/. In C_1C_2, C_1 is less sonorous than C_2 except for /s/ clusters.

2. /ŋ/ does not appear in onsets (unlike Malay which has words like *nganga* /ŋaŋa/).

3. /v ð z ʒ/ do not form part of onset clusters, especially /ʒ/, which is rare in onset.

4. /t d θ/ plus /l/ do not form permissible onset clusters.

5. /h/ does not appear in codas (unlike Malay which has words like *boleh* /bɒleh/).

6. Coda clusters of nasal plus oral stop are only acceptable if the two stops share the same place of articulation, for example, *bump* /bʌmp/.

7. /lg/ is not a permissible coda cluster.

(McMahon, 2002, p. 106)

One word can consist of one, two or more syllables. *Pie* is an example of a monosyllable, one-syllable word. Examples of disyllables, two-syllable words, are *table* and *hidden*, while *immediate* and *unsophisticated* are examples of polysyllables, containing more than two syllables (Borden & Harris, 1984, p. 131). Because of the different theories defining a syllable, there are different ways of syllabifying a word, that is, drawing boundaries between syllables, and therefore determining what consonants are in the coda of the first syllable and which consonants are in the onset of the second syllable. This is discussed in the next subsection.

2.1.2 Syllabifying a Word

Syllabifying a word like *unrest* and *undo* is easy if it is based on a morphological criterion (Low & Grabe, 1999; Turk & Shattuck-Hufnagel, 2007), where a word is syllabified based on its morphological make up. Only one division would leave legal syllables, that is, *un + rest* and *un + do*. According to Ladefoged (2006, p. 242), the difficulty of syllabifying a word is in determining "whether a consonant is the coda of

one syllable or the onset of another", for example, the word *happy*. There are three options, that is, /hæ + pi/, /hæp + i/ or the ambisyllabic /hæpi/ or /hæp + pi/.

To solve the problem of syllabifying, Pulgram (1970, p. 48-51) posits five rules. Rule 1 is to establish the boundaries of the sections to be syllabified so they become syllable boundaries. Rule 2 is to divide each section provisorily in such a way that a syllable boundary occurs after every vowel, that is, that all syllables are open. Rule 3 states that after the division (Rule 2), if a syllable has a vowel which does not exist at the end of a word, then a consonant or consonants, from the onset of the subsequent syllable must be moved to this syllable, to provide it with a permissible coda (Pulgram, 1970, p. 48). Rule 4 states that if a syllable cannot be open because the consonant or consonants that would form the onset of the subsequent syllable do not exist in the beginning of a word, then a consonant or consonants, must be moved to the preceding syllable as coda, to reduce the onset to a permissible word-initial shape, thus closing the syllable (Pulgram, 1970, p. 50). The last rule states that the burden of irregularity is borne by the coda not the following onset, if the transfer from the aforesaid rules leads to an inadmissible syllable-final group of consonants (Pulgram, 1970, p. 51).

It is fairly easy to decide on the boundaries of the sections in a word to be syllabified, as stated in Rule 1, for example, the word *table*, would be syllabified as *ta + ble*. Rule 2 above states that all syllables should be made open. A syllable is open when it ends with a vowel, while a closed syllable ends with a consonant. For example, Knowles (1987, p.

108) gives the following examples; the first syllables of *bacon* and *butter* are open, and the first syllables of *banking* and *bungle* are closed.

To explain Rule 3, Pulgram (1970, p. 48) gives the example of the word *establish*. Following Rule 2, the word should be syllabified as *e + sta + blish*, but because /æ/ in the second syllable does not occur syllable-finally in English, it is syllabified as *e + stab + lish*. To make clear Rule 4, Pulgram (1970, p. 50) uses the example of syllabifying the word *employ*. This word cannot be syllabified as *e + mploy* because in English, the word-initial *mpl* does not occur. Therefore, it is syllabified as *em + ploy*.

In syllabifying English words, the process is completed by Rule 4. Pulgram (1970, p. 51), noted, with regard to Rule 5, that "the contingency to create a non-permissible coda arises only rarely, and when it does it generally discloses the presence of an unassimilated loan (a category which includes so-called learned words, which are loans from an earlier stage of the same language)". He gives the example of a Spanish word *transcribir*, which cannot be syllabified *tra + nscribir* or *tran + scribir* because both *nscr* and *scr* are impossible word-initially; hence is syllabified *trans + cribir*, even though the final *ns* is also non-permissible (Pulgram, 1970, p. 51).

