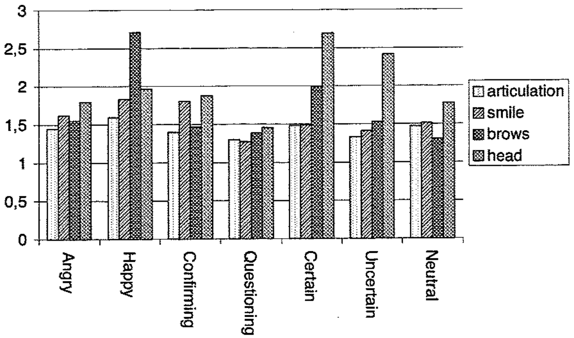

Figure 3. The effect of focus on the variation of several groups of MPG-4 FAP parameters, for different expressive modes

While much more detailed data on facial movement patterns is available in the database, we wanted to show the strong effects of focal accent on basically all facial movement patterns. Modelling the timing of the facial gestures and head movements relating to differences between focal and non-focal accent and to differences between expressive modes promises to be a fruitful area of future research.

Acknowledgements

This paper describes research in the CTT multimodal communication group including also Loredana Cerrato, Mikael Nordenberg, Magnus Nordstrand and Gunilla Svanfeldt which is gratefully acknowledged. Special thanks to Bertil Lyberg for making available the Qualisys Lab at Linköping University. The work was supported by the EU/IST projects SYNFACE, PF-Star and CHIL, and CTT, supported by VINNOVA, KTH and participating Swedish companies and organizations.

References

Beskow, J., L. Cerrato, B. Granström, D. House, M. Nordstrand & G. Svanfeldt, 2004. The Swedish PF-Star Multimodal Corpora. *Proc. LREC Workshop, Multimodal Corpora: Models of Human Behaviour for the Specification and Evaluation of Multimodal Input and Output Interfaces*, Lisbon, 34-37.

Beskow, J., O. Engwall & B. Granström, 2003. Resynthesis of Facial and Intraoral Articulation from Simultaneous Measurements. *Proc. ICPhS 2003*, Barcelona, 431-434.

Granström, B., D. House & J. Beskow, 2002. Speech and gestures for talking faces in conversational dialogue systems. In B. Granström, D. House & I. Karlsson (eds.), *Multimodality in language and speech systems*. Dordrecht: Kluwer Academic Publishers, 209-241.

Granström, B., D. House & M. Swerts, 2002. Multimodal feedback cues in human-machine interactions. *Proc. Speech Prosody 2002*, Aix-en-Provence, 347-350.

Nordstrand, M., G. Svanfeldt, B. Granström & D. House, 2004. Measurement of articulatory variation in expressive speech for a set of Swedish vowels. *Journal of Speech Communication 44*, 187-196.

Lund University, Centre for Languages & Literature, Dept. of Linguistics & Phonetics
Working Papers 52 (2006), 13–16

A Study of Simultaneous-masking and Pulsation-threshold Patterns of a Steady-state Synthetic Vowel: A Preliminary Report

Ulla Bjursäter
Department of Linguistics, Stockholm University
ullabj@ling.su.se

Abstract

This study will be a remake in part of Tyler & Lindblom "Preliminary study of simultaneous-masking and pulsation-threshold patterns of vowels" (1982), with the use of today's technology. A steady-state vowel as masker and pure tones as signals will be presented using simultaneous-masking (SM) and pulsation-threshold (PT) procedures in an adjustment method to collect the vowel masking pattern. Vowel intensity is changed in three steps of 15 dB. For SM, each 15 dB change is expected to result in about a 10-13-dB change in signal thresholds. For PT, the change in signal thresholds with vowel intensity is expected to be about 3-4 dB. These results would correspond with the results from the Tyler & Lindblom study. Depending on technology outcome, further experiments can be made, involving representations of dynamic stimuli like CV-transitions and diphthongs.

1 Introduction

This study is an attempt to partially replicate Tyler & Lindblom "Preliminary study of simultaneous-masking and pulsation-threshold patterns of vowels" (1982). Their intention was to investigate the effect of the two different masking types as well as the role of suppression in the coding of speech spectra.

Suppression, or lateral inhibition, refers to the reduction in the reaction to one stimulus by the introduction of a second (Oxenham & Plack, 1998). The ability of one tone to suppress the activity of another tone of adjacent frequency has been thoroughly documented in auditory physiology (Delgutte, 1990; Moore, 1978). In speech, suppression can be used to investigate formant frequencies.

