Phonetica
Editor: K. Kohler, Kiel

Separatum
Publisher: S. Karger AG, Basel
Printed in Switzerland

Phonetica 1989; 46: 197–203

P-Center Judgments Are Generally Insensitive to the Instructions Given

D. H. Whalen[a], *André M. Cooper*[a,b], *Carol A. Fowler*[a,c]

[a] Haskins Laboratories and [b] Yale University, New Haven, Conn., [c] Dartmouth College, Hanover, N. H., USA

Abstract. The perceptual moment of occurrence of a syllable, its P-center, has frequently been examined by instructing subjects to adjust a series of speech sounds until they sounded isochronous. The present two experiments examined the effect of changing the instructions. In addition to the overall isochrony instructions, we asked subjects to align pairs of syllables so that the syllable onsets, vowel onsets, or syllable offsets sounded isochronous. In the first experiment, 3 of 4 subjects showed no difference among the first three instruction sets, and the changes introduced by the fourth went in the wrong direction. All subjects found it impossible to make alignments with respect to offsets. In the second experiment, vowel durations of two versions of some stimuli differed by 100 ms, to enhance the difference in syllable rhyme durations. Two subjects received the same instruction sets as in Experiment 1, and again found alignment with respect to offsets impossible. These subjects showed differences among the other instruction sets, although the direction and magnitude of the differences indicated that they had not succeeded in changing their timing criteria. The results indicate that P-center alignments are the syllable timing judgments that subjects most naturally make, and they may, indeed, be the only isochrony judgments that subjects can make reliably.

Introduction

The perceptual moment of occurrence of a speech sound is not its acoustic onset. Furthermore, the P-center, as it has been named, does not correspond to any single acoustic landmark in the speech signal. This is true of the acoustically defined onset [Morton et al., 1976; Marcus, 1976, 1981; Fowler, 1979], local or global intensity peaks [Allen, 1972; Marcus, 1981], the 'center of gravity' [as defined by Howell, 1984, 1988, and refuted by Fowler et al., 1988], the number of initial consonants [Cooper et al., 1986], or the phonetic quality of the vowel [Fox and Lehiste, 1987]. Attempts to define the P-center in terms of acoustic parameters alone [Marcus, 1976, 1981; Howell, 1984, 1988; Vos and Rasch, 1981] have failed [Fowler et al., 1988].

Nonetheless, some doubts remain about the best interpretation of the findings from the timing studies done to date.

One confusion has arisen from a frequently cited figure of Fowler and Tassinary [1981]. In that study, speakers produced monosyllables to the beat of a metronome. The phonetic identity, of the initial consonants was varied, while the syllable rhyme (in this case, /ɑd/) was the same throughout. When the syllables were displayed so that their position relative to the metronome pulse lined up, the vowel offsets happened to line up closely as well. During presentations of our work, several colleagues at Haskins Laboratories speculated that the judgments might, in fact, be about syllable offsets rather than onsets. Since changes in the duration of the syllable rhyme can change P-center judgments, that speculation is incorrect [Marcus, 1981; Fox and Lehiste, 1987; Cooper et al., 1988: also see Fowler and Tassinary, 1981, pp. 530–531]. Nevertheless, we began to wonder whether subjects did, or even could, use different alignment strategies in the method-of-adjustment technique. The present experiments were designed to see whether subjects can align syllables isochronously with respect to points other than the P-center.

Three line-up points other than the P-center were selected. The first was the acoustic onset of the syllable. Perhaps if subjects are instructed to line them up, they could do so, regardless of the rest of the syllable. The second was the vowel onset, which is the acoustic landmark whose variation in timing with respect to the onset of the preceding consonants is most like that of the P-center. The third, already discussed, was the syllable offset.

If the different sets of instructions given to subjects result in different timing alignments, then the P-center literature must be interpreted cautiously. Not only would the exact instructions given to subjects be seen as crucial for determining the correct interpretation of the results, but it would be clear that timing patterns other than P-center alignment could have been studied. Alternatively, if the different instructions do not affect subjects' timing judgments, then the interpretations of the previous literature would be bolstered. Not only would the exact nature of the instructions be irrelevant, but the notion that the P-center is influenced by the entirety of the speech signal [Marcus, 1981; Cooper et al., 1986] would receive further support.