Echoing Pulgram's rules of syllabification, Kahn (1976) suggests that the onset "be made as long as it legitimately can be; then form a legitimate coda", in what is known as, the Maximum Onset Principle (MOP). The MOP is a universal principle but different languages allow different kinds of syllables. To explain this, Gussenhoven and Jacobs

(1998, p. 151) give the example of the word *basta*. According to them, languages like Dutch, English and German allow /st/ in the onset, thus syllabifying this word as /bʌ + stʌ/, but Spanish does not, therefore this word is syllabified as /bʌs + tʌ/ in Spanish.

Using the MOP, a word like *feeder* would have /d/ as the onset of the second syllable and not the coda of the first. This is permissible because both /fi:/ and /də/ are legitimate word-initial syllables in English. In the example of the word *happy* earlier, Ladefoged (2006, p. 242) proposes three options, /hæ + pi/, /hæp + i/ or the ambisyllabic /hæpi/ or /hæp + pi/. Following the MOP, the first option is the answer. However, according to Gussenhoven and Jacobs (1998, p. 167), there are two post-MOP syllabification rules, liaison and right capture. Liaison causes a word-final consonant to be in the onset of a following vowel-initial word and it applies across word boundaries, for example, in *five eggs*, the consonant /v/ is both syllable-final and syllable-initial (Gussenhoven & Jacobs, 1998, p. 167). Right capture causes the first consonant of a weak syllable to serve as a (final) coda consonant of a preceding syllable which applies within the domain of the foot (left-dominant in English) (Gussenhoven & Jacobs, 1998, p. 167). In a word like *happy*, the /p/ will be the final consonant of the first syllable and the first consonant of the second syllable (Gussenhoven & Jacobs, 1998, p. 167).

As for the alternation between the clear /l/ and dark /ɫ/, Knowles (1987) confirms that in a word like *yelling*, the /l/ is syllable-initial, but not in *welding*. According to Knowles (1987, p. 108-109), the post-vocalic /l/ in *weld* is dark, and remains so in *welding* where /d/ is on the boundary; in *yell* /l/ is dark, but in *yelling* it becomes clear, as it is the

boundary segment. This statement is supported by McMahon (2002, p. 111) who explains that where there are two syllables, for example, the word *hilly*, the /l/ must be in the onset of the second syllable, "where it automatically surfaces as clear". In making the choice of syllabification for the derivative –*fully*, Low and Grabe (1999) decided that the liquid should form the onset of the following syllable /fʊ/ + /li/, and not form the coda of the first syllable /fʊl/ +/i/ or be ambisyllabic and shared between the syllables /fʊl/ +/li/. This decision was made based on observing the choice of the /l/ produced by their subjects, where "all speakers produced a clear /l/ in all instances where the test item contained the derivational suffix –*fully*" although measurements were not provided in the research (Low & Grabe, 1999, p. 43).

2.1.3 Suprasegmentals

Suprasegmentals are features superimposed on syllables. The suprasegmental or prosodic features of a language are variations larger than individual segments. They are overlaid upon a word, phrase, or sentence. All suprasegmental features are described in relation to other items in the same utterance. As Ladefoged (2001, p. 15) puts it, it is "the relative values of pitch, length, or degree of stress of an item that are significant". The suprasegmentals are a direct bridge to meaning, revealing as they do the attitudes and feelings of the speaker in ways the segmental information alone can never do. Principal suprasegmental features are length, tone, intonation and stress (Ladefoged, 2006, p. 237).

Variations in length can affect single segments and also whole syllables, and these are also considered to be suprasegmental features. While this "general pattern seems to be

preserved across languages, suggesting that it is physiologically expedient … English shows a larger difference ... suggesting a learned overlay" (Borden and Harris, 1984, p. 134). According to Ladefoged (2006, p. 244), in most varieties of English, variations in lengths are completely allophonic. He gives the example of the words *bat* and *bad*. The vowel in the word *bad* tends to be longer than the vowel in the word *bat* because "vowels are always longer before voiced consonants than before voiceless consonants" (Ladefoged, 2006, p. 244). The vowel length is longer before voiced consonants (versus unvoiced) and longer before continuants (versus stops), resulting in a phenomena called pre-fortis clipping. According to Ashby and Maidment (2005), pre-fortis clipping refers to the shortening of a vowel due to the presence of a following fortis consonant, that is strong, long, loud, aspirated and always voiceless.