In the original article, the authors (Tyler & Lindblom, 1982) constructed an experiment masking steady-state synthetic pure tones by simultaneous and pulsation-threshold patterns of vowels. Their vowels were synthesized on an OVE 1b speech synthesizer (Fant, 1960) with formant frequencies, bandwidths and intensities as approximate values for Swedish. In this study only one of the vowels from the original experiment is synthesized, using Madde, a singing synthesizer (<www.speech.kth.se/smptool/>) instead of OVE 1b.

In this experiment, the original vowel masking patterns will be used on the Swedish vowel /y/, a vowel that, according to Tyler & Lindblom (1982), is particularly useful in testing the role of suppression in speech as it has three closely spaced high frequency formants (F2, F3 and F4). F2 and F4 have about the same frequency as in the vowels /i/ and /e/, and a distinct perception of these three vowels must depend on good frequency resolution of F3 (Carlson et al., 1970; Bladon & Fant, 1978; Tyler & Lindblom, 1982).

Tyler, Curtis & Preece (1978) have shown that vowel masking patterns from a forward masking (FM) procedure preserves the formants better than patterns obtained using simultaneous masking (SM).

2 Method

2.1 Subject

All collected data will be from an experienced listener who will receive about 30 minutes of practice with the test procedure.

2.2 Procedure

The subject's hearing level will be controlled and set in connection with the experiment in the phonetics laboratory at the Department of Linguistics, SU, by Philips SBC HP890 headphones. The tone will be presented at different intensities to get a baseline point.

The tests are constructed in a graphic programming language, LabView (<www.ni.com/labview/>).

The Swedish vowel /y/ is synthesized in Madde, an additive, real-time, singing synthesizer (<www.speech.kth.se/smptool/>). The vowel formant frequencies and bandwidths used in this study (Table 1) are the same as used in Tyler & Lindblom (1982), and defined in Carlson, Fant & Granström (1975) and in Carlson, Granström & Fant (1970).

Table 1. Formant frequencies (F_1, F_2, F_3 and F_4), bandwidths and Q-values in Hertz for the vowel /y/.

/y/	*Frequency*	*Bandwidth*	*Q*
F_1	255 Hz	62.75 Hz	4.06
F_2	1930 Hz	146.5 Hz	13.17
F_3	2420 Hz	171.0 Hz	14.15
F_4	3300 Hz	215.0 Hz	13.35

The procedures of the two simultaneous-masking (SM) and pulsation-threshold (PT) tests are the same as in the Tyler & Lindblom (1982) study, although in this case with only one vowel instead of three.

In the SM procedure, the vowels will be presented 875 ms (between 50% points), repeated with a 125-ms silent interval. Three pulses of the pure tone will appear within each masking. These pulses start 125 ms after the vowel onset, continue for 125 ms and are separated by 125 ms. Rise/fall times (between 10% and 90% points) are 7 ms for both signal and masker.

In the PT procedure the masking vowel and the pulsating signal alternate, with duration of 125 ms each. In order to assist the subject in the task, every fourth signal (125 ms) is omitted. Rise/fall times are 7 ms and the signal and the vowel are separated by 0 ms.

F_0 levels vary between 80 Hz, 120 Hz and 240 Hz. The vowel intensity changes parametrically over a range of 45 dB in three steps of 15 dB. Intensity levels alternate between 55.5, 70.5 and 85.5 dB SPL, representing low-voiced, medium and strong speech. The presentation order of the vowel's fundamental frequency and intensity will be randomized.

The testing with PT-values and SM will alternate every 10th minute. Each condition will be presented until five estimates are registered. The stimuli will be presented monaurally, to the right ear. The subject will be instructed to adjust the level of the signal to a just noticeable pulsation threshold level; all answers are automatically registered in the LabView programme.

All data will be analyzed in SPSS 14.0 and MS Excel 2002.

3 Results

The technical solution of the test in LabView is currently under construction. The results of the tests are expected to correspond with the results from the Tyler & Lindblom (1982) study. Some variation may occur due to individual variations between subjects.

4 Discussion

The expectations of this study are to get data that concurs with the results from the Tyler & Lindblom (1982) study.

The results from the Tyler & Lindblom (1982) study show that the masking pattern received with the PT method delineates the vowel's formant frequencies better than the pattern received with the SM pattern. Suppression only occurs when two sounds are presented simultaneously, as in the SM procedure, which seems to result in the signal needing a higher intensity to be detected.

The difference between the SM and PT measurements were very small at low signal frequencies and quite large at high signal frequencies (Tyler & Lindblom, 1982). One of the explanations offered were that the high-intensity F_1 suppressed the activity caused by the higher formants, resulting in lower PT in the high-frequency regions.

