

A Study of Isochrony Theories in English

Masaki TSUDZUKI and Atsunori KAMIYA

要 旨

伝統的に英語は強勢が等時間隔に繰り返し生じる言語であると考えられてきた。この概念が音声学者の間で注目されるようになったのはPike (1945)とAbercrombie (1964)に拠るところが大きい。特にAbercrombie (1964: 217)が、等時性が生じる強勢間を「フット」(foot:脚)と名付けたことで、英語のリズム構造を理論的・体系的に説明することが容易になり、日本の英語音声教育においてもこの考えが浸透した。

このように英語の等時性はフット間に生じる現象として認識されてきたが、その一方で、近年のサウンドスペクトログラフなどの音響分析機器の飛躍的な進歩に伴って、言語直感に基づいた英語フットの等時性に疑問を投げかける実証的音声リズム研究が相次いで報告されてきた(Shen & Peterson, 1962; Bolinger, 1965; O'Connor, 1965; Lea, 1974; Nakatani, O'Connor & Aston, 1981)。生成面に焦点をあてた等時性理論が音響実験により否定されたことによって、多くの研究者が等時性は聴き手の知覚レベルに存在するものではないかと考えるようになった。実際、様々な知覚実験を通して、英語フットの等時性が妥当な現象であることが立証されてきた。これらを代表するような実験に拠ると、英語母語話者は実際の発話時間よりもフット間を等時間隔に知覚しているというのである(Lehiste, 1977; Donovan & Darwin 1979; Benguerel & D'Arcy, 1986)。

結局、Classe (1939)が述べたように、どの程度まで等時性を認めるか、さらにどの視点から等時性について論じるのかが重要なのであると筆者らは考える。

本研究では、これまで報告されてきた等時性もしくは強勢拍リズムに関する代表的な研究を検証しつつ、英語の等時性の妥当性について種々な角度から考察を試みる。

1. Introduction

All spoken languages are classified as either stress-timed languages or syllable-timed languages (Jones, 1918; Pike, 1945; Abercrombie, 1967, etc.). In stress-timed languages, the intervals between stresses theoretically have equal length regardless of the number of syllables within each foot, whereas in

syllable-timed languages, successive syllables are of equal duration regardless of the segmental makeup of the syllable (Abercrombie, 1964). English, Russian, and Arabic fall into the first category, and French, Spanish, Telugu, and Yoruba are examples of the second category (Dauer, 1983). A third category, known as “mora-timed” languages, exemplified by Japanese and Ancient Greek, has been proposed by many linguists (Han, 1962; Ladefoged, 2010). However, it has been pointed out that mora-timed languages are similar to syllable-timed languages; therefore, they can be subsumed into the class of syllable-timed languages (Kubozono & Honma, 2002).

In stress-timed languages, the intervals between stressed syllables have equal duration, and the duration of each syllable lengthens or compresses, according to the number of unstressed syllables within each foot (Teschner & Whitley, 2004). This means that the duration of each interval is almost the same, even though the numbers of syllables in each foot are different (Dauer, 1983). On the contrary, in syllable-timed languages, the temporal adjustment seen in stress-timed languages is not observed and each syllable is produced at an even interval of time. Therefore, the duration of each utterance is determined by the number of syllables within each utterance.

Up to the present time, two rhythmic classifications proposed by Pike (1945) and later elaborated by Abercrombie (1964) have been widely accepted among linguists (Bolinger, 1965; O’Connor, 1967; Halliday, 1970; Uldall, 1971; Corder, 1973; Catford, 1977; Major, 1981). Abercrombie (1964) introduced the concept of “foot” as an isochronous unit, specifying that each English foot consists of a stressed syllable followed optionally by weak syllables, up to but not including the next stressed syllable. Consecutive feet are of equal duration, regardless of the number of syllables in each foot. For example, the sentence “Walk down the path to the end of the canal.” is divided into feet as follows (Roach, 2009): |Walk |down the |path to the |end of the ca|nal.|

However, Abercrombie’s theory does not allow weak syllables to be included in a foot if they appear at the beginning of a sentence. For example, “The” in Example (1) below is considered to be an “anacrusis” (i.e., a redundant element) because it is not included in Abercrombie’s foot (Takebayashi, 1996).

