

THE PERCEIVED NATURALNESS OF THE WORD-FINAL DENTAL STOP IN THE ENGLISH OF NATIVE SPEAKERS OF SWEDISH

K.P.H. Sullivan and Y.N. Karst

Phonetics Section
Department of Linguistics
Umeå University, Sweden.

ABSTRACT - The duration of consonants after long vowels is shorter than after short vowels in Swedish. This is a fact of which much is made in the instruction of Swedish as a foreign language. In English the difference can either be considered not to exist. In the teaching of English to native Swedish speakers little, if any, importance is placed on the 'removal' of this complementary vowel:consonant length relationship. This paper reports on work undertaken to assess the perceptual importance of acquiring English consonant length. Consonant length was found to be perceptually unimportant, yet the degree of aspiration after the consonantal closure was found to be perceptually noticeable.

INTRODUCTION

The duration of consonants after long vowels is shorter than after short vowels in Swedish (Eiert, 1964; Löfqvist, 1976). This is a fact of which much is made in the instruction of Swedish as a foreign language. The acquisition of this distinction is central to the development of a good Swedish accent (Thorén, 1989). In English the difference can either be considered not to exist (Kent and Moll, 1969) or to be negligible except in medial poststress position (Lisker, 1967;1969). The acoustics of the English stop is described in Zue (1976) and the acoustic differences between the English and the Swedish stops is clearly described in Woods (1975). In the teaching of English to native Swedish speakers little, if any, importance is placed on the 'removal' of this complementary vowel:consonant length relationship. This paper reports on work undertaken to assess the perceptual importance of acquiring English consonant length.

Two experiments were conducted. In the primary experiment the duration of the closure was altered to produce pairs of stimuli consisting of one word with the Swedish V:/t/ ratio and one word with the English V:/t/ ratio. In the auxiliary experiment the closure was removed totally for the stimuli presented to the listeners in order to assess the perceptual impact of the different degrees of aspiration.

PROCEDURE

The Primary Experiment: The Material

The carrier sentence "I saw a ____" was used to elicit 10 recordings of each of the conditions. The conditions were {p, t, k, s}{t, l, i} and result in the English words: pit, peat, tit, teat, kit, keat, sit, seat. They also result in a set of Swedish words: pit, pitt, tit, titt, kit, kiit and sit, sitt. The last pair of words are in fact not Swedish words when pronounced using /k/. The subject was asked to read these words beginning with a /k/. This he did without any problem. In Swedish the conditions were placed in the carrier sentence "Jag såg en ____", which has the same meaning as the English carrier sentence.

The recordings were made in the phonetics laboratory's anechoic chamber in the Department of Linguistics, Umeå University. A Panasonic SV-3700 DAT recorder was used with a Neumann U87 microphone and a wind shield. The order in which the conditions were placed in the carrier sentence was randomised.

The recordings were transferred to the department's network of Sun Ultras, where the duration of the vowel and the duration of the closure in the word-final /t/s were measured using the package ESPS/waves+™. The closure was measured from onset of the closure to the beginning of the aspiration. The measured duration are shown in Table 1 for the native speaker of English, Table 2 for the Swedish learner of English and Table 3 the Swedish data from the Swedish learner of English.

Word	vowel length (msec)	/t/ length (msec)	vowel:/t/ ration	μ (vowel length): μ (/t/ length) ratio
pit	$\mu=94$ $\sigma=13$	$\mu=120$ $\sigma=20$	$\mu=1:1.312$ $\sigma=0.350$	1:1.282
peat	$\mu=137$ $\sigma=13$	$\mu=100$ $\sigma=13$	$\mu=1:0.745$ $\sigma=0.160$	1:0.734
tit	$\mu=102$ $\sigma=10$	$\mu=112$ $\sigma=13$	$\mu=1:1.111$ $\sigma=0.205$	1:1.096
teat	$\mu=138$ $\sigma=15$	$\mu=108$ $\sigma=12$	$\mu=1:0.793$ $\sigma=0.142$	1:0.782
kit	$\mu=92$ $\sigma=10$	$\mu=102$ $\sigma=11$	$\mu=1:1.117$ $\sigma=0.167$	1:1.107
keat	$\mu=141$ $\sigma=11$	$\mu=104$ $\sigma=14$	$\mu=1:0.744$ $\sigma=0.121$	1:0.739
sit	$\mu=95$ $\sigma=10$	$\mu=106$ $\sigma=17$	$\mu=1:1.128$ $\sigma=0.264$	1:1.107
seat	$\mu=144$ $\sigma=13$	$\mu=87$ $\sigma=16$	$\mu=1:0.613$ $\sigma=0.138$	1:0.606

Table 1. Durations of vowel and /t/-consonant segments in the native English speaker's data