Tyler & Lindblom (1982) also propose that the suggested suppression effects for steady-state vowels also could occur for all speech sounds, although in natural speech, the duration for which the vowel achieves its target is typically very short.

Depending on the outcome of the technology used in the PT and SM procedures, the program used in the test can be extended to further investigations of the effects of the two masking procedures on representations of dynamic stimuli, like CV-transitions and diphthongs.

Acknowledgements

The LabView test is being constructed with the invaluable help of Ellen Marklund, Francisco Lacerda and Peter Branderud, SU.

References

Bladon, A. & G. Fant, 1978. A two-formant model and the cardinal vowels. *Speech Transmission Laboratory Quarterly Progress Status Report (STL-QPSR 1)*, 1-8.

Carlsson, R., G. Fant & B. Granström, 1975. Two-formant Models, Pitch and Vowel Perception. In G. Fant & M.A.A. Tathham (eds.), *Auditory Analysis and Perception of Speech*. London: Academic.

Carlson, R., B. Granström & G. Fant, 1970. Some studies concerning perception of isolated vowels. *Speech Transmission Laboratory Quarterly Progress Status Report (STL-QPSR 2/3)*, 19-35.

Delgutte, B., 1990. Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers. *Journal of the Acoustical Society of America 87* (2), 791-809.

Fant, G., 1960. *Acoustical Theory of Speech Production*. The Hague: Mouton.

Moore, B.C.J., 1978. Psychophysical tuning curves measured in simultaneous and forward masking. *Journal of the Acoustical Society of America 63* (2), 524-532.

Oxenham, A.J. & C.J. Plack, 1998. Suppression and the upward spread of masking. *Journal of the Acoustical Society of America 104* (6), 3500-3510.

Tyler, R.S., J.F. Curtis & J.P. Preece, 1978. Formant preservation revealed by vowel masking patterns. *The Canadian Speech and Hearing Convention*, Saskatoon, Saskatchewan.

Tyler, R.S. & B. Lindblom, 1982. Preliminary Study of Simultaneous-Masking abs Pulsation Threshold Patterns of Vowels. *Journal of the Acoustical Society of America 71* (1), 220-224.

Lund University, Centre for Languages & Literature, Dept. of Linguistics & Phonetics 17
Working Papers 52 (2006), 17–20

Youth Language in Multilingual Göteborg

Petra Bodén[1] and Julia Groβe[2]
[1]Dept. of Linguistics and Phonetics, Centre for Languages and Literature, Lund University
petra.boden@ling.lu.se
[2]Institute of Swedish as a Second Language, Dept. of Swedish, Göteborg University
julia.grosse@svenska.gu.se

Abstract
In this paper, the results from a perception experiment about youth language in multilingual Göteborg are presented and discussed.

1 Introduction

1.1 Language and language use among young people in multilingual urban settings

The overall goal of the research project 'Language and language use among young people in multilingual urban settings' is to describe and analyze a Swedish variety (or set of varieties) hereafter called SMG (Lindberg, 2006). SMG stands for 'Swedish on Multilingual Ground' and refers to youth varieties like "Rinkeby Swedish" and "Rosengård Swedish". In the present paper, we address two of the project's research questions: SMG's relation to foreign accent and how SMG is perceived by the adolescents themselves.

1.2 Purpose of the perception experiment

In the perception experiment, Göteborg students are asked to listen for examples of "gårdstenska" (the SMG spoken in Göteborg) in recordings from secondary schools. The purpose is to identify speakers of SMG for future studies and to test the hypotheses that 1) monolingual speakers of Swedish can speak SMG and 2) speakers of SMG can code-switch to a more standardized form of Swedish. Foreign accent, defined here as the result of negative interference from the speaker's L1 (first language), cannot occur in the Swedish that is spoken by persons who have Swedish as their (only) L1, nor can foreign accent be switched off in certain situations.

2 Method

Stimuli were extracted from the research project's speech database and played once (over loudspeakers) to a total of 81 listeners. The listeners were asked to answer two questions about each stimulus: *Does the speaker speak what is generally called gårdstenska?* (*yes* or *no*), and *How confident are you about that?* (*confident, rather confident, rather uncertain* or *uncertain*). The listeners were also asked to answer a few questions about who they believed typically speaks gårdstenska. The 19 stimuli used in the experiment were approximately 30 second long sections that had been extracted from spontaneous (unscripted) recordings made at secondary schools in Göteborg. The listeners in the experiment were students from the same two schools as the speakers.

After having collected the answer sheets, a general discussion on SMG was held in each class.