However, the number of syllables does not determine the time duration of the utterance. According to Borden and Harris (1984, p. 132), there is "a tendency toward alternating stressed and unstressed syllables in English with stressed syllables occurring at fairly regular intervals, a phenomenon called *isochrony*". Baskaran (2005, p. 48) explains that isochrony is where an increase in the number of syllables does not affect the time used for the utterance.

Another suprasegmental feature is tone, which is the overall behaviour of the pitch. A falling tone descends from a higher pitch to a lower pitch. A rising tone ascends from a lower pitch to a higher pitch. These tonal events are more obvious in one-syllable utterances where the meaning is made clear by the direction of the pitch movement at the

end of the utterance, not by grammatical means or additional lexis. For example, the following two words are the same, but (a) is a statement and (b) is a question.

(a) "Yes."

(b) "Yes?"

In utterance, (a) ends with a falling tone while (b) ends with a rising tone.

The pitch pattern in a sentence is known as the intonation. Intonation refers to the changes in pitch when articulating an utterance. It is the melody of speech. It refers to the rise and fall of voice pitch over phrases and sentences. Intonation occurs even in non-tone languages, such as English. According to Borden and Harris (1984, p. 132), intonation is the "perceived changes in fundamental frequency (F0)" that can express differences in attitudes and meaning. For example, the sentence "Today is Tuesday" said with a falling intonation is a declaration, but when said with a rising intonation, it turns the declaration into a question (Borden & Harris, 1984, p. 132).

Falling intonation is used for "various types of statements, communicating a feeling of completeness" and accompanies the decrease in intensity in the breath group. This means that pitch as well as subglottal pressure decreases in time and this decrease is accompanied by both intensity and F0 declines, also known as an "unmarked breath group" (Borden & Harris, 1984, p. 133). Some examples of statements using the falling tone are *That was a good show* (to declare that the show was good), *I like it very much* (to declare that something is liked very much) and *He gave his address* (to confirm that the address was given) (Baskaran, 2005, p. 50).

Pitch rise indicates a question to be answered with yes or no, for example, *Is it ready?*, can be answered with a yes or no (Borden and Harris, 1984, p. 133). It may also indicate that a sentence is incomplete, for example, *As I think about it...* (Borden and Harris, 1984, p. 133). Baskaran (2005, p.50) notes that the rising tone "can be used for statements or when an utterance is not final (inconclusive) and open-ended" and gives the examples of, among others, *I shan't be long (so don't worry)* and *Oh, come on! Don't be upset (it's not the end of the world)*. Pitch rise can also be used by speakers to "hold the floor" during a discussion. If a speaker pauses to think in the middle of a phrase, with the pitch rising, they are less likely to be interrupted than if the pause occurred at a fall in intonation (Borden and Harris, 1984, p. 133).

North American English intonation patterns usually rise during the first part of an utterance and fall at the end (jump up, step down). This is generally true of declarative sentences and of questions which are impossible to answer with yes or no (Borden and Harris, 1984, p. 132). According to Baskaran (2005, p. 51), the rise-fall tone "indicates statements of some conviction, sometimes also used to express a feeling of genuine warmth, surprise or shock" and gives the examples of, among others, *We thought that was marvellous*, *That's wonderful* and *She's an utter fool*.

In other words, intonation marks syntactic contrasts (phrase endings, interrogation versus declaration), changes meaning, and signals attitude and feelings (Borden and Harris, 1984, p. 133). It "is used to group together words which grammatically and semantically form coherent units" in spoken language (Katamba, 1989, p. 244). Excitement, including

some kinds of anger and states of enthusiasm, is often accompanied by large shifts in intonation while calm, subdued states, including some forms of grief, anger, peacefulness, and boredom are characterised by a narrow range of intonation variation. How a person feels is known as often by how he says his message as by the message itself. The meaning of what a person is saying not only "depends on the semantic content of the words as well as the grammatical and logical relationships between the words used in an utterance" but also on the "context of utterance (the linguistic context in which the words are uttered an the background situation known to be relevant by both the speaker and the hearer); ... on the paralinguistic features, for example, gestures, facial expression and voice quality employed; and ... the intonation used by the speaker" (Katamba, 1989, p. 246).

2.2 Stress

Stress is a suprasegmental feature applied to hold syllables and not to individual contoids or vocoids. Stress accompanies words, phrases or sentences to emphasise or focus the important element(s) in them (Baskaran, 2005, p. 46). Therefore, stress can be defined in terms of "something a speaker does in one part of an utterance relative to another" (Ladefoged, 2001, p. 93). Other phonologists describe a stressed syllable as "auditorily more salient than the rest of the elements in the string of which it is a part" (Katamba, 1989, p. 221) and that stress is a "relational feature" as it makes a syllable more prominent and more audible than neighbouring syllables (Ashby & Maidment, 2005, p. 154). According to Katamba (1989, p. 221), the "main phonetic ingredients of stress are pitch, length and loudness". This is echoed by Ashby and Maidment (2005, p. 156), who

state that a stressed syllable is made more prominent and more audible than neighbouring syllables by a combination of the same factors.

