Stress and cues to relative prominence in English and French: A perceptual study

Dan Frost

IUT2, Université Pierre Mendès-France, Grenoble, France
frost@iut2.upmf-grenoble.fr

The relative prominence of syllables is essential to the segmentation of speech and therefore a crucial component of language comprehension, acquisition and learning. Incorrect placement and marking of prominence in English by non-native speakers can lead to problems in comprehensibility. Because the English and French phonological systems are so different, especially in the domain of stress, this can cause serious difficulties for many French speakers learning English. Indeed, some authors have posited the existence of 'stress deafness' in certain individuals. I suggest that French and English native speakers listen differently for stress, attributing different importance to the acoustic cues of F0, duration, amplitude and formant structure. This study focuses on the relative importance of these four cues with both English and French stimuli for English and French native speakers, and the results support the hypothesis.

One of the major difficulties when studying the phenomenon of stress stems from the fact that it is both an abstract concept and a concrete phonetic realisation in communicative situations. To make matters worse, terms such as 'stress' and 'accent' have been used variously over the years to refer both to the abstraction and the utterance-level reality (Ladd 2008: 52). As Beckman & Edwards (1994: 8) explain, 'stress is not a paradigmatic specification like tone or vowel quality. Rather it is a syntagmatic structural specification. It is one of the devices that a language can use to set up a hierarchical organization for its utterances'. At the utterance level, a stressed syllable is more prominent than the syllables which surround it, but as the first section of this paper outlines, opinions have differed over what causes this prominence. In this paper, the term 'stress' is used to refer to the concept, and where necessary, the term 'prominence' is used to distinguish the phonetic realisation of a stressed syllable.

French and English, despite sharing much of their vocabulary, are vastly different phonetically and phonologically. This is particularly evident in the domain of stress, where the two languages differ so greatly that many French speakers find it very difficult to perceive prominent and also unstressed syllables in English, which often leads to serious comprehension problems. These difficulties are so important for certain individuals that some French authors have posited the existence of 'stress deafness' with regard to languages such as English, where stress plays a defining role (Dupoux & Peperkamp 1999, Dupoux, Peperkamp & Sebastián-Gallés 2001, Peperkamp & Dupoux 2002).

Syllables are perceived as more or less prominent according to several acoustic cues, and the relative importance of these cues is not the same in French and in English. As relative

Journal of the International Phonetic Association (2011) 41/1
doi:10.1017/S0025100310000253

prominence is not simply an acoustic phenomenon but also a perceptual phenomenon, its perception depends on picking up on these acoustic cues. In the first section of this paper, the nature of stress and prominence in both languages is examined briefly. It is suggested that a reason for 'stress deafness' may be that French and English native speakers may not process these cues in the same way. This is an example of linguistic transfer, a concept discussed briefly in the second section. In order to investigate the possibility of this particular aspect of linguistic transfer, a cross-linguistic perceptual experiment was devised; in the third and principal section of this article, this experiment and its results are presented.

1 Stress

Stress has always seemed to resist all attempts at definition: the closer one investigates the nature of stress, the more difficult it becomes to define. One of the reasons for this is that a stressed (or prominent) syllable is distinguished not only by acoustic features, but it is also a perceptual phenomenon, i.e. in defining it, one must account not just for its production, but also for its perception by the interlocutor. Couper-Kuhlen (1985) defines stress as 'nothing more than the fact that in a succession of spoken syllables or words some will be perceived as more salient or prominent than others' (Couper-Kuhlen 1985: 19). Stress is also therefore a relative and not an absolute feature, and the stressed syllable must be defined as prominent in relation to its surrounding environment. Stress plays a role in three main areas: lexical recognition, syntactic processing and discourse structure (Cutler, Dahan & van Donselaar 1997). However, this study is not concerned with the functions of stress, but in how important the acoustic cues are in distinguishing the prominent syllable at the phonetic level.

1.1 Stress in English

One thing is certain about stress, which is that unusually for human language, it is highly iconic (Pennington 1996: 137), i.e. the acoustic effort involved in marking stress coincides with the relative magnitude of the stress. The prominent syllable is marked by variations in four acoustic cues: fundamental frequency (F0), amplitude, duration and formant structure. The perceptual correlates of these four acoustic cues are respectively: pitch, volume, length and a different timbre to the vowel. Phonologists and phoneticians do not all agree about which of these cues is the most important in English, but most seem to agree that stress in English is usually marked by variations in the fundamental frequency of the prominent syllable, which led Bolinger (1958) to posit the existence of a 'pitch accent' in English. It is important to remember, however, that as prominence is a communicative feature, it may manifest itself in many different ways: the prominent syllable does not necessarily bear a pitch accent, a syllable may be metrically strong without necessarily being stressed, etc. (Ladd 2008: 61).

Stress is a relative phenomenon, and most authors agree that it is marked by several degrees of prominence. There is some disagreement as to how many levels exist in English; for example, according to Pennington (1996: 131–132), between four and six levels suffice for a detailed transcription. Other authors, such as Cruttenden (1986: 21), distinguish four levels: primary stress, secondary stress, tertiary stress and unstressed.

Aside from its form, stress also has different functions. Firstly, in English, stress exists at the level of the word: this is generally referred to as word stress or lexical stress. An example of word stress in English is the word *uni'versity* /junɪˈvɜːsɪti/, where the primary word stress occurs on the third syllable. If any syllable other than the third syllable is pronounced with more prominence than the other syllables, this is considered to be a speech error. The nature and behaviour of stress above the level of the word in English is a great deal more complex. Any syllable, stressed or otherwise, may also receive additional weighting. This additional weighting above the level of the word has sometimes been referred to as 'sentence stress',

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

but it is perhaps more helpful to think of it as focus, as it may be placed on different syllables according to the utterer's discursive intention. Focus 'gives prominence to the syllables that are lexically stressed, primarily by assigning them a pitch accent' (Xu & Xu 2005: 160). Syllables receiving focus will therefore be more prominent than the other stressed syllables. Focus generally occurs on the stressed syllable of the last lexical word (i.e. a verb, noun, adverb or adjective) of the sentence, and in such cases is often referred to as the nucleus, or nuclear stress. With longer sentences composed of several prosodic groups, each prosodic group or tone-unit may also contain a more prominent syllable. In addition to word stress and focus, any stressed syllable may be given ADDITIONAL, or contrastive stress; this type of stress, also known as 'narrow focus' is a language universal and is highly unpredictable (Lado 2008: 216). Although word stress is fixed in English, focus projection will depend on the information structure of the sentence in question. The rhythm of English is therefore composed of a succession of more or less prominent syllables, i.e. unstressed and variously stressed syllables.