Word	vowel length (msec)	/t/ length (msec)	vowel:/t/ ration	μ (vowel length): μ (/t/ length) ratio
pit	$\mu=123$ $\sigma=15$	$\mu=187$ $\sigma=12$	$\mu=1:1.550$ $\sigma=0.248$	1:1.525
peat	$\mu=201$ $\sigma=19$	$\mu=159$ $\sigma=15$	$\mu=1:0.980$ $\sigma=0.106$	1:0.791
tit	$\mu=133$ $\sigma=16$	$\mu=192$ $\sigma=46$	$\mu=1:1.442$ $\sigma=0.359$	1:1.435
teat	$\mu=186$ $\sigma=13$	$\mu=184$ $\sigma=22$	$\mu=1:0.987$ $\sigma=0.98$	1:0.986
kit	$\mu=125$ $\sigma=9$	$\mu=208$ $\sigma=14$	$\mu=1:1.675$ $\sigma=0.184$	1:1.666
keat	$\mu=181$ $\sigma=14$	$\mu=174$ $\sigma=20$	$\mu=1:0.967$ $\sigma=0.135$	1:0.962
sit	$\mu=134$ $\sigma=12$	$\mu=204$ $\sigma=15$	$\mu=1:1.527$ $\sigma=0.166$	1:1.517
seat	$\mu=196$ $\sigma=17$	$\mu=182$ $\sigma=17$	$\mu=1:0.940$ $\sigma=0.134$	1:0.932

Table 2: Durations of vowel and /t/-consonant segments in the Swedish learner of English's English data

Word	vowel length (msec)	/t/ length (msec)	vowel:/t/ ration	μ (vowel length): μ (/t/ length) ratio
pit	$\mu=124$ $\sigma=13$	$\mu=223$ $\sigma=15$	$\mu=1:1.816$ $\sigma=0.235$	1:1.793
peat	$\mu=204$ $\sigma=13$	$\mu=181$ $\sigma=16$	$\mu=1:0.888$ $\sigma=0.1$	1:0.885
tit	$\mu=122$ $\sigma=6$	$\mu=230$ $\sigma=21$	$\mu=1:1.885$ $\sigma=0.187$	1:1.881
teat	$\mu=201$ $\sigma=16$	$\mu=184$ $\sigma=11$	$\mu=1:0.922$ $\sigma=0.97$	1:0.916
kit	$\mu=124$ $\sigma=14$	$\mu=227$ $\sigma=31$	$\mu=1:1.861$ $\sigma=0.369$	1:1.835
keat	$\mu=200$ $\sigma=19$	$\mu=186$ $\sigma=19$	$\mu=1:0.924$ $\sigma=0.118$	1:0.927
sit	$\mu=130$ $\sigma=11$	$\mu=243$ $\sigma=13$	$\mu=1:1.883$ $\sigma=0.177$	1:1.872
seat	$\mu=209$ $\sigma=10$	$\mu=216$ $\sigma=17$	$\mu=1:0.973$ $\sigma=0.068$	1:1.033

Table 3: Durations of vowel and /t/-consonant segments in the Swedish learner of English's Swedish data

For each of the eight words the one whose vowel duration was nearest the mean was chosen as the sample which would be manipulated for presentation to the listeners. The words were selected from the data collected using the English carrier sentence. The length of the consonant closure was calculated using ration of the mean of the vowel to the mean of the consonant: μ (vowel length): μ (/t/ length). A stimulus was created for each sample with the English V:/t/ ratio, based on the native English speaker's data, and another with the Swedish V:/t/ ratio, based on the data collected using the Swedish carrier sentence read by the Swedish learner of English. Table 4 shows the durations for the manipulated stimuli based on the native English speaker's data and Table 5 shows the durations for the manipulated stimuli based on the Swedish Learners of English's data. In the Tables 'Eng' means that the closure length has been manipulated to result in the English V:/t/ ratio; the 'Sv' means that the closure length has been manipulated to result in the Swedish V:/t/ ratio. The manipulation was carried out in ESPS/waves+™ by either clipping out a segment of the closure or by adding in a segment of silence in the middle of the existing closure. The vowels were not manipulated. This technique did not result in any sound quality degradation which sometimes occurs with manipulated lpc-resynthesis.

