COMPARISON OF SOUND PRESSURE LEVEL AND LOUDNESS(GF) MEASUREMENTS ON SPEECH SOUNDS

Carl Ludvigsen and Nina G. Thorsen

1. Introduction

Lately, an instrument called a loudness analyzer (HP 8051A Loudness Analyzer), which operates according to Zwicker's method of estimating loudness, has been made commercially available.

Zwicker's method is based on a division of the signal in critical bands. In each critical band the signal is weighted according to the sensitivity of the ear. The contribution of the sub-signals are added up, allowance being made for masking phenomena. - The output of the loudness analyzer is called loudness(GF)*and is measured in sones(GF). "G" refers to critical bands (Frequenzgruppe) and "F" refers to a frontal sound field.

The aim of this investigation was to compare sound pressure level (SPL) measurements with loudness(GF) measurements on some consonants and vowels.

The investigation must be regarded as a pilot experiment since only two subjects were involved and the material was recorded once only, by each subject.

2. The material

The material consisted of isolated bisyllabic words of the CV:Cə type with trochaic stress pattern, which is a common type in Danish. C represents s, f, n, m, and V represents i e ɛ a y ø œ u o ɔ ɑ ɒ. The consonants and

---

*) It is also possible to get an output called loudness(GD), where "D" refers to a diffuse sound field.

vowels were combined in all possible combinations, forming a list of 192 different words. - The material was originally designed for a different purpose, namely an examination of the relations between fricatives, nasals, and vowels. We found it suitable for the present purpose too, since the unvoiced fricatives, the nasals, and the vowels are three acoustically distinct classes of speech sounds with very different spectra.

The words were recorded once by two subjects, one male (CL) and one female (NT), (the authors), who both speak a form of advanced Standard Copenhagen Danish. - We attempted to produce the words at a comfortable constant level, but we cannot exclude that our knowledge of the purpose of the experiment has influenced the reading of the words. This question will be treated again briefly in section 5.1.1.

3. Instrumental set-up

The recordings were made in the sound studio at the Institute of Phonetics. The subjects were placed comfortably at a distance of approximately 1 meter from a free field response microphone (B&K, type 4131).

The response of a free field microphone is adjusted so as to deliver a signal which corresponds to the SPL value obtained if no microphone were present. In the B&K, type 4131 microphone this is accomplished by a damping of the frequencies above 2000Hz of approximately 6dB/octave. This characteristic of the microphone has been calculated for use in free, plane sound fields, 0^{o} incident, and this situation is approached by placing the microphone at a distance of 1 meter from the subject.

The signal from the microphone was picked up on a tape recorder via a microphone amplifier (B&K, type 2603).

A pure tone of 1000Hz preceded and followed the recordings for the sake of calibration.

3.1. Processing of the speech signals

3.1.1. Tracing of words

The signal from the tape recorder was fed into different analyzing devices, the outputs of which were traced on an ink writer[1]. The analyzing devices were

(1) fundamental frequency detector
(2) intensity meter
(3) loudness analyzer.

3.1.2. Tracing of isolated speech sounds

The peak mode operation of the loudness analyzer implies that a segment with a small loudness(GF) value immediately preceded by a segment with a larger loudness(GF) value will not cause any further deflection of the loudness curve. - To render possible loudness(GF) measurements of such weaker segments all segments, except the schwa vowels, were isolated on a segmentator and recorded on tape. The segmentator is described by Thorvaldsen (1970). The rise and fall times of the segments were set at 5ms. - The segments were then analyzed in the same way as the words (cf. 3.1.1.).

4. Calculations

The data were subjected to statistical treatment. Four variables were included in the process

(1) the duration of each segment
(2) the maximum SPL of each segment
(3) the maximum loudness(GF) of each segment

---

1) Instruments which are not further specified are listed in ARIPUC vol. 4, 1969, p. II-V.

(4) the $\log_2$ of the maximum loudness(GF) of each segment.

The total population of sound segments analyzed was treated in various sub-populations, as well as pooled. This was done in order to watch the variation in the SPL and loudness(GF) measurements in vowels in different surroundings before pooling these measurements.

