ON THE INFLUENCE OF THE INTERNAL STRUCTURE OF A SYLLABLE ON THE P-CENTER-PERCEPTION

Peter. M. Janker
Forschungsschwerpunkt Allgemeine Sprachwissenschaft, Berlin,
Institut für Phonetik und Sprachliche Kommunikation der LMU, München, Germany

ABSTRACT

The two experiments described here investigate whether or not the internal structure of a syllable has an influence on the p-center-perception. Subjects had to perform a synchronisation task by tapping to sequences of German monosyllabic words with either different nuclei or varying complexity within the syllable shell.

INTODUCTION

In recent years various investigations [1-7] have been undertaken to gain knowledge about the parameters influencing the 'moment of occurrence' or the so called p-center [8]. Some of the models proposed suggest that the actual acoustic make-up of the phonological segments of a syllable is responsible for its p-center position hence the internal structure is a parameter which should not be neglected. Evidence for the assumptions made is mainly taken from experiments with synthesized artificial sound or speech stimuli.

EXPERIMENTAL DESIGN

To test whether or not the internal structure or complexity of the syllable has an influence on the p-center position two synchronisation experiments with naturally spoken stimulus material were carried out.

Stimulus material

For experiment one (VQ) a set of monosyllabic stimuli with phonologically identical shell but different nuclei was produced. To build the stimuli having a short vowel (abrupt cut) the German words <Still, Stell, Stall> [ʃtɪl, ʃtɛl, ʃtal], for the stimuli with a long vowel (smooth cut) the German words <Stil, Stehl, Stahl> [ʃti:l, ʃte:l, ʃta:l] were used.

The overall duration of the stimuli with smooth cut is approximately 80 ms longer than that of the stimuli with abrupt cut despite some compensational coda shortening in the case of smooth cut.

For experiment two (CS) a set of monosyllabic stimuli with increasing complexity in head and coda but phonologically identical nucleus was produced with the German words <Schal, Stahl, Strahl, Schalt, Strahlt, Schalst, Stahlst, Strahlst> [ʃa:l, ʃta:l, ʃtra:l, ʃa:lt, ʃtra:lt, ʃa:lst, ʃta:lst, ʃtra:lst].

The overall length of the stimuli varied between about 400 ms and 620 ms with a stronger tendency for compensatory shortening / lengthening with increasing / decreasing complexity in the head.

All recordings were performed in a soundproofed studio using an Electro Voice 631B microphone and a Sony DAT recorder. All words were well pronounced (explicitly demonstrated) in focus position within the frame sentence <Ich habe das Wort _____ gesagt.> (I said the word _____.)[1]. The recordings were transmitted via Digidesign AudioMedia II and then segmented and downsampled to 20 kHz using Signalyze on the Macintosh.

Method

These stimuli, as well as a control stimulus (click signal: 5 ms, 1 kHz tone burst), were presented binaural using a Sennheiser HD 250 headphone under computer control (DEC VaxStation VS

1) n.b. that the position in the German phrase is not sentence final.

3200, Distec DA-converter, Krohn-Hite 3750 filter) with 20 kHz sample rate and lowpass filtered at 6 kHz (24 dB/oct).

30 subjects had to listen and tap in synchrony to sequences built of 15 repetitions of the same stimulus with an inter stimulus interval of 700 ms and an inter sequence interval of 1400 ms. The stimulus sequences, chosen in random order regardless of the experiment they belonged to were grouped in blocks of 10. The subject starts the presentation of the next block by pressing the return key. Each stimulus sequence was given four times with at least two different intermediate sequences. A sequence was repeatedly presented as long as the subject did not start to tap. To register the taps a 5 x 10 cm capacitive sensory field was used. Before the presentation of the target stimuli subjects were familiarized with the data acquisition procedure using the click signal, the sound [pst] and the word <Schwimmst> [ʃβɪmst] as stimulus material.

Overall 25200 taps were registered. For analysis the taps to the first three and the last two presentations within a sequence were omitted (leaving 16800).

Subjects

30 subjects (13 female, 17 male) took part in the experiments. All of them had a 10 minute introduction on using the computer and the stimulus presentation program, none of them had participated in a former experiment on rhythm perception.