1.2 Stress in French

The situation in French is no less complex than in English. Comparing French to other European languages, some authors have even concluded that French is a language without stress. This situation becomes clearer when we accept that although French does not have word stress, there is relative prominence, although to a lesser degree than in English. Rossi (1979) for example, after conducting several experiments on the production and perception of stress, concluded that French was a language without stress, in the sense that stress and intonation in French, both by their nature and by their function, do not constitute two distinct entities (Rossi 1979: 39). However, empirical research has shown that French listners do rely on certain acoustic cues in the rhythm of the language to segment speech (Wenk & Wioland 1982: 196). It is generally accepted, however, that French does have stress. Dahan & Bernard (1996: 342) list final stress, secondary stress and contrastive stress as the three categories of stress in French. Although no real consensus is apparent, most authors do agree that the last syllable of each tone unit in French is marked out from the others.

Stress in French depends on the separation of sentences into prosodic units or tone units, which are given different names by different authors, both in English and in French. Many authors in English simply refer to these prosodic units as 'rhythmic groups' (e.g. Cutler et al. 1997). In fact, two levels of prosodic unit may be distinguished in French. The smaller of the two, a 'stress group' (Di Cristo 1998) or an 'Accentual Phrase' (Jun & Fougeron 1995), may contain several syllables or even words. Di Cristo states that French has a 'rhythmic stress', which is 'regularly assigned to the final full syllable (i.e. not containing a schwa) of the last lexical item of a stress group'. He describes a stress group as 'a prosodic unit containing a stressed syllable preceded by a number of unstressed ones' (Di Cristo 1998: 4). The larger prosodic units, such as 'intonative units' (Di Cristo 1998) or 'Intonation Phrases' (Jun & Fougeron 1995), may contain several of the smaller units. Jun & Fougeron (2000) later add a third unit, the 'Intermediate Phrase'.

In terms of production, the group-final syllable in French is marked most notably by an increase in its duration (Benguerel 1973, Di Cristo 1998, Lacheret-Dujour & Beaugendre 1999: 41, Jun & Fougeron 2000, Astésano 2001), but this may also be in part a contextual effect. As Astésano (2001: 54) states, it may be that increased syllabic duration in French is a result of the presence of stress, or it may be a possible component of stress. Whichever is the case, this syllable is consistently longer than the others, and is generally marked by the pitch contour, often with a rise, but when the stress group occurs at the end of an intonation unit, the stressed syllable will frequently be accompanied (at least in declarative utterances) by a fall in both F0 and in amplitude (Di Cristo 1998: 4). However, these modifications may simply result from the position of the syllable at the end of the rhythmic group (Faure & Di Cristo 1973: 234). During an utterance, the acoustic energy invested by the speaker often diminishes

towards the end of each prosodic unit, particularly intonative units, as the articulatory force diminishes, and in French, this phenomenon is especially evident. The phenomenon of final syllable lengthening is also apparent in isolated words, which has been interpreted by some as a sort of word stress in French (e.g. Delattre 1965, Dahan & Bernard 1996). Whatever the interpretation of this 'final accent', it has no discriminating value, and although it may aid in segmenting speech, it adds nothing to the meaning of the word on which it happens to occur.

1.3 Stress perception studies

This study owes a great deal to the classic stress perception studies carried out over half a century ago (Fry 1955, 1958). As stated in the introduction, stress is tremendously difficult to define, perhaps because it is not merely a phenomenon of production, but also of perception. Daniel Jones highlighted the problem in 1956: 'Stress may be described as the degree of force with which a sound or syllable is uttered. It is essentially a subjective action' (Jones 1956: 245). Advances in technology in the 1950s and 1960s (notably the Haskins Pattern Playback Synthesiser) led to new possibilities in the domain of experimental phonetics. Numerous studies were carried out on the perception of the acoustic correlates of stress, especially in English, to attempt to establish which were the most important.

In 1955, Dennis Fry used the Haskins equipment to modify the acoustic parameters of word pairs of the type *ˈpermit* (noun) *perˈmit* (verb), and he found that duration and intensity had an effect on stress perception, the former more than the latter (Fry 1955: 767–768). In a second series of tests, Fry (1958) introduced the F0 cue and found it to be more reliable than duration. He also raised the question of formant structure as a cue to stress perception, which he further examined several years later, but his results, although seeming to attribute greater importance to vowel quality than to duration, were limited by the technology available to him (Fry 1965: 431). Bolinger, in the same year and with the same equipment as Fry, carried out fourteen tests, resulting in the proposal of a theory of pitch accent in English (Bolinger 1958). The central notion was that of 'pitch prominence', which he described as 'a rapid and relatively wide departure from a smooth or undulating contour' (Bolinger 1958: 112). The importance he gave to stress as a perceptual phenomenon is highlighted in the following definition: 'When only one item is given pitch prominence, it is heard as accented. The pitch movement may be UP TO, DOWN TO, or DOWN FROM the accented syllable' (op. cit.: 127). As for the importance of duration as a cue, Bolinger explained it away in typically pragmatic fashion: 'A pitch obtrusion requires time for its execution. When the pitch accent is embraced completely by a single syllable, the syllable is lengthened to accommodate the necessary range of pitches . . . Figuratively speaking, it is there IN ORDER to make room for the accent (op. cit.: 138).

Lieberman (1960), again using word pairs, chose to examine production. He examined the acoustic cues of utterances, and corroborated the results of the perception tests of both Fry and Bolinger, finding that 'the stressed syllable had a higher fundamental frequency than the unstressed syllable of the same utterance in 90 per cent of the cases, a higher peak envelope in 87 per cent, and a longer duration in 66 per cent' (Lieberman 1960: 397). Robert Jenkins found that the cues to stress perception were, in order of importance, 'pitch', 'timbre' and 'loudness' (Jenkins 1961: 1557). Morton & Jassem (1965: 178) concurred, finding that 'fundamental frequency changes were by far the most effective in producing universally accepted stress-marking', especially rises in F0.

Interest in stress in French was admittedly less pronounced during this period, but in 1962, André Rigault used the PAT speech synthesis equipment at Edinburgh University to conduct perception tests on French subjects using synthesised words and sentences. He found that increased frequency consistently produced an impression of increased prominence (Rigault 1962: 738). Rigault's findings went against the traditional view that stress in French was primarily a durational phenomenon (Delattre 1938), but supported those of an earlier study made using an oscilloscope (Parmenter & Blanc 1933). The findings of the 1950s and 1960s

were summarised by Lehiste (1970) who examined data for five European languages: French, English, Polish, Hungarian and Swedish:

> The intensity of speech sound may be influenced by respiratory effort, by degree of opening of the vocal tract, and by the interaction between fundamental frequency and formant frequency. The perception of stressedness appears to be based on a number of factors, the most influential of which is fundamental frequency. Other phonetic correlates of stress, besides fundamental frequency and intensity, include vowel quality and duration. There is a high degree of interaction between stress and other suprasegmental features. (Lehiste 1970: 153)

It would seem therefore that in terms of production, stress in French is consistently marked by an increased duration, whereas F0 plays an important role in terms of perception.