Stressed syllables "tend to have higher pitch than their non-stressed counterparts" (Katamba, 1989). Pitch is the speed of vibration of the vocal cords and also relative frequency perception, for example perpetually high-pitched or low-pitched. The speed of vibration is determined by the number of times the vocal folds open and close per second. The more rapid the vibration or the more frequently the vocal folds open and close, the higher the pitch. Women generally have higher pitched voices compared to men, as the frequency of vibration in women's speech varies between 150 and 320 Hz, whilst men's frequency varies between 80 and 210 Hz (Ashby & Maidment, 2005, p. 154). A syllable is perceived as stressed when there is pitch height, or when there is a change in pitch.

Stressed syllables tend to have longer duration. This is due to several reasons. Bertinetto (1977) suggests that stressed syllables are regularly longer than unstressed syllables as unstressed syllables undergo compression. According to Ladefoged (2001, p. 93), a stressed syllable "frequently has a longer vowel than the same vowel would be if it were unstressed". This could be due to the reduction of vowels in unstressed syllables and the full articulation of vowels in stressed syllables. The duration of a stressed syllable also depends on the weight of the syllable, whether it is a light or heavy syllable (Knowles, 1987). A light syllable is typically short and has one of the vowels /i, e, a, o, ʌ, u/ immediately before the segment at the syllable boundary, while a heavy syllable is

typically long and contains either a heavy vowel or an approximant in the closing sequence before the segment at the syllable boundary (Knowles, 1987, p. 104).

Stressed syllables are louder than unstressed syllables, but according to Katamba (1989, p. 221), loudness is a "much less important parameter than pitch or length". According to Ladefoged (2001, p. 93), a stressed syllable is "often, but not always, louder than an unstressed syllable". The amount of air that passes through the vocal cords is known as the amplitude of vibration, producing the corresponding loudness or softness of one's voice (Baskaran, 2005, p. 22). Clark and Yallop (1990, p. 282) note that "loudness seems to be the least salient and least consistent of the three parameters of pitch, duration and loudness" at least when signalling stress.

An increase of the respiratory muscles when a greater amount of air is pushed out of the lungs and a change of laryngeal muscles activity can result in a significant change in pitch which causes variations in stress. Pitch changes due to variations in laryngeal activity can transpire whether stress changes transpire or not. Pitch changes can affect the meaning of the sentence as a whole. For instance, stress, when used for emphasis, can express disdain for children in general, "not that **child**!", or dislike of a particular child "not **that** child". The use of changing F0, perceived as the pitch change or intonation contour of a phrase or sentence is particularly effective in expressing difference in attitude (F0 would increase for the stressed words in the example above) and differences in meaning [for example, declarative into interrogative, etc] (Borden and Harris, 1984, p. 132). Stress and operative words help to tell a story and make an argument. And the

prosodic features alone can signal opposite meanings [for example, admiration or sarcasm on the same word].

Although stress is used to emphasise important elements in words, phrases or sentences, it should not be confused with the term 'emphasis'. According to Mangold (1975, p. 2), the term 'emphasis' is often synonymous with stress, however, "in the last decades the use of the term 'emphasis' as referring to sound [stress] seems to have considerably decreased" and taken over to a certain extent by the term 'prominence'. Katamba (1989, p. 221) describes stress as "primarily a matter of greater auditory prominence". Mangold (1975, p. 2) writes that the prominence of a syllable may be due to strong stress, but it may also be due to other features of pronunciation and particularly to the inherent quality of sounds, to the length of syllables or to intonation.

According to Ladefoged (2001, p. 93), a stressed syllable is produced by "pushing more air out of the lungs in one syllable relative to others" which results in it having a "greater respiratory energy than neighbouring unstressed syllables", being pronounced with "greater amount of energy" than an unstressed syllable. It may also have "an increase in laryngeal activity" (Ladefoged, 2001, p. 93). When a speaker pushes more air from the lungs to stress a syllable, the loudness increase due to the extra activity of the laryngeal muscles. Borden and Harris (1984, p. 132) explain that more articulatory effort is involved in producing the stressed syllable, thus, F0 usually increases for the stressed syllable, and the formants for the stressed vowels reflect articulatory achievement of target positions.