Example (1)

The| boy be | longs to the |music | club.|
 foot foot foot foot

The most important point of Abercrombie’s theory is that feet are independent of syntax. Thus a foot boundary is allowed to occur within a word. For example, the word “belongs” in the above example can

be allocated to two separate feet.

Abercrombie's foot concept was inherited by Halliday (1970), who introduced the idea of "silent stress," which carries the stress before an anacrusis, producing a foot at the beginning of a sentence. Therefore, Halliday's theory regards the word "the" in Example (2) below as part of the first foot.

Example (2)

|* The| boy be|longs to the |music |club.| (* indicates a silent stress)
foot foot foot foot foot

Although many other linguists have proposed theories of English rhythm (Gimson, 1985; O'Connor, 1967; Jassem, Hill & Witten, 1984, etc.), there is no definite consensus among them. However, Abercrombie's theory now receives widespread support. For example, Yamauchi (2000) compared three major theories developed by Abercrombie, O'Connor, and Bolinger. She divided several sentences into rhythm units according to the methods proposed by them. The measurement of the duration of each rhythm unit showed that Abercrombie's theory is most appropriate as an isochronous way of determining the rhythm unit of English. Moreover, many rhythm studies have been reported based on Abercrombie's theory, providing an impressive body of evidence in his favor (Cutler, 1980).

Abercrombie's theory has clearly attracted widespread support. Yet many linguists are skeptical about it because many studies have failed to show an isochronal tendency among English feet when inter-stress intervals were measured.

In this study, we will consider the validity of English isochrony by reviewing previous studies on English speech.

2. Isochrony in English

It has traditionally been claimed that English utterances have a regular timing of stressed syllables (Abercrombie, 1964). This isochronal phenomenon can easily be observed and many arguments have been made in its favor. For example, in idiomatic phrases such as "*tea* or *coffee*" or "*cup* and *saucer*," it is not common for the word order to be reversed like "*coffee* or *tea*" or "*saucer* and *cup*." This is because the former consists of repetitions of strong and weak syllables, unlike the latter. So the former is more appropriate for a stress-timed language. Similarly, lexical stresses on certain words and compounds tend to move so that there is more regularity between feet. Thus "*Japanese tea*" and "*Thirteen boys*" are heard

frequently instead of “*Japanese tea*” and “*Thirteen boys*.”

With the advance of technology in recent years, many linguists have used empirical methods to investigate the validity of isochrony from the viewpoint of production. The first technological experiment on isochrony was performed by Classe (1939), who measured inter-stress intervals (ISIs) using a kymograph (a scientific instrument which measures changes in sound pressure). Although Classe failed to show absolute isochrony either in production or perception, he did not dismiss the notion of isochrony, but regarded it as a tendency to speak in rhythmic units that are perceived as isochronous.

2. 2. Isochrony in Production

Although many studies of English rhythm have been reported, the concept of isochrony remains elusive. This is partly because of the wide range of methods used by researchers, which makes it difficult to compare the results.

We will now review several experiments which have claimed to show evidence against isochrony in English.

(1) Shen & Peterson (1962)

In this experiment, subjects were asked to read written prose, and the intervals between succeeding primary stresses were measured by using a machine. There were three subjects, and each subject read a different prose. The intervals between the primary stresses ranged from 410 to 1820 ms for the first subject, from 380 to 2500 ms for the second subject, and from 550 to 3610 ms for the third subject. The intervals were too variable to support isochrony, even when the secondary stress was considered to carry the unit of rhythm. Thus, Shen & Peterson rejected the notion of isochrony in English.

(2) Bolinger (1965)

Six subjects were asked to read two very long sentences. The intervals between the stresses were measured, which showed considerable variation among inter-stress intervals (ISIs). One fourth of the intervals (13 out of the 53 intervals) were particularly long in duration. They had approximately twice the length of the shortest interval. Bolinger suggested that syllable structures and semantic factors affected the length of the inter-stress intervals. Bolinger therefore also found the theory of isochrony untenable.