Finally, recent efforts have often been concentrated on production, although there have also been perception studies, sometimes producing conflicting results. Studies on stress in French have also attributed a greater role to F0 than to duration as a cue, for example, Dahan & Bernard (1996). Heldner (2001, 2003) tested the reliability of intensity and spectral emphasis (frequency band-filtered intensity) as acoustic correlates of focus in Swedish. He found that although spectral emphasis was a more reliable predictor of the words receiving focus, overall intensity also increased in words containing a syllable receiving focus. Sluijter & van Heuven (1996) found, however, that overall intensity was not a reliable correlate of stress (Sluijter & van Heuven 1996: 2482). As stress depends so much on the PERCEPTION of prominence, it was important to conduct perceptual studies to explore this further. In a study involving two perception tests, Sluijter, van Heuven & Pacilly (1997: 510) found that intensity was not a reliable cue on the perception level either. Cambier-Langeveld & Turk (1999), in a study of vowel-lengthening as an indicator of focus in English and Dutch, found that in English not only was the prominent syllable lengthened, but that all the syllables in the accented word received some lengthening. Xu & Xu (2005) found 'consistent alignment of f0 valley with the onset of stressed syllable, and consistent alignment of f0 peak with the offset of stressed syllable when the syllable is non-word-final or word-final but not focused' (Xu & Xu 2005: 194). It would seem from these more recent studies, that F0 and formant structure and duration all work together, but that intensity is less important than previous studies (e.g. Fry 1955) showed. Further perception studies are necessary to help understand the picture, and cross-linguistic studies remain a useful way of providing insight into the nature of the perception of prominence.

1.4 Some differences between English and French

Perhaps the most well-known and well-documented prosodic difference between French and English is that of isochrony (Pike 1945). The traditional view is that French is a syllable-timed language, in that it accords a more or less equal period of time to each syllable (apart from the final syllable of each breath group, as previously stated). Of course, syllables in French are not all perfectly equal in length, as explained above, but the resistance of French to reduce its syllables contributes to the 'myth' of the isosyllabic nature of French (Astésano 2001: 33). English, on the other hand, is generally held to be stress-timed, as the distance in time between stressed syllables is fairly constant, no matter how many syllables occur between those stressed syllables. As a consequence, English often reduces unstressed syllables (schwa is the prime example of a reduced vowel) and certain vowels or even syllables may disappear altogether. This is not to be confused with vowel length opposition, which is also present in English but not in French. It seems to be the reduction or the suppression of unstressed syllables which causes the most difficulties for native French speakers when they attempt to reconstruct the meaning from running speech in English (Pennington 1996: 146–148). Since Pike posited the dichotomy of stress- and syllable-timed languages, certain authors have agreed with the assertion that English is isochronous (Abercrombie 1965: 18), but many authors have consistently sought to question the validity of the two categories, as Bertinetto (1989) points out. For example, Jenkins (2000: 149) states that 'the concept of stress-timing

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

Table 1 A comparison of the position of the most prominent syllable in English and French (Delattre 1965: 29).

Words of	1st	2nd	3rd	4th
1 syllable	100%			
2 syllables	0%	100%		
3 syllables	0%	0%	100%	
4 syllables	0%	0%	0%	100%

French

Words of	1st	2nd	3rd	4th
1 syllable	100%			
2 syllables	74%	26%		
3 syllables	55%	39%	6%	
4 syllables	33%	36%	29%	2%

English

appears to have little basis in reality'. Wenk & Wioland (1982: 204) prefer the terms 'trailer-timed' for French and 'leader-timed' for English, terms clearly influenced by the fact that primary word stress in English more often occurs on a syllable towards the beginning of a word than at the end. In a classic study comparing several European languages, Delattre (1965: 21) found the position of the stressed syllable to be a major difference between French and English (Table 1).

Finally, as Peter Roach has said on the subject of isochrony: 'one is obliged to conclude that the basis for the distinction is auditory and subjective – a language is syllable-timed if it sounds syllable-timed' (Roach 1982: 78). This supports my view that stress is as much (if not more) a perceptual phenomenon as an acoustic one.

To summarise the behaviour of stress in English and French, six major differences can therefore be established:

(i) English has different types of stress: word stress (or lexical stress), focus and contrastive stress (or narrow focus) which may be separated into between four and six levels. French however, does not have word stress and marks out the final syllable of each prosodic group.
(ii) The two languages use F0 (i.e. the 'pitch accent') differently.
(iii) Of the four cues to stress, F0 seems the most important in terms of perception in both languages; however, duration plays a more important role in the production of stress in French.
(iv) Stress in English tends to occur towards the beginning of the word, whereas in French, the last syllable of tone units is more prominent.
(v) In terms of isochrony, English tends to regulate its rhythm by the stressed syllables, interspersed with unstressed syllables, whereas French tends to give a more equal value to all syllables (apart from the final syllable of prosodic groups).
(vi) Unstressed and destressed syllables are generally reduced to a greater or lesser degree in English.

When learning a second language (L2), certain characteristics of the learner's mother tongue (L1) influence the learning process. When two languages differ as greatly as do English and French in terms of stress, it is to be expected that linguistic transfer should occur. In the next section, the question of linguistic transfer and how it contributed to the hypothesis is examined briefly.

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

2 Linguistic transfer

Linguistic transfer, often referred to as L1 interference, is a major factor in L2 acquisition. Whereas some errors present in learners' interlanguage are individual and some are universal, most are related directly or indirectly to the learners' L1. Children seem to learn their L1 (or even several languages) with little or no conscious effort, but adults find the task more challenging. There are many reasons for becoming a less efficient language learner with age, and although some aspects of language learning such as learning new lexical items may continue all one's life, the effects of the L1 will be felt most strongly when it comes to the phonology of a language (Pennington 1996).

In the early half of the twentieth century, the Prague Linguistic Circle were interested in linguistic transfer in relation to phonological features and coined the term 'phonological deafness' (Polivanov 1931, Trubetzkoy 1939). It is interesting to note that in the formulation of this concept, the Czech linguists were more interested in perception than in production. However, it is the relationship between perception and production which interests many researchers in second language acquisition. Most of the perception models since the 1950s highlight the active nature of perception (Rost 2002, Tatham & Morton 2006). Indeed, more recent work with neuroimaging has confirmed that areas of the brain which are active during language production are also active during language perception, which would seem to support 'active' models of perception (Mildner 2006: 35).