In English, the term 'stress' denotes the prominence of syllables in single words, also known as lexical stress. However, not every word is stressed in a sentence because of sentence stress, also known as tonic stress, rhythmic stress and accent. These two are described in the following subsections.

2.2.1 Lexical Stress

Lexical stress, also known as word stress, governs the pronunciation of words. Knowles (1987, p. 117) states that the stress of a word "depends on whether it belongs to the native 'Germanic' or the borrowed 'Classical' part of the vocabulary". Germanic words are stressed on the first syllable, while in Classical words, the stress position "depends on the weight of the syllables of the word, and partly on the final suffix" (Knowles, 1987, p. 119). The stress position is in the first syllable of Germanic words because "many Germanic words are monosyllables ... so the question where the accent falls does not arise in the first place" (Knowles, 1987, p. 117). The position of stress does not change when a Germanic suffix is added to a word. Knowles (1987, p. 117) gives the examples of *singing* being stressed on the same syllable as *sing* and *friendly* stressed the same as *friend*. There are, however, some exceptions to the Germanic rule. Words that begin with the prefix *a-*, for example, *abroad* and *ashore* are stressed on the second syllable (Knowles, 1987, p. 118) as "this prefix derives historically from the preposition *on*, and like other prepositions before a noun phrase, it is unaccented". Also, according to Knowles (1987, p. 118), verbal prefixes, such as *be-*, *for-* and *with-*, are not stressed, for example, *begin*, *forbid* and *withstand*, "presumably because when the rule first came into operation, the prefixes were not really felt to be part of the verb".

As for words with Classical origin, the rules are more complex, and "are breaking down in modern English, with many irregularities as a result"; for example, while *adjective* is stressed on the first syllable, *objective* is stressed on the second syllable (Knowles, 1987, p. 119). Knowles (1987, p. 119) states that "accent is assigned from right to left, starting at the end of the word, and working towards the beginning" and suggests that the rule of thumb is that "accent falls on heavy syllables: a heavy syllable is accented unless it is immediately followed by another heavy syllable, and a light syllable is accented only before an unaccented syllable". Not counted for are inflectional endings such as *–us*, *–a*, *–um*, *–o* and *–i*, rather like Germanic suffixes, "so that a preceding light syllable is not accented" (Knowles, 1987, p. 119). He gives the example of the words *aroma* and *cinema*, where the former is stressed on the second syllable while the latter, the first syllable (Knowles, 1987, p. 119). In both words, the *–a* is left out of account; in the first word, the heavy *–ro* is stressed while in the second word, the stress is passed on to the first syllable from the light second syllable (Knowles, 1987, p. 119). Another rule for stressing words of Classical origin is not to stress the last syllable of these words (Knowles, 1987, p. 120). Latin scholars would recognise the endings of *liberty*, *personal*, *glorious* and *resident* as stressed syllables, but in English, these words "retain the property characteristic of accented syllables of blocking the accent on the syllable immediately preceding" which means the stress position falls two syllables to the left, so these words are stressed on the first syllable (Knowles, 1987, p. 120).

The main function of word stress in English is to show the syntactic relationships between words and parts of words. It gives distinct form to the words, for example, noun,

verb, and adjective. Knowles (1987, p. 118) explains that "verbs of two syllables of Latin origin are typically accented on the root... while the corresponding nouns are accented on the prefix". Historically, according to Knowles (1987, p. 118), words such as *convict*, *transfer*, *permit*, *progress*, *export*, *survey* and *protest*, were "formerly accented on the root, and that the accent has shifted on to the prefix first in the noun, and later in the verb". In other words, stress is used to show the difference in meaning between two identical words which sound the same but have different grammatical functions, that is, phonologically identical, but morphemically different. There are many noun-verb oppositions, such as 'an insult, to insult'; 'a permit, to permit'; 'an overflow, to overflow'; 'an increase, to increase'. In all these pairs of words, the noun has the stress on the first syllable and the verb has it on the last. The placement of the stress indicates the syntactic function of the word (Ladefoged, 2001, p. 94).

The addition of a suffix to create a new word may result in an adjustment of the stress placement. Some suffixes do not change the place of stress when a suffix is added, for example, *-able*, *-en*, *-ful*, *-ed*, *-er*, *-est*, *-ing*, *-ive* and *-ly*. These are called stress-neutral suffixes. Examples of suffixes that cause a change in the place of stress are, *-ian*, *-ial*, *-al*, *-ic*, *-ical* and *-ity*, are called stress-imposing suffixes (Ashby & Maidment, 2005, p. 159). The difference is illustrated in the following table (see Table 2.1).

