Phonotactically determined allophones of the j phoneme in Swedish

Björsten, S.

journal: TMH-QPSR
volume: 37
number: 2
year: 1996
pages: 005-008

http://www.speech.kth.se/qpsr
Phonotactically determined allophones of the j phoneme in Swedish

Sven Björsten
Department of Linguistics, Stockholm University

Abstract
The realization of the j phoneme in Swedish varies between a glide and a fricative. This study describes the phonotactical distribution of these variants in 3 subjects from the Stockholm area. It was found that syllable initial single j-, and j in initial clusters of the type obstruent + j, vary between a glide and a fricative, while j in initial clusters of the type nasal + j, syllable final single -j, and j in all final clusters studied, are realized as a glide. There could be either one or two underlying targets for j/ giving rise to this distribution.

Introduction
According to Lindblad (1980) the phonetic realization of the j phoneme in Swedish varies between a fricative and a glide. The realization as a glide is the more common. There is, however, no study which tries to map the distribution of the different allophones. The aim of the present study was to investigate the distribution of fricative and glide variants of j/, and if they are phonotactically determined in careful, “laboratory” speech. Also considered were the questions: Does the distribution of allophones obey the sonority hierarchy, or is there some other explanation for the distribution?

Phonotactics of j/ in Swedish.
/j/ occurs singly in both syllable initial and syllable final position in Swedish, and also occurs in a fairly large number of initial and final clusters. For a listing of these, see Elert (1970). The two component clusters can be classified into four groups: Initial clusters where j/ is preceded by an obstruent (eg. bj-, fj-), initial clusters where j/ is preceded by a nasal (mj-, nj-), final clusters were j/ is followed by an obstruent (eg. -jd, -jg) and final clusters were j/ is preceded by a liquid or nasal (eg. -nj, -nj). Examples from all these groups, and also singly occurring initial and final /j/, were considered in this study.

Method
Subjects and word material
A list of Swedish words containing /j/ in most attested phonotactical combinations was prepared. The list contained 33 words, 3 occurrences of each, in random order. The list was read by 3 subjects, 1 male and 2 females, all of them native Swedish speakers, approximately 20 years old and raised in the Stockholm area. The subjects were not aware of the purpose of the study. The list was read in a careful but not over-articulated way in an anechoic room and the words were recorded with a digital tape recorder, and then transferred to computer files (sampling frequency 16 kHz). The word material was then analysed using three different methods: manual analysis of spectrograms, LPC analysis, and a perception test.

Spectrograms
The entire word material was printed as spectrograms with a frequency range of 0-8 kHz. Most of the spectrograms were then independently examined by 2 persons, the author and an experienced phonetician. The amount and degree of blackness during the j segment in the frequency range 5-8 kHz was evaluated. This range was chosen because it was shown in Lindblad (1980) that this area of the spectrum differs between fricative j/ and glide j/. The spectrograms were divided into 5 classes (1-5), class 1 containing the least extensive blackness
Figure 1. Results of analysis of spectrograms. The histograms show the percentages of j-sounds (in 8 different types of clusters) falling into specific classes when evaluating the spectrograms. There are 5 classes, from class 1 with the least amount of noise (glide) to class 5 with the greatest amount of noise (fricative). Nj = initial nasal + j, -Sj = final sonorant + j.
Figure 2. Result of LPC analysis. On the Y axis is the K value (see explanation in text), which is higher the more noise the j segment contains. On the X axis is type of cluster. Nj- = initial nasal + j. The central box comprises 50% of the data, half above and half below the median which is marked by the center horizontal line. The H spread is the difference between the hinges. The inner fences are defined as the lower hinge - (1.5 * H spread) and the upper hinge + (1.5 * H spread). Likewise, the outer fences are the lower hinge - (3 * H spread) and the upper hinge + (3 * H spread), respectively. Outside data values (marked by asterisks) are any data values beyond either inner fence; far outside data values (marked by empty circles) are any data values beyond either outer fence. The outermost data value on each end that is still not beyond the corresponding inner fence is called an adjacent value. The whiskers are drawn from each hinge to the corresponding adjacent value.

in the relevant area and class 5 containing the most extensive blackness. Classes 2-4 were intermediary. Low classes were considered containing j with a glide articulation and high classes were considered containing j with a fricative articulation, but a boundary between the articulations was not established because there was a continuum between the two extremes. The classifications obtained by the two evaluators were similar (Spearman Rank-Order Correlation Coefficient rho = 0.662). The results of one of the evaluators are presented in Fig. 1.