Lado's (1957) contrastive analysis hypothesis suggests that for a given learner, language points which are similar in L1 and L2 will be easier to learn, whereas those which are less similar will be harder to master. Eckman (1977) continues in the same vein as Lado, but integrates the theory of markedness. According to Eckman, markedness has a bearing on the ease with which a feature may be acquired. Eckman explains his 'markedness differential hypothesis' as follows:

> The areas of difficulty that a second-language learner will have can be predicted on the basis of a comparison of the first language and the target language such that:
>
> (1) Those areas of the target language that are different from the first language and are relatively more marked than in the first language will be difficult.
> (2) The degree of difficulty associated with those aspects of the target language that are more different and more marked than in the native language corresponds to the relative degree of markedness associated with those aspects.
> (3) Those areas of the target language that are different from the first language but are not relatively more marked than in the first language will not be difficult.
>
> In this conceptualization, markedness is defined in the following terms: 'A phenomenon or structure X in some language is relatively more marked than some other phenomenon or structure Y if cross-linguistically the presence of X in a language implies the presence of Y, but the presence of Y does not imply the presence of X'.
>
> (Eckman 1977: 321)

Eckman's markedness differential hypothesis could contribute to understanding the problems of some French learners of English, as stress in English when compared to French corresponds to the concept of 'structure X' in Eckman's explanation. Following Eckman's arguments cited above, as stress is more marked in English than in French, it will be more difficult to learn for native French speakers. More recent research backs up this idea; for example, Rasier & Hiligsmann (2007: 59) suggest that their study 'provides experimental support for Eckman's view that markedness is an important factor in predicting and explaining learning difficulties, especially the cases of prosodic transfer between the learners' L1 and L2'.

Flege (1995) proposes an evidence-driven model to account for L2 speech acquisition, the speech learning model (SLM). His model contains four postulates and seven hypotheses which are intended to account for the interaction of L1 and L2 phonetic subsystems throughout

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

a learner's life via the creation of memory representations, which Flege calls 'phonetic categories'. The perceived differences between categories are fundamental, and L2 sounds may be mapped onto L1 categories, which exist in a common 'phonological space' (Flege, Schirru & MacKay 2003: 469). However, as Brown (2000: 9) points out, Flege's model fails to explain the exact nature of this mapping process. A more recent model, MacWhinney's (2008) 'Unified Competition Model' takes its very name from the competing forces at play between a learner's L1 and L2 in the language-learning process. MacWhinney prefers the term 'entrenchment' and contends that the learner creates 'self-organizing maps' (SOMs) of the target language and proposes a series of mechanisms to explain the processes of entrenchment.

Whereas the majority of second language acquisition (SLA) models are based mainly on studies of segmental features, this study concerns the interaction between the prosodic features of English and French. In a study on the perception of stress in French, Dolbec & Santi (1995) suggested the existence of a 'linguistic filter', which, they claimed, was particularly present with prosodic features. They suggested that the L1 acted as a linguistic filter which conditions or orients the interpretation of the acoustic signal (Dolbec & Santi 1995: 46).

Dupoux & Peperkamp (1999), Dupoux et al. (2001), and Peperkamp & Dupoux (2002) originally developed the concept of 'stress deafness' whilst working on French and Spanish corpora and subjects but then enlarged scope of their research to take in other European languages, most notably English. According to Dupoux & Peperkamp (1999), the listening apparatus of interlocutors is tuned to a greater or lesser degree according to their L1. As a result, they often have difficulties when exposed to a language other than the one to which they were exposed as children (Dupoux & Peperkamp 1999: 203). Although the term 'stress deafness' is somewhat dramatic, stress in English does clearly create great difficulties for certain French learners of English. It is reasonable to assume that the problems which many native French speakers experience in deciphering authentic running speech in English stem largely from the four main differences which were listed at the end of Section 1 above. In order to investigate this question more fully, a study of stress perception was conducted, which is presented in the next section.

3 The experiment

3.1 The hypothesis

As the term 'stress deafness' implies, certain French native speakers appear unable to identify stressed syllables in English. As explained in Section 1 above, there are four acoustic cues to prominence: F0, amplitude, duration and formant structure. In terms of perception, these four acoustic cues correspond to pitch, volume, length and timbre. Based on what is known about the production of stress in English and French, it may be assumed that speakers of these two languages listen for stress differently. To be more precise, the relative importance attributed to each of the perceptual correlates of the acoustic cues of prominence may vary according to the L1 of a given speaker. One may also wonder whether, given the phenomenon of linguistic transfer, this relative importance of the cues is carried over into the L2. These questions led us to formulate the following hypothesis:

> French native speakers and English native speakers do not perceive prominence in the same way in English and in French.

The experiment is therefore an attempt to determine the relative importance which the subjects attribute to the acoustic cues and/or their perceptual correlates according to their native language, and according to the language they are listening to. The cues which were focused on were F0 (pitch), duration (length) and formant structure (timbre). The variable of amplitude (volume) was not included, as existing research on English and French has

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

shown that F0 and duration are the most important cues in the two languages, respectively, as discussed earlier. The cue of formant structure was added, as syllable reduction is a phenomenon which is much more common in English than in French, and is commonly held to be a problem area for French learners by language teachers and resource developers in France.

3.2 Subjects

This study featured twenty subjects, of which ten were native English speakers and ten were native French speakers. All were aged between twenty and twenty-five years. The French subjects were all studying History and Geography, whereas the group of English speakers was more heterogeneous, comprising a mixture of students studying in France on the Erasmus university exchange programme. The language proficiency of the subjects was not tested and varied between subjects, although none could be described as being near native-speaker level in their L2.