RESULTS

The data showed a large intersubject variability but according to 'Duncan's multiple range test' out of the 30 subjects 21 had been able to perform the experimental task as intended showing a low intrasubject variability of the tapping positions for the respective stimulus and not having unusual values for skewness and kurtosis.

The intersubject variability can be seen in Figure 1 which shows the tapping positions for the click stimulus. The control

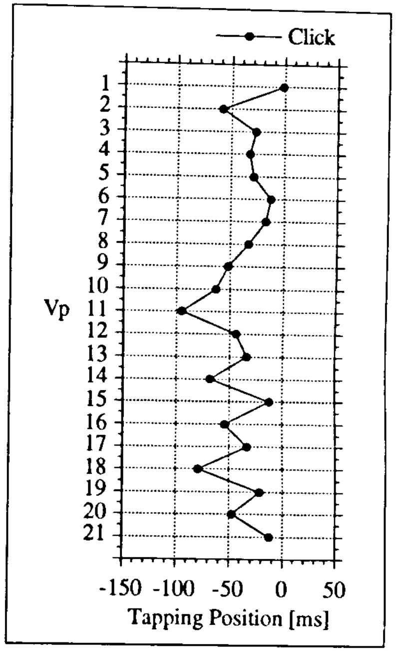

Figure 1: Tapping positions for the control stimulus (click signal) showing the amount of anticipation and the intersubject variability.

stimulus had been presented to be able to compensate for the known effect of anticipation common in repetitive tapping task experiments. On average, the subjects tapped 39,27 ms before the physical onset of the click, which is in good agreement with findings reported elsewhere [9-12].

In Figure 2 the measured as well as the neutralized (anticipation corrected) tapping positions are given in relation to the durations of syllable head, nucleaus and coda for both experiments. An effect of the different stimuli on the location of the tapping position in relation to the stimulus onset can clearly be seen. This effect, however, may simply be caused by the different durations of the stimuli.

Experiment VQ

The stimuli for experiment one (VQ) were chosen to show whether or not the

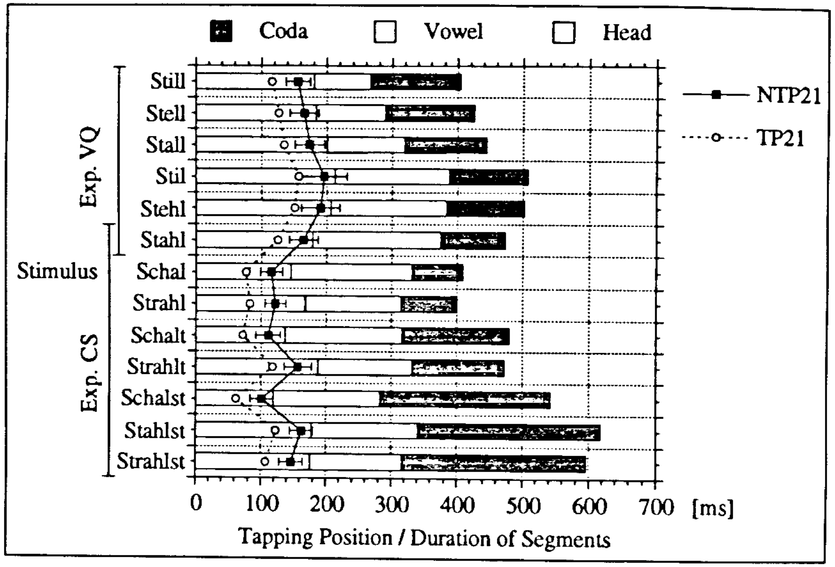

Figure 2: Mean (TP21) and neutralized (anticipation corrected) tapping position (NTP21 with SD) in relation to the stimulus onset with indicated duration of head, nucleus vowel and coda for the stimuli of experiments VQ and CS.

varying influence of different energy distributions known from experiments with synthetic stimuli can be replicated with naturally spoken stimulus material.

However, there seem to be no evidence that the differences in the abruptness of the vowel ending (i.e. [ʃtal] vs. [ʃta:l]) or vowel quality (order: [ʃte:l] -> [ʃti:l] vs. [ʃtɪl] -> [ʃtɛl], second with later tapping position) are responsible for any differences in the measured tapping positions, although there is a slight tendency for the smooth cut stimuli - which are also longer - to show later ones.