Experiment 1

Subjects were asked to align syllable pairs so that the whole syllable or one of three acoustic events was isochronous. The alignments were made by modifying the intervals between two syllables in continuous alternation.

Method

Stimuli. The syllables used were /bɑd/, /tɑd/, and /ʃɑd/. These were selected because they have three distinct P-centers, when measured from the acoustic onset [Fowler and Tassinary, 1981]. Several examples of each syllable were recorded on audio tape. These were then digitized (10-kHz sampling rate, 12-bit resolution), and one token of each was selected for use in the experiment. The durations were 414, 414, and 549 ms for /bɑd/, /tɑd/, and /ʃɑd/, respectively. The durations of the voiceless segments at the beginning of the syllable (either burst plus aspiration or fricative noise) were 13, 66, and 179 ms, respectively.

Subjects. Three of the subjects (AMC, CAF, and

MRS) had participated in other timing studies [Cooper et al., 1986, 1988]. The fourth subject (HE) was a young male volunteer, who was paid for his services. The first two subjects were two of the authors, and the last two were naive to the purpose of the experiment. All subjects were native speakers of some variety of American English, though HE had spent his childhood in Jamaica.

Equipment and Procedure. Each syllable was paired with each of the other two, giving six permutations. For each pair, the first syllable presented was considered the standard, and the other syllable's timing could be changed relative to it. At the beginning of each trial, there was a 50-ms gap between the offset of the first syllable and the onset of the second; the interval between the offset of the second syllable and the onset of the next repetition of the first was the longest possible interstimulus interval (ISI). The subjects would then modify this ISI in accordance with the instructions given.

Two sets of equipment were used. Each involved computer control of the ISI, but one used keyboard presses to change the ISI (used by CAF and MRS), and the other used an analog potentiometer (for AMC and HE). The other difference between the two setups was in the interval between the onset of like syllables. For the first system this interval was the duration of those syllables plus 550 ms. For the second system, there was a constant 1,400 ms from the onset of the fixed syllable to the next repetition of this syllable.

Four blocks, each with a different set of instructions, were run for each subject. The instructions were to adjust the timing of the syllables until: (1) the whole syllable sounded isochronous as if timed to the beat of a metronome, (2) their onsets were isochronous, (3) their vowel onsets were isochronous, or (4) their offsets were isochronous. All subjects received the blocks of instructions in this order. The syllable pairs were presented in random order. Subjects could take as long as they needed to make their adjustments. There were four repetitions of each of the six pairings for CAF and MRS, and six for AMC and HE.

Results

The manipulation of syllable order within a trial (e.g. /bɑd/-/ʃɑd/ versus /ʃɑd/-/bɑd/) did not affect the results, so all measurements were adjusted to ignore this factor. Each alignment value reflects the divergence away from the midpoint of the /bɑd/-to-/bɑd/ interval (or the /tɑd/-to-/tɑd/ interval for trials without /bad/ stimuli). Positive numbers show that the /tad/ or /ʃɑd/ syllable's onset was earlier than this midpoint, reflecting agreement with P-center measurements made in the past.

Table 1. Mean timing alignments for four subjects in three instruction conditions, Experiment 1

Pair	Instruction set		
	P-center	Acoustic onset	Vowel onset
b-ʃ	122	131 (0)	124 (166)
b-t	42	39 (0)	33 (53)
t-ʃ	75	95 (0)	73 (113)

Numbers in parentheses are the values that should have been obtained if the subjects had been able to do the task.

Statistical analyses were performed only on the first three instruction sets. None of the four subjects completed the offset alignment task, claiming that they had no idea what they were doing. We suspected that the extreme difficulty of the task might have been due to the lack of variation in the durations of the syllable rhymes, so we did not interpret this refusal further.