3.3 Stimuli

In the tradition of the experiments carried out at Haskins labs in the 1950s, this study uses two-syllable word pairs where the positioning of word stress depends on the grammatical category. It therefore focused entirely on syllable prominence at the word level, but as the stimuli were generated from words read aloud in both English and French, and word stress does not occur in French, it cannot be said that word stress was compared in the two languages. Where this study differs from previous studies is in its use of the word pair protocol not only for the English stimuli but also for the French stimuli. The stimuli consisted of four words: two English and two French. For the English words, *ˈtransfer* /ˈtrænsfɜː/ (noun) and *transˈfer* /trænsˈfɜː/ (verb) were chosen. As discussed in Section 1 above, French does not have word stress, but it was possible to use the local regional variety of French to constitute the French stimuli. In standard French, when the letter *-e* is present in the orthography in word-final position, it is not pronounced. In the various accents of the Midi region of France, however, this vowel, known in French as '*e-caduc*' or '*e-muet*', is generally realised. A recent study showed it to be present in between 52% and 86.8% of cases where a final *-e* was present in the orthography (Durand 2009). Whilst not actually the same as an English schwa, this vowel was as close as possible in French to the final vowel of the English noun *ˈtransfer*. The meridional French schwa is closer to the French /ø/ in unstressed position (Durand 2009: 134), making it a little more short, tense, rounded and front than /ɜː/, and more close, front and rounded than schwa. Using this idea, the French words *boîte* /bwatə/ and *boiteux* /bwatø/ were chosen as stimuli. This pair of words, whilst not being distinguished by word stress, behave acoustically in a fashion similar to the two English words which have stress on the first and second syllable, respectively. Because the meridional accent was used, both utterances are composed of two syllables (CVCV): in *boîte*, the first syllable is more prominent than the second, and in *boîteux*, the second syllable is more prominent than the first. It is important to add that in normal (i.e. non-experimental) conditions, native speakers of either standard or meridional French perceive *boîte* as a one-syllable word and *boîteux* as a two-syllable word, even when uttered with the meridional accent, whereas a non-French speaker would perceive two syllables for each word. However, in the experimental conditions described in this paper, none of the subjects questioned the possibility of identifying one of two syllables for the stimuli.

The four original words were recorded in the anechoic chamber of the Laboratoire Parole et Langage at the Université de Provence, Aix-en-Provence. The two French words were read by a French native speaker from the town of Marseilles, and the two English words by a native British English speaker. The stimuli were then validated by ten native speakers for each language. The validation process entailed playing the stimuli as isolated words and also

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

embedded in utterances to the twenty native speakers and asking them to rate the stimuli as sounding 'authentic' (i.e. uttered by a native speaker) or 'not authentic'. All the stimuli were validated, and several of the French speakers commented on the strong meridional accent.

Next it was necessary to synthesise the four words in order to create the stimuli for the experiment. MBROLA[1] was used to modify the duration and the formant structure and Momel[2] to modify the F0 curve. The Momel algorithm allows us to modify the F0 curve by factoring the raw curves into two components: a macromelodic component, which corresponds to the global pitch contour of the utterance, and which is independent of the nature of the constituent phonemes, and a micromelodic component consisting of deviations from the macromelodic curve and which is determined entirely by the segmental constituents of the utterance and is independent of the macromelodic component. The four stimuli all had rise-fall peaks on the stressed syllable, and the Momel software enabled us to manually lower the peak on the stressed syllable and raise the peak on the unstressed syllable at regular intervals until they reached the opposite values to the starting positions, whilst the algorithms used by Momel ensured a smoothing of the contour. This operation was performed over regular (linear) intervals.

In total, 100 stimuli were created from the four original words, 50 from the English word-pair and 50 from the French word-pair (see Table 2). For the purposes of explaining this process, a word stressed on the first syllable is called W1 and a word stressed on the second syllable W2. Five stimuli were generated from W1 in each language, with vowel durations varying between W1 and W2 (i.e. three intermediate stimuli). As vowel structure was also an element which was to feature in the test, a way had to be devised of introducing different formant structures into the stimuli. To have modified all of the formants for each stimulus would have been very difficult and time-consuming, so instead of actually modifying the formant structure, the process described above was simply repeated, but using W2 to generate the five stimuli for each language. This produced ten words for each language, where the vowel quality was determined either by W1 or by W2. The 50 experimental stimuli for each language were obtained from these 'words' by modifying the F0 curve in five stages from W1 to W2. In this way, 50 stimuli were obtained for each language where F0, duration and formant structure could each be identified as somewhere between W1 and W2. For example, a stimulus could be *transfer* (W1D4F3). This stimulus would originate from W1 (therefore the formant structure is akin to *'transfer* not to *trans'fer*), the relative duration of each syllable is four-fifths of the way between *'transfer* and *trans'fer*, and the F0 curve is three-fifths of the way between *'transfer* and *trans'fer*. These five steps are not intended to correspond to the levels of stress mentioned above (Cruttenden 1986, Pennington 1996) but simply to establish five distinct gradations of the parameters for the purposes of the protocol.

3.4 Protocol

The subjects' task was simple: they had to tick a box corresponding to the syllable which they thought was stressed (*accentuée* in French). The instructions asked them specifically to tick either the box labelled W1 or W2 (*'transfer*/*trans'fer*) or M1/M2 (*boîte*/*boiteux*) depending on which word they thought each stimulus most resembled. For example, the instructions in English were:

> Please tick the box labelled W1 if you think the word you hear resembles the noun *'transfer* (i.e. stressed on the first syllable), and tick the box labelled W2 if you think the word you hear most resembles the verb *trans'fer* (i.e. stressed on the second syllable).

[1] MBROLA is a speech synthesis program developed at the Faculté Polytechnique de Mons, Belgium (Dutoit & Pagel 1996).

[2] Momel (Modelling Melody) is a program developped by the Laboratoire Parole et Langage, Aix-en-Provence (Hirst, Di Cristo & Espesser 2000).

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

Table 2 The generation of the final 50 stimuli for each language: 25 stimuli from W1 and 25 from W2. All 25 stimuli have different F0 curves (steps 1–5), and each stimulus was obtained from one of the 5 duration-modified 'words'.

Original word	W1 > W2																								
Duration steps	D1					D1					D3					D4					D5				
F0 steps (=25 stimuli)	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Original word	W2 > W1																								
Duration steps	D1					D1					D3					D4					D5				
F0 steps (=25 stimuli)	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

Table 3 Percentage results of 'correct' answers for both tests and both groups of subjects (i.e. W1 was identified as originating from W1 and W2 from W2).