The statistical treatment involved calculations of mean, standard deviation, range, skewness, mean error, and 95% and 99% confidence limits. Each subject was treated separately.

5. Results

5.1. The vowels

(1) A cursory glance at the tracings of the whole words tells us that the course of the loudness(GF) and SPL curves are not always correlated. - In most words the loudness(GF) and SPL maxima were distributed in the same way over the two syllables, i.e. either the full vowel had the greater loudness(GF) and SPL values, or both maximum values were reached in the unstressed vowel. - However, there were instances where the maximum value of loudness(GF) in a word was reached during the unstressed schwa, but the maximum SPL value was measured in the stressed vowel, and vice versa.

(2) In Table I are given the means of the duration, SPL, and loudness(GF) values for the (isolated) vowels.

Table I

Means of duration, SPL, and loudness (GF) for 12 vowels, all surroundings pooled, 16 examples of each vowel. Two subjects, NT and CL.

	duration [ms]		SPL [$dB/2x1o^{-5}N/m^2$]		loudness(GF) [sones(GF)]	
	NT	CL	NT	CL	NT	CL
i	2o9	2o6	74,6	73,1	1o,7	1o,8
e	251	225	7o,2	74,5	1o,9	14,2
ɛ	264	251	68,5	72,3	15,6	15,3
a	268	266	68,4	72,3	17,9	16,9
y	241	218	75,3	74,4	12,o	11,2
ø	26o	239	7o,3	74,1	12,1	14,3
œ	273	26o	69,4	71,9	17,2	15,2
u	256	22o	74,1	73,8	1o,3	1o,o
o	264	251	7o,4	72,1	1o,1	1o,8
ɔ	268	269	7o,o	7o,8	15,1	12,3
ɑ	288	281	71,6	69,9	23,1	16,7
ɒ	279	262	7o,4	7o,9	16,4	12,4

The SPL and loudness(GF) means were subjected to a ranking correlation test. The ranks (1, 2, ...) were assigned to each value according to its magnitude.

(a) The ranks of the two subjects' SPL means were not significantly correlated. The Spearman rank correlation coefficient, r_s, was calculated to be o,2o. In order for the rank correlation to be significant at the 95% level it should exceed o,5o.

(b) However, the ranks of the two subjects' loudness(GF) means were significantly correlated. The r_s value was calculated to be o,86. In order for the correlation to be significant at the 99% level it must exceed o,7o, only. - This must be due to the fact that the SPL measures depend on the total power of the sound, whereas the loudness(GF) measures are dependent also on the shape of the spectrum, which dependency has been decisive for the output of the loudness analyzer.

(c) The ranks of the loudness(GF)/SPL means for the subject NT were negatively correlated, r_s = -o,52, which exceeds the demand for significance at the 95% level. The same ranks for the subject CL were likewise negatively correlated, r_s = -o,3o, which value is not significant at the 95% level.

5.1.1. Vowel SPL and loudness(GF) vs. F1-frequency

The SPL means in Table I are displayed graphically in Fig. 1. It will be seen that in most vowels small SPL corresponds to high F1. Correlation calculations yielded: The rank correlation between SPL means and F1-frequency for both subjects was found to be negative at a level of significance greater than 95%[2]. This is surprising in the light of earlier investigations, cf. for instance House, Fairbanks and Stevens (195o) and Lehiste and Peterson (1959). The reason may very well be found in the reading of the words. For example, in trying to produce all the words with the same strength we may subconsciously have compensated for the smaller SPL of the close vowels. Furthermore, the results we obtained gave rise to a renewed inspection of the word lists and the tape. This revealed that the originally intended randomization had failed in certain respects. - We cannot exclude that a certain rhythmic effect may have been active, but a phonetically trained listener did not detect

---

2) Vowels of approximately the same F1 were tied in four groups (iyu)(eøo)(εœɔ)(aɑɒ). These ties were incorporated in the ranking tests.