The mean measurements for the three pairings in the three instruction sets are shown in table 1. An analysis of variance with the factors Instruction Set, Subject, and Pairing was run on the individual trial values. There was no main effect of Instruction Set ($F(2,276) < 1$); however, effects of Subject and Pair were significant ($F(3,276)$

= 7.67, $F(2,276) = 34.96$, $p. < 0.001$, respectively). Interactions between Instruction Set and Subject, and between Subject and Pair were also significant ($F(6,276) = 2.89$, $F(6,276) = 8.70$, $p. < 0.01$ for the two, respectively). Individual subject analyses showed that subject HE had a significant effect of Instruction Set ($F(2,63) = 3.46$, $p. < 0.05$). However, his means for the acoustic onset condition were larger than for the P-center and vowel onset conditions (144 vs. 83, and 89 ms), indicating that he introduced *more* deviation from acoustic isochrony when he attempted to align acoustic onsets.

As a final check to ensure that the amount of variability in the data was not so extreme as to obscure a difference between the P-center and acoustic onset conditions, we modeled perfect performance on the acoustic onset condition by subtracting from each value the mean for that pair for that subject. Thus the means for that condition came out to zero, which would have indicated perfect alignment of the acoustic onsets. If that had been the case, then the instruction set would have had a significant effect ($F(1,168) = 50.11$, $p. < 0.001$). Thus if there had been a sizable effect, this amount of data would have discovered it. The one subject who showed an effect showed it in the wrong direction.

Discussion

None of the subjects succeeded at all in aligning the syllables with respect to acoustic onsets. All but one subject made statistically indistinguishable temporal alignments for each of the other three instruction sets. In fact, the one subject who showed a difference in the instruction sets tended to push the syllables farther from acoustic onset isochrony in his 'acoustic onset' alignments than in his P-center alignments. Even the vowel onsets were more accurately aligned (though not significantly more) in the P-center instruction condition than in the vowel onset condition. In short, no matter what the instructions were, subjects generally seemed to base their timing judgments on the P-center.

Experiment 2

In Experiment 1, subjects may have found the syllable offset alignment task difficult to perform because correct alignment of the offsets would make the P-centers very close as well. Thus it may have been difficult to ignore the more robust P-center alignment and focus in on the more difficult offset alignment, despite the instructions. In Experiment 2, larger variations in the duration of the vocalic segment were introduced into some of the syllables in an effort to make the syllable offset instructions easier to carry out.

Method

Stimuli. Several tokens of /bɑ/, /tɑ/, and /ʃɑ/ were recorded on audio tape and digitized as described for Experiment 1. One token of each type was chosen such that all the vocalic segments were roughly equal in intensity and fundamental frequency. The original durations were 414, 412, and 538 ms for /bɑ/, /tɑ/, and /ʃɑ/, respectively. The onset-to-vocalic-segment durations were 27, 79, and 157 ms for /bɑ/, /tɑ/, and /ʃɑ/, respectively. Two additional stimuli were made from /bɑ/ and /ʃɑ/. A shortened version was made by deleting whole pitch periods until approximately 50 ms had been removed. A lengthened version was made by repeating the same pitch periods that had been removed from the shortened versions. Thus shortened /bɑ/ was 364 ms in duration, and lengthened /bɑ/ was

464 ms. Shortened /ʃɑ/ was 486 ms, and lengthened /ʃɑ/ was 590 ms.

Subjects. Two subjects participated in this experiment. They were AMC, who had participated in the previous experiment, and a young male research assistant (RFP), who had participated in a previous study [Cooper et al., 1986], but was naive to the purpose of the present study. Both were native speakers of American English.

Equipment and Procedure. The set-up with the keyboard response was used. For this experiment, however, all of the stimuli were adjusted within a 1,400-ms window regardless of the duration of the stimuli.

To reduce the number of judgments that had to be made, only some combinations of stimuli were used. The shortened and lengthened versions of /bɑ/ were paired with original /tɑ/ and /ʃɑ/, to give four of the pairs. The shortened and lengthened versions of /ʃɑ/ were paired with original /tɑ/, to give the other two. The same instructions were given as in Experiment 1. AMC gave six judgments on each pair with each syllable starting the series, and RFP gave eight.

Results

The greater variation in the stimulus durations in this experiment was not enough to make the subjects confident at all in their offset alignments. Subjects had such difficulty in performing the offset alignment task that they eventually gave up. One subject (AMC) went through much of this condition, but quit when he discovered that his last three answers had been ones which allowed the minimum (50 ms) of silence between one syllable and the next, i.e., the equivalent of the starting position. No data will be reported from this instruction set.