	English stimuli		French stimuli	
	W1 as W1 ('*transfer*)	W2 as W2 (*trans'fer*)	W1 as W1 (*boîte*)	W2 as W2 (*boiteux*)
English subjects	58.8%	56.6%	66.2%	80.2%
French subjects	59.6%	61%	62.2%	70.2%

There was a limited response time as the recording of the stimuli was played non-stop at a constant rate of one stimulus per second. The boxes were labelled W1 and W2 for the English test and M1 and M2 for the French test. The subjects listened to the stimuli through headphones in a language laboratory. First, the subjects read the instructions. Then they completed a warm-up test consisting of ten items in their L1. Next they listened to the 50 stimuli played in random order twice each in their L1. After a few minutes' pause, the subjects completed the same warm-up test, and then the same test of 50 items played twice in their L2. The experiment may therefore be represented as:

$$L_2 \times W_2 \times D_5 \times F_5 \times R_2 \times N_2 < S_{10} >$$

where:

L = language
W = word generated from W1 or from W2
D = modification of duration
F = modification of F0 curve
R = repetition of stimuli
N = native language of subjects
S = number of subjects

3.5 Results

For each of the two tests (one for the French stimuli and one for the English stimuli), there was one dependent variable (the subjects' answers) and four independent variables, WORD (if the subject 'correctly' identified whether the stimulus had been generated from W1 or W2), F0 CURVE, DURATION and NATIVE LANGUAGE. From the first ANOVA table of the results, it was clear that the factor WORD was present in all the significant results (results were considered significant if p was less than .05). In other words, the quality of the vowels played a primary role in the stress judgements of all the subjects; the fact that the stimuli had been generated from either W1 or W2 was a very important factor in influencing the subjects' choice. The fact that segmental cues were so important is an interesting result in itself, but in order to better understand the other factors at play, the results were recoded to change the way the factor WORD was used. Henceforth, if a subject chose the first syllable for a stimulus created from W1 or the second syllable for a stimulus created from W2, that answer was coded R (for RIGHT) and W (WRONG) if not. In this way, it is not the actual initial word itself used to generate each stimulus which is taken into account, but only whether an initial word has influened the subjects' decisions. The results of both tests for both groups of subjects are shown in Table 3 and represented as a histogram in Figure 1, with percentages for the number of times W1 was perceived as W1 and W2 was perceived as W2. It is clear from these results that, as mentioned above, in the majority of cases the word was 'correctly' identified, i.e. the formant structure proved to be a very powerful indicator of the prominent syllable.

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

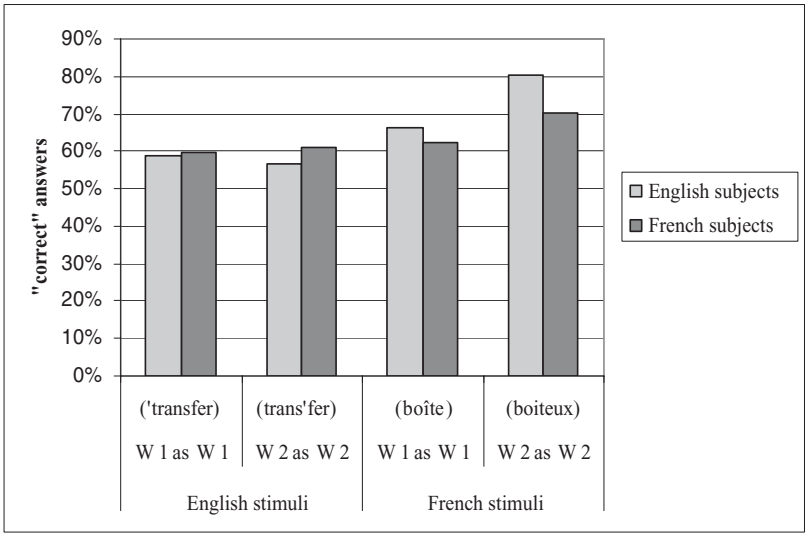

Figure 1 Percentage results of 'correct' answers for both tests and both groups of subjects (i.e. W1 was identified as originating from W1 and W2 from W2).

Table 4 English stimuli (recoded without WORD) ANOVA table for 'RIGHT/WRONG'.

	DF	Sum of squares	Mean square	*F*-value	p-value
WORD	1	0.008	0.008	0.039	.8444
DURATION	4	21.205	5.301	25.519	<.0001
WORD × DURATION	4	0.927	0.232	1.116	.3474
F0 CURVE	4	41.140	10.535	50.713	<.0001
WORD × F0 CURVE	4	0.992	0.248	1.194	.3115
WORD × F0 CURVE	16	3.005	0.188	0.904	.5642
WORD × DURATION F0 CURVE	16	6.073	0.380	1.827	.0232
NATIONALITY	1	0.338	0.338	1.627	.2032
WORD × NATIONALITY	1	0.162	0.162	0.780	.3773
DURATION × NATIONALITY	4	1.417	0.354	1.705	.1461
WORD × DURATION × NATIONALITY	4	1.303	0.326	1.568	.1802
F0 CURVE × NATIONALITY	4	3.182	0.795	3.829	.0042
WORD × F0 CURVE × NATIONALITY	4	1.478	0.370	1.799	.1304
DURATION × F0 CURVE × NATIONALITY	16	4.363	0.273	1.313	.1799
WORD × DURATION × F0 × NATIONALITY	16	2.507	0.157	0.754	.7389
Residual	1900	394.700	0.208		

The significant results for the English test are shaded in the ANOVA table (Table 4).

They may be summarised as follows:

(i) The greater the modification of the duration of the vowel segments of the stimuli, the greater the effect on the stress judgements of both groups of subjects.
(ii) The same effect was observed for modifications of the F0 curve, but to a slightly more pronounced degree.
(iii) The English-speaking subjects were more strongly influenced by modifications of the F0 curve than were the French speaking subjects.

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

Table 5 French stimuli (recoded without WORD): ANOVA table for 'RIGHT/WRONG'.

	DF	Sum of squares	Mean square	F-value	*p*-value
WORD	1	2.546	2.546	13.013	.0003
DURATION	4	20.253	5.063	25.879	<.0001
WORD ×DURATION	4	1.827	0.457	2.335	.0535
F0 CURVE	4	5.223	1.306	6.674	<.0001
WORD × F0 CURVE	4	6.456	1.614	8.250	<.0001
WORD × F0 CURVE	16	8.747	0.547	2.794	.0002
WORD × DURATION F0 CURVE	16	3.602	0.225	1.151	.3015
NATIONALITY	1	5.611	5.611	28.678	<.0001
WORD × NATIONALITY	1	0.146	0.146	0.746	.3878
DURATION × NATIONALITY	4	0.551	0.138	0.704	.5893
WORD × DURATION × NATIONALITY	4	1.097	0.274	1.402	.2308
F0 CURVE × NATIONALITY	4	1.611	0.403	2.058	.0839
WORD × F0 CURVE × NATIONALITY	4	0.476	0.119	0.608	.6567
DURATION × F0 CURVE × NATIONALITY	16	3.519	0.220	1.124	.3257
WORD × DURATION × F0 × NATIONALITY	16	2.562	0.116	0.819	.6654
Residual	1900	371.727	0.196		

For the French test, the significant results from the ANOVA tests are shaded in Table 5. They may be summarised as follows:

(iv) Both groups of subjects identified stimuli generated from W1 more often than stimuli created from W2, i.e. the segmental cues present in W1 were a greater influence on subjects' stress judgements than those present in W2.