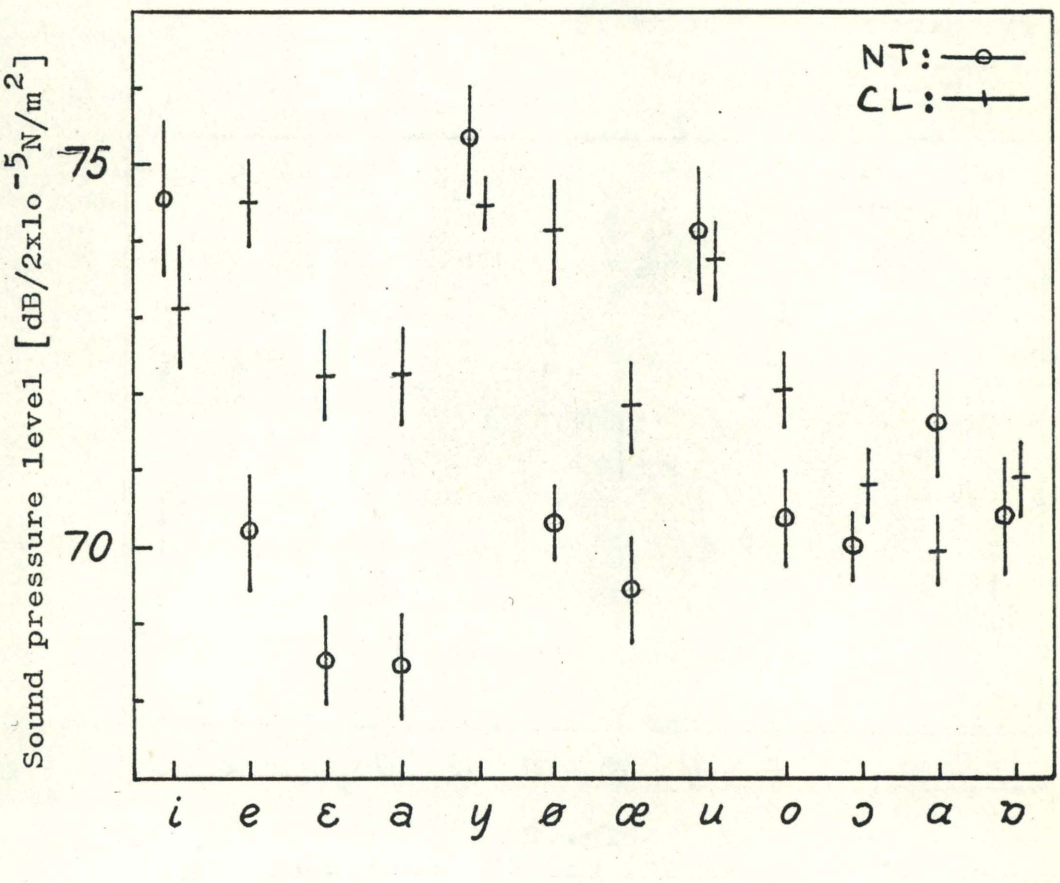

Fig. 1
Means of SPL and 95% confidence limits for 12 vowels, all surroundings pooled, 16 examples of each vowel for each of the two subjects, NT and CL.