The measurements were transformed as in Experiment 1, with positive numbers indicating that /tɑ/ and /ʃɑ/ had acoustic onsets earlier than /bɑ/. For the manipulation of vowel duration, this had the unfortunate effect that the predicted difference went in one direction for /bɑ/ and in the other direction for /ʃɑ/. Thus the effect of shortening or lengthening the vocalic segment shows up first as an interaction between the Duration factor and the Pair factor.

Table 2. Mean timing alignments for two subjects in three instruction conditions, Experiment 2

Pair	Instruction set		
	P-center	Acoustic onset	Vowel onset
b(shortened)-ʃ	156	103 (0)	137 (130)
b(lenghtened)-ʃ	133	105 (0)	122 (130)
b(shortened)-t	22	26 (0)	14 (52)
b(lengthened)-t	13	28 (0)	4 (52)
t-ʃ(shortened)	106	81 (0)	92 (78)
t-ʃ(lengthened)	132	104 (0)	126 (78)

Numbers in parentheses are the values that should have been obtained if the subjects had been able to do the task.

The results are shown in table 2. An analysis of variance was performed on the individual trial values with Instruction Set, Subject, Pair, and Duration as the factors. The first three main effects were significant ($F(2,468) = 11.46$, $p < 0.001$; $F(1,468) = 7.14$, $p < 0.01$; $F(2,468) = 378.55$, $p < 0.001$, respectively). As expected, Duration was not significant as a main effect ($F(1,468) = 0.85$) but only as an interaction with Pair ($F(1,468) = 14.15$, $p < 0.001$). There were further interactions of Instruction Set and Subject, ($F(2,468) = 9.65$, $p < 0.001$), Instruction Set and Pair, ($F(2,468) = 8.49$, $p < 0.001$), and Subject and Pair ($F(2,468) = 3.83$, $p < 0.05$).

Tests of each of the three pairings of the instruction sets revealed that each was distinct from the other, with grand means of 93, 74, and 82 ms for P-center, acoustic on-

set, and vowel onset instructions, respectively. The acoustic onset adjustments were generally closer to physical isochrony than the P-center adjustments were. However, the acoustic adjustments were still much closer to the P-center values than they were to physical isochrony. The vowel onset values were close to the physical target of 87 ms. However, other aspects of the data complicate the picture that these mean values depict. The effect of syllable duration on the isochrony judgments was roughly equivalent for the P-center and vowel onset instructions. If the subjects had succeeded in aligning the syllables by the vowel onset, the duration of the vocalic segment should not have had any effect on the timing judgments. Further, a test of the variances in the syllable onset and vowel onset conditions revealed that P-center alignments accounted for the data better than the physical targets for those conditions. For each condition, we generated two sets of variances. One was based on the P-center alignment for that cell, the other, on the physical target value for the condition. These physical target values were 0 ms for the syllable onset condition and, in the vowel onset condition, 130, 52, and 78 ms for /bɑ/-/ʃɑ/, /bɑ/-/tɑ/, and /tɑ/-/ʃɑ/, respectively. A test of the absolute values of these differences showed that the judgments were less different from the P-center alignments than they were from the physical targets in both the syllable onset (39 ms vs. 80 ms, $F(1,156) = 73.20$, $p. < 0.001$) and the vowel onset conditions (29 ms vs. 39 ms, $F(1,156) = 37.23$, $p. < 0.001$). Whatever it was that subjects were doing differently, it was not likely to have been finding the syllable or vowel onset. Though the onset instructions elicited judgments significantly different from the P-center instructions, these judgments were still significantly closer to P-center values than to their respective physical targets. Further, there were some cases where the vowel onset judgments were closer to the acoustic onset physical target than the vowel onset physical target (/bɑ/-/tɑ/), and vice versa (/tɑ/-/ʃɑ/).