(v) As for the English test, the greater the modification of the duration of the vowel segments of the syllables in the stimuli, the greater the effect on the stress-judgements of both subject groups.

(vi) Modifications to the F0 curve of the stimuli appeared to be significant, but no discernible pattern was perceived in relation to subjects' stress judgements.

(vii) English-speaking subjects were more influenced by modifications to the formant structure with the French stimuli than the French-speaking subjects.

(viii) English-speaking subjects behaved differently to French-speaking subjects independently of the other factors (F0, formant structure or duration).

3.6 Discussion

The results of this study support the hypothesis that French and English native speakers listen to stress differently in English and in French, albeit for a small sample of subjects and in limited phonological contexts. The experimental design of this study was not devised to prove or disprove the concept of 'stress deafness', but it does seem that some sort of linguistic transfer has occurred. The most evident result is the importance of segmental cues in identifying the stressed syllable, for both groups of subjects. Both from the results shown in Table 3 and from the ANOVA tests, it is evident that for the French stimuli, vowel quality proved even more reliable as an indicator than for the English stimuli. It is likely (although not certain) that this is due to the way the experiment was constructed. In English, it is normal to reduce the second vowel of a word such as '*transfer* to a schwa, and in various communication situations, the second vowel may indeed be anywhere between the full vowel of *trans'fer* and the reduced vowel of '*transfer*. The same cannot be said for the two second vowels of the French stimuli, where any variation which may be observed in a communicative context would not be due to vowel reduction through destressing but to other factors, such as dialectal or idiolectal influences. Another possible explanation is that in manipulating the F0 curve with

https://doi.org/10.1017/S0025100310000253 Published online by Cambridge University Press

the Momel algorithm, the formant structure of the vowels was affected. Putting aside these two potential experiment-based explanations for this result, it may well be that segmental factors are indeed an important cue in English for identifying the prominent syllable. Certain authors maintain that with stress-timed languages such as English, segmentation of the speech signal is achieved through the alternation of strong and weak syllables (Cutler et al. 1997: 148). If this is the case, segmental cues, and in particular vowel quality, would be more important than the cues of F0, amplitude and duration.

The most interesting result, at least given the literature cited in Section 1 above, is the fact that the English-speaking subjects relied more on F0 than the French subjects for the test in their own language (point (iii) above). Also, F0 modifications influenced the stress judgements for both groups of subjects in the English test, but not in the French test. This holds with the view that the pitch contour which accompanies the prominent syllable is of great importance in English and that, at least subjectively, the cue of pitch is relatively more important in English than in French, and perhaps the most important of all the cues. The perception studies of the 1950s and 1960s all point to F0 being the most important cue for English speakers, but the picture is less clear for French speakers. F0 is perhaps the most important cue for French speakers, too, but duration seems to play a larger role than for English speakers (see Astésano 2001 for a literature review). This may also be due to the fact that French does not have vowel length opposition as English does, which 'frees' vowel duration as a cue for prominence. The results support the view that F0 is an important cue in the perception of stress for speakers of both languages, but it is more reliable for English stimuli: the results for the English test showed that F0 modification was a reliable cue for stress position in English, whereas in French, although F0 was picked up on by the subjects, they were unable to use the cue effectively.

The other result which deserves comment is the fact that English-speaking subjects were more influenced by vowel quality for the French test than were the French native speakers (point (vii) above). A possible reason for this may be that the English subjects resorted to segmental cues in the absence of sufficiently salient cues in F0, in other words, the lack of a pitch accent in French. This would seem to be a clear case of linguistic transfer, and it may be an example of what Ellis (2008) calls 'contingency', i.e. resorting to another contingent linguistic cue in the absence of the cue which would be used in the L1: 'a cue with high availability but low reliability may initially be used over a cue that is of lower availability, even though it is in fact more reliable' (Ellis 2008: 378). Indeed as Cutler et al. (1996: 147–148) point out, the importance of the opposition between 'full' stressed and reduced unstressed syllables in English may mean that segmental cues, and in particular vowel quality, are more important for native English speakers than for native French speakers. It is difficult to explain, however, without further experimentation, why the English and French subjects behaved differently independent of the factors of vowel quality, F0 and duration.

4 Conclusion

The results support the hypothesis that English and French native speakers listen for stress differently and that some form of linguistic transfer is apparent. Nevertheless, these findings cannot necessarily be generalised to other contexts. In order to confirm these results, more studies must be undertaken, involving more subjects, a greater variety of stimuli and targeting stress in other contexts, not simply at word level.

Acknowledgements

I would like to thank Daniel Hirst and his colleagues at the Laboratoire Parole et Langage, Université de Provence, Aix-en-Provence, for all their help and encouragement.

References

Abercrombie, David. 1965. A phonetician's view of verse structure. In David Abercrombie, *Phonetics and linguistics*, 16–25. London: Oxford University Press.

Astésano, Corinne. 2001. *Rythme et accentuation en français. Invariance et variabilité stylistique*. Paris: Editions L'Harmattan, Collection Langue et Parole.

Beckman, Mary E. & Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. In Patricia Keating (ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III*, 1–33. Cambridge: Cambridge University Press.

Benguerel, André-Pierre. 1973. Corrélats physiologiques de l'accent en français. *Phonetica* 27, 21–35.

Bertinetto, Pier Marco. 1989. Reflections on the dichotomy 'stress' vs. 'syllable-timing'. *Revue de Phonétique Appliquée* 91–93, 99–130.

Bolinger, Dwight. 1958. A theory of pitch accent in English. *Word* 14, 109–149.

Brown, Cynthia. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. In John Archibald (ed.), *Second language acquisition and linguistic theory*, 4–63. Malden, MA: Blackwell.

Cambier-Langeveld, Tina & Alice Turk. 1999. A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics* 27, 255–280.

Couper-Kuhlen, Elizabeth. 1985. *An introduction to English prosody*. Tübingen: Niemeyer.

Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.

Cutler, Anne, Delphine Dahan & Wilma van Donselaar. 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40(2), 141–201.

Dahan, Delphine & Jean-Marc Bernard. 1996. Interspeaker variability in emphatic accent production in French, *Language and Speech* 39(4), 341–374.

Delattre, Pierre. 1938. L'accent final en français: accent d'intensité, accent de hauteur, accent de durée. *The French Review* 12(2), 141–145

Delattre, Pierre. 1965. *Comparing the phonetic features of English, French, German and Spanish: An interim report*. Heidelberg: Julius Groos.