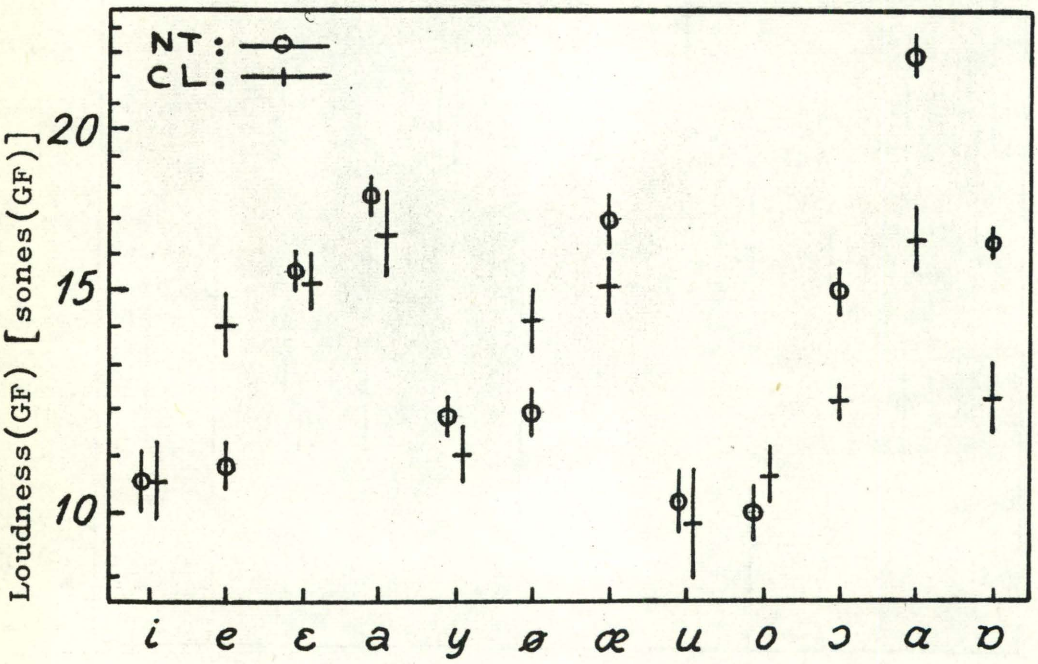

Fig. 2
Means of loudness(GF) and 95% confidence limits for 12 vowels, logarithmic scale, all surroundings pooled, 16 examples of each vowel for each of the two subjects, NT and CL.

any particular rhythm during a casual listening to the tape. Thus, our results cannot invalidate what is a very common experience, namely that vowels with low F1 have smaller SPL than vowels with high F1.

On the other hand, the means of the loudness(GF) obviously rise with rising F1, cf. Fig. 2. We found a significant positive correlation between loudness(GF) and F1. For both subjects the rank correlation calculations yielded values that exceed the demand for significance at the 99% level. - This is in accordance with the frequency response of the loudness analyzer.

5.1.2. Vowel duration

In Table I can be seen that vowel duration increases with increasing mouth opening, which corresponds well with earlier findings, cf. for instance House and Fairbanks (1953).

5.2. The consonants

In Table II are given the means of the duration, SPL, and loudness(GF) values for the (isolated) consonants.

TABLE II

Means of duration, SPL, and loudness(GF) for 4 initial and 4 intervocalic consonants, all surroundings pooled, 48 examples of each consonant. Two subjects, NT and CL.

	duration [ms]		SPL [$dB/2x10^{-5}N/m^2$]		loudness(GF) [sones(GF)]	
	NT	CL	NT	CL	NT	CL
s-	161	19o	59,1	57,4	7,8	7,8
f-	133	161	52,3	52,1	4,3	5,o
n-	122	115	73,1	71,2	8,2	6,8
m-	1o7	113	72,8	71,3	7,9	6,9
-s-	1o8	1o7	59,4	57,4	7,4	6,7
-f-	11o	111	51,6	52,8	4,7	4,3
-n-	79	66	77,5	7o,1	9,4	5,1
-m-	88	91	78,4	7o,2	1o,4	5,4

The SPL value of s is only 5-7dB higher than that of f. This difference is smaller than that found in many earlier investigations (15-20dB). This may be due to the operation of the free field response microphone, cf. section 3. - The loudness(GF) difference, however, is evident, since the noise of s lies in a frequency domain where the ear, and therefore the loudness analyzer, is very sensitive.

6. Final remarks

An evaluation of the loudness analyzer as a tool in speech research must await further experiments, which must obviously involve listening tests.

References

Fairbanks, G., A.S. House and E.L. Stevens 195o: "An Experimental Study of Vowel Intensities", JASA 22, 4, p. 457-459.

House, A.S. and G. Fairbanks 1953: "The Influence of Consonant Environment upon the secondary Acoustical Characteristics of Vowels", JASA 25, 1, p. 1o5-113.

Lehiste, I. and G.E. Peterson 1959: "Vowel Amplitude and Phonemic Stress in American English", JASA 31, 4, p. 428-435.

Thorvaldsen, P. 197o: "Construction of a Segmentator", ARIPUC 4/1969, p. 8-14.