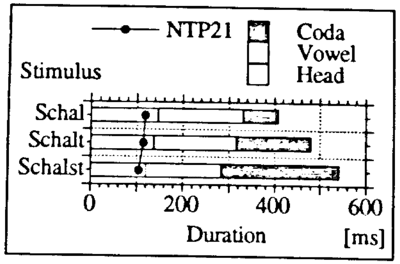

Figure 3: Variation of the coda showing no influence on the tapping position

Therefore this experiment with naturally spoken stimuli with identical shell does not replicate the findings with synthetic stimuli and holds no evidence that a different vowel quality or vowel quantity (abruptness) is in itself - in naturally spoken stimuli accompanied by compensational shortening or lengthening - of any influence on the p-center perception.

Experiment CS

The second experiment (CS) was carried out to investigate whether or not the internal structure of the head and the coda of a natural spoken syllable is of influence on the p-center-perception.

In accordance with the literature the duration of the initial consonance (head) clearly had an influence on the tapping position. However, looking at figure 2 the measured tapping positions do not indicate that the compositional structure of the head is of importance. This is also the case for the internal structure of the coda. Furthermore and in opposition to the literature the greatly varying duration of the

coda in [ʃa:l, ʃa:lt, ʃa:lst] shows no influence at all (Figure 3).

Therefore experiment CS does not show any evidence that the complexity of the syllable shell is of importance for the p-center-perception.

CONCLUSION

Overall there is no evidence for any influence of the internal structure of a syllable on the p-center-perception. With respect to the used natural spoken stimuli the neutralized tapping positions closely follow the consonance-vowel transition, thus the duration of the initial consonance still seems to be the best reference to determine the location of the p-center.

REFERENCES

[1] Howell, P. (1988), "Prediction of p-center location from the distribution of energy in the amplitude envelope: Part I & II", *Perception & Psychophysics*, vol. 43, pp. 90-93 & 99.

[2] JANKER, P. M. (1989), "Der Einfluß von Segmentdauer- und Amplitudenmanipulation auf die P-center-Position einfacher CV-Silben", *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, vol. 27, pp. 71-141.

[3] JANKER, P. M., POMPINO-MARSCHALL, B. (1991), "Is the p-center influenced by 'tone'?", *Proc. 12th ICPhS* Aix-en-Provence, vol. 3, pp.290-293

[4] KÖHLMANN, M. (1984), "Rhythmische Segmentierung von Sprach- und Musiksignalen und ihre Nachbildung mit einem Funktionsschema", *Acoustica*, vol. 56, pp. 192-204.

[5] MARCUS, S. M.(1981), "Acoustic determinants of perceptual center (P-center) location", *Perception & Psychophysics*, vol. 30(3), pp. 247-256.

[6] POMPINO-MARSCHALL, B. (1989), "On the psychoacoustic nature of the P-center phenomenon", *Journal of Phonetics*, vol. 17, pp. 175-192.

[7] POMPINO-MARSCHALL, B. (1990), *Die Silbenprosodie. Ein elementarer Aspekt der Wahrnehmung von Sprechrhythmus und Sprechtempo*, Tübingen: Niemeyer.

[8] MORTON, J., MARCUS, S.M., & FRANKISH. C. R. (1976), "Perceptual centres (P-centers)", *Psychological Review*, vol. 83, pp. 405-408.

[9] ASCHERSLEBEN, G., PRINZ, W. (1992), "What gets synchronized with what in sensorimotor synchronisation", *Paper 13/1992*, MPI für psychologische Forschung, München.

[10] DUNLAP, K. (1910), "Reactions on rhythmic stimuli, with attempt to synchronize.", *Psychological Review*, vol. 17, pp. 399-416.

[11] RADIL, T. et al. (1990), "Stimulus anticipation in following rhythmic acoustical patterns by tapping.", *Experimentia*, vol. 46, pp. 762-763.

[12] JANKER, P. M. (1993), *Sprechrhythmus, Silbe, Ereignis. Eine experimentalphonetische Untersuchung zu den psychoakustisch relevanten Parametern zur rhythmischen Gliederung sprechsprachlicher Äusserungen*, Dissertation, LMU München.