Figure 3. Result of the perception test. On the Y axis is the class membership as determined by the subjects in the perception test. The values on the scale are the means of the 14 subjects. The higher the Y value, the more fricative the j segment. On the X axis type of cluster. Nj- = initial nasal + j, -Sj = final sonorant + j. For an explanation of the box diagram, see Fig. 2.

LPC-analys

The reason for using this method was the possibility of getting the noise level of the j segment expressed as a numerical value so that the continuous nature of the variation between the different varieties of j could be shown. LPC is a method of mathematically predicting the continuation of the signal. Two parameters, residual signal energy and signal energy, were calculated for overlapping frames (length: 20 ms overlap: 5 ms) during the entire length of the j segment. Signal energy described the total energy of the signal prior to preemphasis, residual signal energy described the amount of irregular energy (calculated after preemphasis) that could not be predicted. Such energy was here considered as noise. A factor K was then calculated, defined as:

\[ K = \text{the mean of the logarithms of the ratios between} \text{ residual signal energy and signal energy for each frame of the} \text{ j segment} \]

K was taken to express the signal level (in bel) referring to the amount of noise present in the segment. A fricative j should have a higher K value than a glide j. The K values were found to correspond well with the amount of blackness of the spectrograms described above. The results are
shown in Fig. 2. Only initial clusters were studied with this method.

Perception test
The j-segments were cut out from all words whose spectrograms had been analysed, and a tape was prepared on which the segments appeared in random order. This tape was used for a perception test in which 14 phonetically trained persons participated. The subjects were asked to judge the auditory characters of the j segments and to assign them to 5 classes (1-5), where class 1 represented the most glide-like pronunciation, and class 5 the most fricative-like pronunciation, classes 2-4 being intermediary. The results are shown in Fig. 3.

Results
The results are most clearly seen in the spectrogram analysis, Fig. 1. The clusters can be divided into two groups: In the first group both high and low classes can be found which means that the realization of /j/ varies between a glide and a fricative pronunciation. To this group belong single initial j- and the initial clusters fj-, bj-, and pj-. In the second group virtually only low classes occur, which means that in these clusters j is pronounced as a glide. To this group belong the initial cluster type nasal + j and all syllable final occurrences of j , that is single final -j, -jd, and sonorant + j. These results are confirmed by LPC-analysis (Fig. 2) and the perception test (Fig. 3). Fig. 2 shows that initial nasal + j clusters have the lowest noise levels (only initial clusters studied). Fig. 3 shows that the four clusters that have the highest median values in the perception test belong to the group which show both glide and fricative pronunciation in the spectrogram analysis, while the four clusters that have the lowest medians belong to the group which shows only glide pronunciation. The results of all 3 analysis methods also hold for all 3 subjects individually.

Discussion
The sonority hierarchy, originally formulated by Jespersen (1926), is presented in a simple form by Katamba (1989). He uses 6 groups (from the least sonorous to the most sonorous): 1. Voiceless obstruents 2. Voiced obstruents 3. Nasals 4. Liquids 5. Glides 6. Vowels. This hierarchy can be used to describe the optimal syllabic structure, highly sonorous elements (usually vowels) forming the syllabic nucleus and the peripheral elements arranged according to sonority, with decreasing sonority as the distance to the nucleus increases. This means that, for instance, initial clusters of the type obstruent + liquid would be optimal and common in the languages of the world, whereas initial clusters of the opposite type (liquid + obstruent) would not. As it is shown in this study that /j/ in final nasal + j and liquid + j clusters is realized as a glide, these clusters in Swedish do not conform to the sonority hierarchy. The other cluster types containing /j/ in Swedish do conform.

The question arises: are the glide and fricative allophones of /j/ manifestations of one or two underlying targets? The data here presented are compatible with either view. If a single glide target is assumed, the notion of strengthening (Kiparsky, 1988) could be evoked to account for the prevalence of fricative pronunciation of initial single j-. Strengthening processes are typical of syllable initial position and are intended to make perception easier. The prevalence of fricative j in initial clusters of the type obstruent + j could be explained through the effect of coarticulation, the noise of the initial obstruent spreading to the adjacent j segment. Alternatively, phonotactics could determine the distribution of two targets. These alternative hypotheses could be tested in future research.

References