(3) O'Connor (1965, 1968)

O'Connor (1965) investigated isochrony under regular conditions, recording by hand and making a click sound at each stress. The average of 15 inter-stress intervals was 518 ms. The difference between the longest and the shortest interval was 88 ms. In spite of these results which seem to support isochrony, strict isochrony was not observed. Moreover, O'Connor (1968) analyzed seven utterances, each consisting of

three monosyllabic feet. O'Connor changed the segmental length of the second foot from three segments to nine segments, while the length of the first foot and the third foot remained constant. The "compression effect," in which syllables in polysyllabic feet are squeezed together in order to maintain equal spacing of stresses, was not observed in his results. In fact, the duration of each foot increased almost linearly as the number of syllables in each foot increased. On these grounds, O'Connor expressed his doubts about the existence of isochrony in production.

(4) Lea (1974)

In this study, the inter-stress intervals (ISIs) of 31 sentences read by eight subjects were measured in order to test the isochronal hypothesis proposed by Pike (1945). Six out of the eight subjects were asked to read a script which was a mixture of various genres, while two other subjects read a monosyllabic script. The results were that the average ISIs increased in proportion to the number of intervening syllables. Therefore, he claimed that there was no support for the theory of isochrony.

(5) Nakatani, L. H., O'Connor, K. D. & Aston, C. H. (1981)

Nakatani et al. (1981) examined isochrony in American English speech by using two-word adjective-noun phrases which were each 3 to 5 syllables long. They tried to find evidence to support a more liberal interpretation of isochrony in English. However, they failed to show the tendency for metrical feet to compress in duration. In fact, foot duration increased almost linearly with foot size, contrary to the hypothesis that the relationship between foot size and foot duration should be inversely proportional. This result led Nakatani et al. (1981) to reject the theory of isochrony in American English speech.

(6) Roach (1982)

Syllable durations of three stress-timed languages were compared with those of three syllable-timed languages, and it turned out that the standard deviations of the syllable durations for both types of languages did not differ significantly. Moreover, Roach (1982) compared the standard deviations of inter-stress intervals (ISIs) among the same six languages and found, contrary to his predictions, that the deviations from isochrony were greater in English than in any of the other five languages. On the basis of these results, he dismissed the concept of isochrony in English.

(7) Dauer (1983)

To investigate isochrony of different languages, Dauer (1983) compared five languages which include stress-timed languages and syllable-timed languages: English, Thai, Spanish, Italian, and Modern Greek. For each language, informants were asked to read a passage from a novel or play. The results did not support the general claim. Inter-stress intervals were no more constant in English than they were in any other language. Furthermore, she claimed that all of these languages are more or less stress-based and

the degree of how much they are stress-based is decided by how much each language depends on stresses. According to the linear graph provided by Dauer (1983), English is the most stress-based language, in which stress plays an important role, while in Japanese, stress is the least important among these six languages. Therefore, her graph indicates that the world's languages can not be classified based on a rhythmic dichotomy.

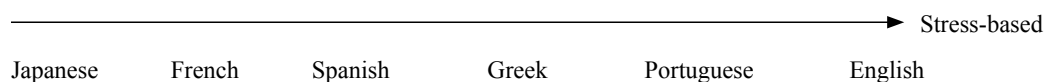


Figure 1: Stress-based languages (Dauer, 1983: 60)

2. 3. Isochrony in Perception

The negative results depending on ISIs (inter-stress intervals) mentioned above led some researchers to suspect that isochrony, rather than being a feature of production, is actually a characteristic of perception (Lehiste, 1977; Donovan & Darwin 1979; Benguerel & D'Arcy, 1986). Among the skeptics, Lehiste (1977) is presumably the most influential. She ran a perceptual experiment to find the noticeable differences in duration. In her experiment, four-foot sentences were chosen and the length of each foot was manipulated. The length of one of the four feet was either decreased or increased by only 10 ms. Thirty listeners were asked to judge which intervals were the longest and the shortest. She discussed the results as follows:

“I reasoned that if listeners cannot identify the actual longest or shortest measures in spoken English sentences, the measures must seem to them to have equal duration. If you cannot tell them apart, they must be alike. Isochrony would then be a perceptual phenomenon.”

(Lehiste, 1977: 256)

Furthermore, she added that:

“The results showed that to get significant agreement among listeners that a given interval was “longest,” an increment was needed that ranged from 30 to more than 100 ms. Differences smaller than 30 ms were never reliably identified. The decrement needed for significant ‘shortest’ judgments ranged also from 30 to 100 ms.”

(Lehiste, 1977: 257)

These statements mean that listeners heard the utterances as more isochronous than they really were. Also, they suggest that the threshold of human perception lies at about 100 ms and thus ISIs below 100 ms are not distinguishable.