Table 2.1: Examples of English Stress-neutral and Stress-imposing Suffixes

Stress-neutral				
-able	comfort	/'kʌmfət/	comfortable	/'kʌmfətəbl/
-en	forgot	/fə'gɒt/	forgotten	/fə'gɒtn/
-ful	wonder	/'wʌndər/	wonderful	/'wʌndəfl/
-ed	interest	/'ɪntrəst/	interested	/'ɪntrəstɪd/
-er	pretty	/'prɪti/	prettier	/'prɪtɪə/
-est	lucky	/'lʌki/	luckiest	/'lʌkɪəst/
-ing	amaze	/ə'meɪz/	amazing	/ə'meɪzɪŋ/
-ive	suggest	/sə'dʒest/	suggestive	/sə'dʒestɪv/
-ly	silent	/'saɪlənt/	silently	/'saɪləntli/
Stress-imposing				
-ian	magic	/'mædʒɪk/	magician	/mə'dʒɪʃən/
-ial	industry	/'ɪndəstri/	industrial	/ɪn'dʌstrɪəl/
-al	accident	/'æksɪdənt/	accidental	/æksɪ'dentl/
-ic	climate	/'klaɪmət/	climatic	/klaɪ'mætɪk/
-ical	hypothesis	/haɪ'pɒθəsɪs/	hypothetical	/haɪpə'θetɪkl/
-ity	electric	/ɪ'lektrɪk/	electricity	/ɪlek'trɪsəti/
Note: Adapted from Ashby and Maidment (2005, p. 160)				

Similar oppositions occur in cases where two words or phrases form compounds, such as 'a walkout, to walk out'; 'a put-on, to put on'; 'a pushover, to push over'. In these cases, there is a stress only on the first element of the compound for the nouns but on both

elements for the verbs. Stress also has a syntactic function in differentiating a compound noun from an adjective followed by a noun. For example, a compound noun 'a hotdog', which is a type of food, and 'a hot dog', which is an overheated dog. Compound nouns have a single stress on the first element, and the adjective plus noun phrases have stresses on both elements (Ladefoged, 2001, p. 94). In English, stress is very important as it is "an essential part of word-shape; words easily become unrecognisable if the stress is wrongly placed" (O'Connor, 1973).

In words that have two or three syllables, also known as polysyllables, it is likely that there will be only one stressed syllable (Ashby & Maidment, 2005, p. 157). Knowles (1987, p. 120) agrees by stating that "polysyllabic words may be long enough for more than one accent to be assigned by the 'Classical' rule". He gives the example of the words *geological* /ˌdʒiə'lɒdʒɪkl/ and *archaeology* /ˌaːki'ɒlədʒi/, among many others, to show that in these two words, the third syllable is stressed, the second syllable unstressed, and another stress falls on the first syllable (Knowles, 1987, p. 120). The stress on the third syllable is more prominent compared to the stress on the first syllable. The loss of accent on the final syllable can occur in related words (Knowles, 1987, p. 120); for example, the words *photographic* /ˌfəʊtə'græfɪk/ and *photograph* /'fəʊtəgraːf/ are not stressed on their final syllables and the *graph* in the latter word loses its accent, sometimes described as 'secondary stress'. However, Knowles (1987, p. 120-121) notes that "it is not the last syllable which is subject to accent loss, but the last morpheme" which means if the final morpheme contains two syllables, for example, –*acy*, –*ary*, –*ory* and –*mony*, both are not stressed. In other words, for words that have many syllables, one

2.2.2 Sentence Stress

Stress is used contrastively in a sentence for emphasis and for differences in meaning. It also expresses emotional attitude – different words are stressed to express different meanings or attitudes. According to Katamba (1989, p. 242), in a sentence, one of the words has a syllable which is more prominent than the rest, and "this is the syllable which has sentence stress – also called tonic stress". According to him, "in unmarked cases, tonic stress goes on the syllable which carries primary word stress in the last lexical item of the tone unit ... often a noun, verb or adjective ... In marked cases, the above rule is superseded by a rule which allows contrastive (or emphatic) stress to fall on grammatical (function) words such as pronouns, prepositions and conjunctions if the speaker wishes, for some communicative reason, to draw attention to such words or alternatively, contrastive stress may occur on non-final lexical items" (Katamba, 1989, p. 242). In unmarked cases, the tonic syllable is underlined in the sentences below:

(a) Hannah is *cooking*.