Di Cristo, Albert. 1998. Intonation in French. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 88–103. Cambridge: Cambridge University Press.

Dolbec, Jean & Serge Santi. 1995. Effet du filtre linguistique sur la perception de l'accent: étude exploratoire. *Travaux de l'Institut de Phonétique d'Aix* 16, 41–60.

Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés. 2001. A robust method to study stress-deafness. *Journal of the Acoustical Society of America* 110(3.1), 1606–1618.

Dupoux, Emmanuel & Sharon Peperkamp. 1999. Fossil markers of language development: Phonological 'deafness' in adult speech processing. In Bernard Laks & Jacques Durand (eds.), *Cognitive phonology*, 168–190. Oxford: Oxford University Press.

Durand, Jacques. 2009. Essai de panorama critique des accents du midi. In Luc Baronian & France Martineau (eds.), *Le français, d'un continent à l'autre: mélanges offerts à Yves Charles Morin* (Collection Les voies du français), 123–170. Québec: Presses de l'Université Laval.

Dutoit, Thierry & Vincent Pagel. 1996. Le projet MBROLA: vers un ensemble de synthétiseurs vocaux disponibles gratuitement pour utilisation non-commerciale. *Actes des Journées d'études sur la parole, Avignon*, 441–444.

Eckman, Fred. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27, 315–330.

Ellis, Nick. 2008. Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention and the limited L2 endstate. In Robinson & Ellis (eds.), 164–194.

Faure, Georges & Albert Di Cristo. 1973. *Phonétique générale et phonétique descriptive du français*. Ann Arbor, MI: University of Michigan Press.

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In Winifred Strange (ed.), *Speech perception and linguistics experience: Issues in crosslanguage research*, 233–277. Timonium, MD: York Press.

Flege, James E., Carlo Schirru & Ian R. A. MacKay. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40(4), 467–491.

Fry, Dennis. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustic Society of America* 27(4), 765–768.
Fry, Dennis. 1958. Experiments in the perception of stress. *Language and Speech* 1(2), 126–152.
Fry, Dennis. 1965. The dependence of stress judgements on vowel formant structure. In Eberhard Zwirner & Wolfgang Bethge (eds.), *6th International Congress of Phonetic Sciences* (ICPhS 6), 306–311. Basel: Karger.
Heldner, Mattias. 2001. Spectral emphasis as a perceptual cue to prominence. *TMH-QPSR* [Speech, Music and Hearing Quarterly Progress and Status Report] 42, 51–57.
Heldner, Mattias. 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics* 31, 39–62.
Hirst, Daniel, Albert Di Cristo & Robert Espesser. 2000. Levels of representation and levels of analysis for intonation. In Merle Horne (ed.), *Prosody: Theory and experiment*, 51–87. Dordrecht: Kluwer.
Jenkins, Jennifer. 2000. *The phonology of English as an international language*. Oxford: Oxford University Press.
Jenkins, Robert. 1961. Perception of pitch, timbre and loudness. *Journal of the Acoustical Society of America* 33(11), 1550–1557.
Jones, Daniel. 1956. *The pronunciation of English*. Cambridge: Cambridge University Press.
Jun, Sun-Ah & Cécile Fougeron. 1995. The accentual phrase and the prosodic structure of French. *13th International Congress of Phonetic Sciences (ICPhS 12)*, vol. 2, 722–725.
Jun, Sun-Ah & Cécile Fougeron. 2000. A phonological model of French intonation. In Antonis Botinis (ed.), *Intonation: Analysis, modeling and technology*, 209–242. Dordrecht: Kluwer.
Lacheret-Dujour, Anne & Frédéric Beaugendre. 1999. *La prosodie du français*. Paris: CNRS.
Ladd, D. Robert. 2008. *Intonational phonology*. Cambridge: Cambridge University Press.
Lado, Robert. 1957. *Linguistics across cultures*. Ann Arbor, MI: Michigan University Press.
Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *Journal of the Acoustic Society of America* 32, 451–454.
MacWhinney, Brian. 2008. A unified model. In Robinson & Ellis (eds.), 341–371.
Mildner, Vesna. 2006. *The cognitive neuroscience of human communication*. New York: Lawrence Erlbaum.
Morton, John & Wiktor Jassem. 1965. Acoustic correlates of stress. *Language and Speech* 8, 159–187.
Parmenter, Clarence E. & A. V. Blanc. 1933. An experimental study of accent in French and English. *Publications of the Modern Language Association of America* 48(2), 598–607.
Pennington, Martha. 1996. *Phonology in English language teaching*. London: Longman.
Peperkamp, Sharon & Emmanuel Dupoux. 2002. A typological study of stress deafness. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory phonology 7*, 203–240. Berlin: Mouton de Gruyter.
Pike, Kenneth. 1945. *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
Polivanov, Evgenij. 1931. La perception des sons d'une langue étrangère. *Travaux du cercle linguistique de Prague* 4, 79–96.
Rasier, Laurent & Philippe Hiligsmann. 2007. Prosodic transfer from L1 to L2: Theoretical and methodological issues. *Nouveaux cahiers de linguistique française* 28, 41–66.
Rigault, André. 1962. Rôle de la fréquence, de l'intensité et de la durée vocaliques dans la perception de l'accent en français. *4th International Congress of Phonetic Sciences, Helsinki* (ICPhS 4), 735–748. The Hague: Mouton.
Roach, Peter. 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In David Crystal (ed.), *Linguistic controversies: Essays in linguistic theory and practice in honour of F. R. Palmer*, 73–79. London: Edward Arnold.
Robinson, Peter & Nick Ellis (eds.). 2008. *Handbook of cognitive linguistics and second language acquisition*. Abingdon: Routledge.
Rossi, Mario. 1979. Le français, langue sans accent? In Ivan Fónagy & Pierre Léon (eds.), *L'accent en français contemporain*, 93–106. Montréal, Paris & Bruxelles: Didier.
Rost, Michael. 2002. *Teaching and researching listening*. Harlow: Pearson.

Sluijter, Agaath & Vincent van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100(4), 2471–2485.
Sluijter, Agaath, Vincent van Heuven & Jos Pacilly. 1997. Spectral balance as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America* 101(1) 503–513.
Tatham, Mark & Kate Morton. 2006. *Speech production and perception*. Basingstoke: MacMillan.
Trubetzkoy, Nikolaj S. 1939. Grundzüge der Phonologie. *Travaux du cercle linguistique de Prague* 7.
Wenk, Brian J. & François Wioland. 1982. Is French really syllable-timed? *Journal of Phonetics* 10, 193–216.
Xu, Yi & Ching Xu. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33, 159–197.