Although her findings were from the viewpoint of perception, they persuaded us to further investigate the isochronal tendency in production. Cutler (1980) indicated this point as follows:

“The perceptual reality of isochrony naturally leads one to suspect that there is indeed an underlying regular rhythm in production, and that it is this underlying rhythm which the listener picks up in spite of the multiple perturbations resulting from segmental variations which obscure it in the acoustic signal.”

(Cutler, 1980: 183)

3. Discussion

Although a strict interpretation of isochrony has been rejected through numerous production experiments, many linguists still support the notion of isochrony in English. The isochrony debate is centered around defining the maximum acceptable differences between the durations of feet which are regarded as being isochronal (Yamauchi, 2000).

Based on Lehiste (1977), we propose that 100 ms is the perceptual threshold between isochronal feet, and that differences under 100 ms can be ignored. If that is the case, then the result of O’Connor (1965), for example, would be in favor of isochrony because the difference between the longest and the shortest interval in his experiment was 88 ms, which is below this perceptual threshold.

The number of intervening syllables is another factor which needs to be taken into account. According to Halliday (1985), a two syllable foot is about one fifth as long as a one syllable foot and the relative duration of a foot containing one to four syllables is as follow:

The number of syllables in the foot:	1	2	3	4
The relative duration of the foot:	1	1.2	1.4	1.6

(Halliday, 1985: 272)

A tendency towards isochrony exists because syllables in polysyllabic feet are squeezed together in order to maintain equal spacing of stresses, which is called the “compression effect.” Yet it turns out that

feet are empirically longer when they contain more syllables. We can thus treat these data as evidence for rejecting a strict interpretation of isochrony in English.

Cross-linguistic research is helpful when evaluating the validity of this argument. For example, in a cross-linguistic study in which English (a stress-timed language) and French (a syllable-timed language) were compared, Grosjean & Deschamps (1975) showed that the pause time ratio in each of the two languages is almost identical, but this equal pause time is organized differently in the two languages: in English, many short pauses tend to appear irregularly, whereas in French, pauses are fewer but longer, and in definite positions. This difference between the two languages is reasonable from a phonological point of view because English speakers tend to use pauses as parts of feet to regulate stress-timed rhythm, but French speakers use pauses to avoid disturbing syllable-time rhythm.

4. Conclusion

In English, there is a well-established theory that stressed syllables tend to occur at regular time intervals. This tendency, which is called “isochrony,” has been a great concern among linguists. However, technological progress in recent years has given scholars many modern tools to measure the intervals between successive stressed syllables. As seen in the previous chapter, the results often show considerable variation among the duration of feet, leading many linguists to reject the validity of isochrony in English. Nevertheless, there remain several reasons to suggest that it is unwise to completely reject the notion of isochrony in English feet. Isochrony in English is not only the object of acoustic study, but has also been investigated from phonological, syntactic, and perceptual points of view, and in fact, many studies in these fields have supported the notion of isochrony in English.

References

- Abercrombie, D. (1964). Syllable quantity and enclitics in English. In Abercrombie, D., Fry, D. B., MacCarthy, P. A. D., Scott, N. C. & Trim, J. L. M. (eds.), *In Honour of Daniel Jones: Papers Contributed on the Occasion of His Eightieth Birthday 12 September 1961*, (pp. 216–222) London: Longman.
- _____ (1965). *Studies in Phonetics and Linguistics*. London: Oxford Univ. Press.
- _____ (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh Univ. Press.
- Benguerel, A. & D’Arcy, J. (1986). Time-warping and the perception of rhythm in speech. *Journal of Phonetics*, 14, 231–246.
- Bolinger, D. L. (1951). Intonation: Levels vs. configurations. *Word* 7, 199–210. Reprinted in Abe, I. & Kanekiyo, T. (eds.), *Forms of English: Accent, Morpheme, Order*. Tokyo: Hakuousha.