(b) She tore her *jacket*.

The tonic stress goes on the syllable of the primary word stress in the words *cooking* and *jacket*. In marked cases, the tonic syllable can be in any word, affecting the way the sentence is interpreted.

(a) *Hannah* is cooking. (Hannah is cooking, not some other person.)

Hannah *is* cooking. (Hannah is cooking, not doing something else.)

(b) *She* tore her jacket. (She tore her jacket, not some other person.)

She *tore* her jacket. (She tore her jacket; she did not do something else to the jacket, for example, she did not wear it or wash it.)

She tore *her* jacket. (She tore her own jacket; she did not tear some other person's jacket.)

The examples above show that in unmarked cases, tonic stress makes a syllable in the word containing new information stand out, while in marked cases, this syllable can be in any word the speaker chooses to highlight.

2.2.3 Measuring Stress

Increased effort, intensity, pitch, duration and a change in formant patterns are indicators of a syllable being stressed (Borden & Harris, 1984, p. 132). Intensity is measured in amplitude (dB), pitch movement, or fundamental frequency (F0) and formants are measured in Hertz (Hz), while length or duration is measured in milliseconds (ms). To measure stress in this study, two of these correlates of stress, duration and F0, were chosen based on the following theoretical underpinnings.

According to Katamba (1989), loudness is less important than pitch or length. This means that the most important perceptual cues to stress are F0 and duration, followed by amplitude. According to Roach (1991, p. 86), the most important factor in determining stress is pitch movement (F0). In addition to these, Low and Grabe (1999) used F0 and

duration in their acoustic analysis of lexical stress placement in Singaporean English (SgE) and British English (BrE). As this present study would compare its results with that of Low and Grabe's (1999) results, F0 and duration were measured in this study.

In Low and Grabe's (1999) study, the duration and F0 of polysyllabic words in phrase-final and phrase-medial positions by the two groups of speakers, SgE and BrE, were measured and compared. In measuring the durations of the final and penultimate syllables of test words, Low and Grabe (1999) hypothesised that there would be stress difference between the two groups because of more phrase-final lengthening in SgE, and therefore would appear in phrase-final position, not phrase-medial position. The results, however, did not support the hypothesis, although phrase-final lengthening contributed towards an account of the cross-varietal difference, which might explain why SgE speakers were thought to stress the final syllable in words in citation forms (Low & Grabe, 1999). Low and Grabe (1999) also measured the F0 of the voiced portions of each syllable, both onset and offset, to distinguish stressed and unstressed syllables. Changes in pitch, or F0, signal the occurrence of stress, and pitch can only be present in voiced portions of a syllable. The evidence did not suggest a cross-varietal difference in stress placement, only a higher F0 target in BrE compared to SgE (Low & Grabe, 1999). Low and Grabe (1999) concluded that SgE and BrE do not differ in lexical stress placement in polysyllabic words, but in the acoustic realisation of stress. This conclusion could have been different if duration measurements were taken and F0 measurements were not. Lengthening is a cue to stress, so analysing the duration measurements alone could lead to the assumption that SgE speakers place lexical stress on the final syllable of polysyllabic words.

In their study of phonetic correlates and stress placement in SgE, Lim and Tan (2001) studied the F0, amplitude and duration of utterances produced by the Chinese, Indian and Malay varieties of SgE. They found that the Chinese used intensity to distinguish stressed and unstressed items, while the Indians used F0 and intensity, and the Malays used all three (Lim & Tan, 2001). The Chinese and the Malays were also found not to distinguish between neutral stress and emphatic stress, while the Indians used duration as a cue for emphatic stress (Lim & Tan, 2001). As for stress placement, Lim and Tan (2001) found that the subjects did not place stress on the item meant to be stressed and did not always stress new or contrastive information. Lim and Tan's (2001) research showed that the phonetic cues for stress in SgE differ from BrE and that each variety, Chinese, Indian and Malay, uses different cues for stressed syllables.