Word	vowel length (msec)	/t/ length (msec)
pit (Eng)	94	121
pit (Sv)	94	169
peat (Eng)	136	100
peat (Sv)	136	120
tit (Eng)	102	112
tit (Sv)	102	192
teat (Eng)	138	108
teat (Sv)	138	126
kit (Eng)	90	100
kit (Sv)	90	165
keat (Eng)	142	105
keat (Sv)	142	132
sit (Eng)	96	106
sit (Sv)	96	180
seat (Eng)	146	88
seat (Sv)	146	151

Table 4: The durations of the vowel and the /t/ closure for the manipulated stimuli based on the native English speaker's data

Word	vowel length msec)	/t/ length (msec)
pit (Eng)	128	164
pit (Sv)	128	230
peat (Eng)	199	146
peat (Sv)	199	176
tit (Eng)	133	146
tit (Sv)	133	250
teat (Eng)	187	146
teat (Sv)	187	171
kit (Eng)	125	138
kit (Sv)	125	229
keat (Eng)	178	132
keat (Sv)	178	165
sit (Eng)	128	142
sit (Sv)	128	240
seat (Eng)	197	119
seat (Sv)	197	203

Table 5: The durations of the vowel and the /t/ closure for the manipulated stimuli based on the Swedish learner of English's data

The Auxiliary Experiment: The material

The stimuli chosen for manipulation in the primary experiment were altered to produce the stimuli for this experiment. The closure of the word-final /t/ was totally removed. The same criteria for the amount of the stimulus to be removed was used as for the initial measuring of the closure's duration. The clipping out of the closure was done using *ESPSwaves+*TM.

Subjects: speakers

The speakers used to collect the material were a male 25 year old native speaker of standard Swedish, who has studied two semesters of English at the University Level and a male 31 year old native speaker of British English with a modified RP accent. The speakers were instructed to read the material at their usual speaking rate and to articulate clearly.

Subjects: listeners

There were 16 subjects aged 15 to 16 years. The listeners all attended North Chadderton High School, Chadderton, Oldham, England and were native speakers of English. North Chadderton High School is a state comprehensive school for boys and girls between the ages of 11 and 18 years. Neither the school nor the listeners received any payment for their participation in the experiment.

THE RECOGNITION TASKS

The Primary Experiment

The manipulated stimuli were presented in pairs with a 30msec pause between the elements of the pair. Each word had two pairs constructed: one with the Swedish V:/t/ ratio first and the other with the English V:/t/ ratio first. This resulted in 32 pairs (8 words x 2 speakers x 2 orders of presentation). The pairs were randomly presented with a 100msec pause between the pairs. Each pair was presented twice, resulting in 64 pairs. The listeners were asked to decide which of the pair was the most natural. They were asked to always make a decision and to follow their intuitions. They were presented with a response sheet on which they had to ring the numbers one or two, depending upon which of the two stimuli they believed to be the most natural.

The Auxiliary Experiment

The manipulated Swedish learner of English's stimuli were presented first in a randomised order. Only after the listeners had turned to the next response sheet were the English speakers manipulated stimuli presented. The listeners were asked to write down what they heard. They were informed that the stimuli had been manipulated, that the word may not sound like a word they knew. They were also asked not to worry about spelling a real word as there were no correct or incorrect answers. The words were presented with a 3 sec pause between them, and the cycle of eight words was repeated twice for each stimulus set. This experiment was conducted prior to Primary Experiment.

RESULTS

The Primary Experiment

The results are shown in Table 6.

Word	Most	Natural	Difference	$\chi^2(1)$
	Swedish V:/t/ ratio	English V:/t/ ratio		
pit	63	65	2	0.008
peat	59	69	10	0.632
tit	55	73	18	2.268
teat	61	67	6	0.195
kit	63	65	2	0.008
keat	64	64	0	0
sit	61	67	6	0.195
seat	61	67	6	0.195
TOTAL	487	537	50	2.345

Table 6: The results of the primary perception test showing the number of times a particular ratio was considered as more natural.

Although there are differences in selection of which ratio is most natural for the various words, ranging from 0 to 18, none of these differences reached the 5 per cent significance level of 3.841 for the χ^2 test with one degree of freedom. Yate's correction was applied in the calculation of these values. Naturally the total difference was also insignificant.

The Auxiliary Experiment

The subjects' responses are shown in Table 7. The word prior to the removal of the closure is listed in the first column, the responses to the word read by the Swedish learner of English are listed in the second column and the responses to the word read by the native English speaker in the third column.