Discussion

The fact that subjects were unable to align syllables with respect to vowel offsets is a strong indication that they did not do so in earlier experiments. This is supported by the fact that the 100 ms change in the duration of the stimuli in this experiment resulted in an overall shift of 15 ms in the timing judgments, rather than none or 100 ms. This is consistent with the small effect of changes in the vocalic segment on the P-center obtained in other research [cf. Marcus, 1976, 1981; Cooper et al., 1988].

Although the subjects in this experiment succeeded in differentiating the three remaining instruction sets, this result does not suggest that subjects were able to align the syllables with respect to acoustic onsets or vowel onsets. That is, the fact that something was done differently did not imply necessarily adjusting the syllables in accordance with the instructions.

Conclusions

Timing judgments for speech stimuli are generally insensitive to the instructions given. Subjects tend to be unable to adopt a specific alignment strategy other than the one they adopt spontaneously. In the first experiment, we found that only one subject showed any differences in alignment under

three sets of instructions, and that was in the wrong direction. In the second experiment, subjects did make distinctions among the instruction sets, but only occasionally in the direction that was appropriate. In both experiments, subjects could not align syllable offsets. We are thus led to two conclusions:

(1) The previous literature on speech timing, especially that using the method of adjustment, is indeed about the timing of P-centers, regardless of the exact nature of the instructions that were given. A corollary of this is that the close alignment of the acoustic offsets in figure 1 of Fowler and Tassinary [1981] is an artifact of the close match in syllable rhyme durations used there.

(2) Not only are there no obvious acoustic markers for the P-center in speech, listeners are unable to make conscious timing judgments based on the acoustic markers of syllable onset, vowel onset, and syllable offset. Instead, the P-center seems to be the only perceptual event at which syllables can be aligned to isochrony.

Acknowledgments

This research was supported by NIH Grant HD 16591, NIH Grant HD 01994 and NSF Grant BNS 8111470 to Haskins Laboratories. We thank Michael Studdert-Kennedy, Steven Marcus, and an anonymous reviewer for helpful comments.

References

Allen, G.: The location of rhythmic stress beats in English: an experimental study. Part I. Lang. Speech, *15:* 72–100 (1972).

Cooper, A. M.; Whalen, D. H.; Fowler, C. A.: P-centers are unaffected by phonetic categorization. Percept. Psychophys. *39:* 187–196 (1986).

Cooper, A. M.; Whalen, D. H.; Fowler, C. A.: The syllable's rhyme affects its P-center as a unit. J. Phonet. *16:* 231–241 (1988).

Fowler, C. A.: 'Perceptual centers' in speech production and perception. Percept. Psychophys. *25:* 375–388 (1979).

Fowler, C. A.; Tassinary, L. G.: Natural measurement criteria for speech: The anisochrony illusion; in Long, Baddeley, Attention and Performance, vol. 9, pp. 521–535 (Erlbaum, Hillsdale, 1981).

Fowler, C. A.; Whalen, D. H.; Cooper, A. M.: Perceived timing is produced timing: a reply to Howell. Percept. Psychophysics *43:* 94–98 (1988).

Fox, R. A.; Lehiste, I: The effect of vowel quality variations on stress-beat location. J. Phonet. *15:* 1–13 (1987).

Howell, P.: An acoustic determinant of perceived and produced anisochrony; in van den Broecke, Cohen, Proc. 10th Int. Congr. Phonet. Sci., pp. 429–433 (Foris, Dordrecht, 1984).

Howell, P.: Prediction of P-center location from the distribution of energy in the amplitude envelope. Part. I. Percept. Psychophys. *43:* 90–93 (1988).

Marcus, S. M.: Perceptual centres; unpubl. doct. thesis Cambridge (1976).

Marcus, S. M.: Acoustic determinants of perceptual center (P-center) location. Percept. Psychophys. *30:* 247–256 (1981).

Morton, J.; Marcus, S.; Frankish, C.: Perceptual centers (P-centers). Psychol. Rev. *83:* 405–408 (1976).

Vos, J.; Rasch, R.: The perceptual onset of musical tones. Percept. Psychophys. *29:* 323–335 (1981).

Received: January 3, 1989
Accepted: June 22, 1989

D. H. Whalen, Ph.D.
Haskins Laboratories
270 Crown St.
New Haven CT 06511 (USA)