In their research on lexical stress contrasts produced by native Mandarin speakers, Zhang, Nissen and Francis (2008) studied the F0 (average), time of F0 peak, the first and second formant frequencies (F1 and F2), duration, intensity and vowel quality of test words to identify the acoustic characteristics of the lexical stress produced. All measurements were done using Praat. They found that Mandarin speakers produced "stressed syllables with significantly higher F0" than American English speakers (Zhang, Nissen & Francis, 2008, p. 4504). In addition, Mandarin speakers were found to produce "F0 peak location significantly earlier in unstressed syllables than that in stressed ones" whereas the difference in F0 peak location between stressed and unstressed syllables for American English speakers was not significant (Zhang, Nissen & Francis, 2008, p. 4505). They also found that Mandarin speakers are not able to manage F0 due to native tonal system

interference. They are also unable to manage vowel quality in an English-like manner due to interference from their native vowel systems. It seems that they have transferred Mandarin stress patterns to achieve English-like stress patterns. It would be of interest to examine if the ten subjects in this study had produced stress contrasts transferred from Mandarin.

Analyses of syllable durations and average intensity showed significant effects of stress produced by both speakers and there were no significant differences between the durations and intensities of stressed syllables or unstressed syllables (Zhang, Nissen & Francis, 2008, p. 4505). This suggests that the acoustic correlates are different according to speaker groups. Zhang, Nissen and Francis (2008, p. 4510) also found that "Mandarin speakers produced English-like vowel reduction in certain unstressed syllables" among other significant differences in F1 and F2 patterns between the two groups of speakers. By measuring the F0, time of F0 peak, F1 and F2, duration, intensity and vowel quality of test words produced by both American and Mandarin speakers, Zhang, Nissen and Francis (2008) were able to make comparisons and finally identify the acoustic characteristics mentioned above.

In their study of the effects of lexical stress patterns on final lengthening in American English words, using acoustic measures, Turk and Shattuck-Hufnagel (2007, p. 449) looked at how an Intonational Phrase boundary affects different phrase-final words and compared final main stress, penultimate main stress and antepenultimate stress. Durational measurements of each syllable, nucleus, onset, coda and rime were taken for

each syllable of the test words. The effects of "varying the phonological nature and intrinsic duration of the final syllable" were also examined by "comparing intrinsically long full-vowel syllables with reduced-vowel final syllables" (Turk & Shattuck-Hufnagel, 2007, p. 449). Where possible, syllabification for the target words was done by morpheme, and otherwise, by following the Maximal Onset Principal described by Pulgram (1970) (Turk & Shattuck-Hufnagel, 2007, p. 451). Their study revealed both similarities and differences of the effects of phrase-final lengthening in American English "compared to earlier studies of German, Dutch, Hebrew, Estonian, and Southern BrE" (Turk & Shattuck-Hufnagel, 2007, p. 464). They found "major lengthening on the final-syllable rime, lesser but reliable lengthening on the main-stress syllable rime, and sporadic minor lengthening of other regions in the final word inconsistent with simple views of boundary-related lengthening", suggesting "a more complex process, involving two different domains and possibly two separate lengthening mechanisms" (Turk & Shattuck-Hufnagel, 2007, p. 464).

Acoustic measurements of stress may allow for a detection of changes to otherwise expected lexical stress placement. Simo Bobda (2010) observed the "variability of word stress in time" in his study of word stress in Cameroon and Nigerian Englishes. He observed that the traditional and newer forms of word stress "are in use today in different varieties of received pronunciation" (Simo Bobda, 2010, p. 61). He gave examples of stress moving within the same word from one position to another, and back to the original position later: *momen'tarily, pri'marily, tempo'rary* to *mo'mentarily, 'primarily, 'temporary*, and back to the original position (Simo Bobda, 2010, p.61).

Another instance is "Guierre's Normal Stress Rule", which is described by Trevian (2007, p. 427) as "initial stressing in two-syllable words and antepenultimate stressing in words of three syllables or more". Over time, it has "generated changes and variants in the words to which they attach". In Trevian's (2007) study of stress-neutral endings in contemporary BrE, a comparison of relevant items was done using Guierre's system of rules which was based on the 12th edition of Daniel Jones' *English Pronouncing Dictionary*, *The Longman Pronunciation Dictionary* by Wells (1990, 2000) and the 15th edition of Jones' *English Pronouncing Dictionary* by Hartman, Roach and Jones (1997). It is found that stress shifts are common in suffixes such as *-ed*, *-ing*, *-ly*, *-atory* and *-able*, mostly brought on by a system of rules presided over by word categories and the quality of the syllable these suffixes attach to (Trevian, 2007). It is concluded that word stress is "subject to variation and changes" in contemporary BrE (Trevian, 2007). Acoustic measurements of stress would be of great use to detect such variation and changes.

2.3 Conclusion

This chapter has presented relevant literature related to the topics central to the study of English word stress. The next chapter presents the research design and a discussion of the research method employed, along with a demonstration and rationale of the procedures involved in the investigation.