- _____ (1965). Pitch accent and sentence rhythm. In Abe, I. & Kanekiyo, T. (eds.), *Forms of English: Accent, Morpheme, Order*. Tokyo: Hakuousha.
- Catford, J.C. (1977). *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh Univ. Press.
- Classe, A. (1939). *The Rhythm of English Prose*. Oxford: Basil Blackwell.
- Corder, S. P. (1973). *Introducing Applied Linguistics*. Harmondsworth, Middlesex: Penguin.
- Cutler, A. (1980). Syllable omission errors and isochrony. In Dochert, H. W. & Raupach, M. (eds.), *Temporal Variables in Speech*, The Hague: Mouton, 183–190.
- Dauer, R. M. (1983). Stress-timing and syllable-timing Reanalyzed. *Journal of Phonetics*, 11, 51–62.
- Donovan, A. & Darwin, C. J. (1979). The perceived rhythm of speech. *Proceedings of the Ninth International Congress of Phonetic Sciences*, 2, 268–274. Copenhagen: Institute of Phonetics.
- Gimson, A. C. (1985). *An Introduction to the Pronunciation of English*, 3rd ed. London: Edward Arnold.
- Grosjean, F. & Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français. *Phonetica*, 31, 144–184.
- Halliday, M. A. K. (1967). *Intonation and grammar in British English*. The Hague: Mouton.
- _____ (1970). *A Course in Spoken English*. London: Oxford Univ. Press.
- _____ (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Han, M. S. (1962). The feature of duration in Japanese. *The Study of Sounds*, 10, 65–80.
- Jassem, W., Hill, D. R. & Witten, I. H. (1984) Isochrony in English speech: Its statistical validity and linguistic relevance. In D. Gibbon and H. Richter (eds.), *Intonation, Accent and Rhythm: Studies in Discourse Phonology*, Berlin: Mouton de Gruyter.
- Jones, D. (1918). *An Outline of English Phonetics*. Cambridge: W Heffers & Sons.
- Kubozono, H. & Honma, T. (2002). *Onsetsu to Mora* [Syllables and Moras]. Tokyo: Kenkyusya.
- Ladefoged, P. N. & Johnson, K. (2010). *A Course in Phonetics*, 6th ed. Boston: Heinle.
- _____ (1996). *Elements of Acoustic Phonetics*. Chicago: University of Chicago Press.
- Lea, W. A. (1974). Prosodic aids to speech recognition: A general strategy for prosodically-guided speech understanding. *Univac Report No. PX10791*. St. Paul, Minn.: Sperry Univac, DSD.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Massachusetts: MIT Press.
- _____ (1977). Isochrony reconsidered. *Journal of Phonetics*, 5, 253–263.
- Major, R. C. (1981). Stress-timing in Brazilian Portuguese. *Journal of Phonetics*, 9, 343–351.
- Nakatani, L. H., O'Connor, K. D. & Aston, C. H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica*, 38, 84–106.
- O'Connor, J. D. (1965). The perception of time intervals. *Progress Report*, 2, Phonetics Laboratory, University College, London, 11–15.
- _____ (1967). *Better English Pronunciation*. Cambridge: Cambridge University Press.
- _____ (1968). The duration of the foot in relation to the number of component sound-segments. *Progress Report*, Phonetics Laboratory, University College, London, 1–6.
- Pike, K. L. (1945). *The intonation of American English*. Ann Arbor, Mich.: Univ. of Michigan Pr.
- Roach, P. J. (1982). On the distinction between “stress-timed” and “syllable-timed” languages. In Crystal, D. (ed.), *Linguistic Controversies: essays in linguistic theory and practice in honour of F. R. Palmer*, London: Edward Arnold.

- _____ (2009). *English Phonetics and Phonology: A Practical Course*, 4th ed. Cambridge: Cambridge Univ. Press.
- Shen, Y. & Peterson, G. G. (1962). Isochronism in English. *Studies in Linguistics, Occasional papers*, 9, 1–36. Dept. of Anthropology and Linguistics, Univ. of Buffalo., Buffalo, NY.
- Takebayashi, S. (1996). *Eigo Onseigaku* [English Phonetics]. Tokyo: Kenkyusha.
- Teschner, R. V. & Whitley, M. S. (2004). *Pronouncing English: A Stress-Based Approach with CD-ROM*. Washington, D. C.: Georgetown Univ. Press.
- Uldall, E. T. (1971). Isochronous Stresses in R.P. *Hammerich-Jakobson-Zwirner*, 205–210.
- Yamauchi, H. (2000). An acoustic study of English rhythm: Degrees of isochrony. *English Phonetics*, 3, 207–225.