Word	Swedish Learner of English	Native Speaker of English
pit	pleased (4), piss (3); piz (1), pees (2); peesed (1) blank (5)	pit (15); tit (1)
peat	pees (12); please (2); pith (1); pessed (1)	peat (16)
tit	tis (13); tiz (1); this (1); tit (1)	tit (15); tet (1)
teat	tease (14); tit (1); bland (1)	teat (15); tit (1)
kit	kiss (6); ciz (2); skis (2); give (1); keas(1); kethed (1); kith (1); skip (1); skiz (1).	kit (15); keat (1)
keat	keys (14); keyed (1); kis (1)	keat (16)
sit	siv (6); siz (3); says (1); siffs (1); sisf (1), sist (1); sifth (1); sit (1); sith (1)	sit (12); set (3); said (1)
seat	cease (7); seas (6); seaps (1); siz (1); sizthed (1)	seat (12); seed (1) tseet (1); teat (1); tit (1)

Table 7: The subjects' responses when the stop closure has been totally removed from the stimulus. The number in brackets after each response indicates the number of times that response was given

From these results it is clear that the existence of the stop closure is particularly important in the case of the Swedish learner of English. The degree of aspiration is greater than for a native English speaker, hence the subjects hear a fricative after the vowel. In the English case the perception, by a clear majority of listeners, of the manipulated stimulus as the same word as prior to manipulation implies that the length of the closure plays little discriminative importance.

DISCUSSION AND CONCLUSION

The results clearly show that the perceived naturalness of a pair of words which only differ in their vocal:/t/ closure ratio is identical. They subjects not only are not used to attending to such a difference, but do not rate such a difference as reducing the perceived naturalness of the stimulus. Of particular interest is that this was as true for words produced with a considerably different ratio (e.g. tit) as for words produced with an more similar ratio (e.g. keat). Considering the problems English learners of Swedish have in acquiring the Swedish complimentary vowel:consonant length ratio, this is perhaps not surprising. The correction of this complimentary length ratio in Swedish learners of English does not appear to be a worthwhile approach to produce a noticeable improvement in their English accent.

The removal of the closure in the stop totally produced a unexpected result. Namely that lack of effect on the perception of the native English speakers words. Although it is known that the degree of aspiration in English is less than in Swedish the perceptual effect when comparing the two sets of responses is extreme. In the manipulated stimuli based on the Swedish learner of English's data the responses are varied and the final stop's aspiration is almost always perceived as a fricative, in direct contrast to that of the native English speakers final stop's aspiration which is almost always still perceived as a stop consonant. Instruction to reduce the degree of aspiration could lead to an improvement in the English accents of native speakers of Swedish.

In order to assess whether this would indeed be the case a follow-up experiment is planned in which the aspiration of the word pairs in one case is the native Swedish aspiration and in the other the native English aspiration. The samples would be manipulated either by editing the aspiration from the productions of one speaker onto the production of the other, or by using re-synthesis techniques to alter the degree of aspiration. Once the manipulated stimuli have been created a similar discriminative perception experiment will be conducted.

Bannert (1994) in his book *På väg mot svenskt uttal* classified the stop consonant system of both Swedish and English as having two series of stop consonants in which the distinctive feature is aspiration rather than the tradition voice/voiceless feature. As the auxiliary experiment reported in this paper has shown that the degree of aspiration is possibly the most important perceptual aspect of the stop consonant in English, perhaps more important than the closure itself, a study investigating the other stop consonant series in the same manner as the experiment reported here, and with the manipulated aspiration as suggested for the words used in the reported experiments, would seem judicious.

ACKNOWLEDGEMENTS

The authors are grateful to the staff and students of the Department of Modern Languages, Chadderton High School, Chadderton, Oldham, England for the helpful and cheerful manner in which they participated in these experiments.

The expenses associated with this research were funded by the Faculty of Arts, Umeå University, Umeå, Sweden.

REFERENCES

- Bannert, R. (1994). *På väg mot svenskt uttal*, (Studentlitteratur: Lund).
- Elert, C.-C. (1964). *Phonological studies of quantity in Swedish*, (Almqvist & Wiksell: Uppsala).
- Kent, R. and Moll, K. (1969). "Vocal-tract characteristics of the stop cognates", *Journal of the Acoustical Society of America*. 46, 1549–1555.
- Lisker, L. (1967). "Measuring stop closure duration from intraoral pressure record", Haskins Labs. Status Report on Speech Research. 7/8, 5.1–5.6.
- Lisker, L. (1969). "Stop duration and voicing in English", Haskins Labs. Status Report on Speech Research. 19/20, 215–230.
- Löfqvist, A. (1976). "Closure duration and aspiration for Swedish stops", Working Papers, Phonetics Laboratory, Department of General Linguistics, Lund University. 13:1–39.
- Thorén, B. (1989). "Ett basprosodiskt grepp om uttalsundervisningen", Andra symposiet om svenska som andraspråk i Göteborg. 288–296.
- Wood, S. (1975). "What is the difference between English and Swedish Dental Stops?" Working Papers, Phonetics Laboratory, Department of General Linguistics, Lund University. 10, 173–193.
- Zue, V.W. (1976) Acoustic characteristics of stop consonants: a controlled study. Reproduced 1980 by the Indiana University Linguistics Club, Bloomington Indiana.