# Visual discrimination between Swedish and Finnish among L2-learners of Swedish

*Niklas Öhrström, Frida Bulukin Wilén, Anna Eklöf and Joakim Gustafsson*
*Department of Linguistics, Stockholm University.*

## Abstract

*A series of speech reading experiments were carried out to examine the ability to discriminate between Swedish and Finnish among L2 learners of Swedish and Spanish as their mother tongue. This group was compared with native speakers of Swedish and a group with no knowledge in Swedish or Finnish. The results showed tendencies, that familiarity with Swedish increased the discrimination ability between Swedish and Finnish.*

## Introduction

Audition is the main modality for speech decoding. Nevertheless, visual information about the speech gestures while listening provides complementary visual cues to speech perception. This use of visual information plays a significant role, especially during noisy conditions (Sumby and Pollack, 1954; Erber, 1969). However, McGurk and MacDonald (1976) showed that the visual signal is incorporated in auditory speech percept, even at favorable S/N ratios. They used dubbed tapes with a face pronouncing the syllables [gaga] and [baba]. When listeners saw the face articulating [gaga], while the audio track was changed for [baba], the majority reported having heard [dada]. Later Traunmüller and Öhrström (2007) demonstrated that this phenomenon also holds for vowels, where the auditory percept is influenced by strong visual cues such as lip rounding. These findings are clear evidence that speech perception in face-to-face-communication is a bimodal rather than uni-modal process.

In case of no available acoustic speech signal, the listener must fully rely on visual speech cues, i.e. speech reading. Visual information alone is in most cases not sufficient for speech processing, since many speech sounds fall into the same visually discriminable category. Homorganic speech sounds are difficult to distinguish, while labial features, such as degree of lip rounding or lip closure are easily distinguishable (Amcoff, 1970). It has been shown that performance in speech reading varies greatly across perceivers (Kricos, 1996). Generally, females perform better than males (Johnson et al., 1988).

The importance of the visual signal in speech perception has recently been stressed in two articles. Soto-Faraco et al. (2007) carried out a study where subjects were presented silent clips of a bilingual speaker uttering sentences in either Spanish or Catalan. Following the first clip, another one was presented. The subjects' task was to decide whether the language had been switched or not from the one clip to the other. Their subjects were either people with Spanish or Catalan as their first language. The other group consisted of people from Italy and England with no knowledge in Spanish or Catalan. In the first group bilinguals performed best. However, people with either Spanish or Catalan as their mother tongue performed better than chance level. The second group did not perform better than chance level, ruling out the possibility that the performance of the first group was due to paralinguistic or extralinguistic signals. Their performance was based on linguistic knowledge in one of the presented languages. Later, Weikum et al. (2007) carried out a similar study, where the speaker was switching between English and French. In this study, the subjects were 4-, 6- and 8-month-old infants, acquiring English. According to their results, the 4-, and 6-month-olds performed well, while the 8-month-olds performed worse. One interpretation is that the 4- and 6-month-olds discriminate on the basis of psycho-optic differences, while the 8-month-olds are about to lose this ability as a result of acquiring the visually discriminable categories of English. These are important findings, since it highlights the visual categories as part of the linguistic competence. The two studies are not fully comparable, since they deal with different languages, which do not necessarily differ in the same way. It cannot be excluded that French and English are so dissimilar visually, that it might be possible to discriminate between them on the basis of psycho-optic differences only. It does however suggest that we might *relearn* to discriminate the L1 from an unknown language.

This study deals with L2 learning. Can we learn an L2 to the extent, that it could be visually discriminable from an unknown language? Can we establish new visual language-specific categories as adults? To answer these questions the visual discriminability between Finnish and Swedish among (i) people with Swedish as their L1, (ii) people with Swedish as their L2 (immigrants from Latin-America) and (iii) people with no knowledge in either Swedish or Finnish (Spanish citizens). The Swedish and Finnish speech sound inventories differ in many ways: The two languages both use front and back rounded vowels. Finnish lacks the difference between in- and out-rounded vowels (Volotinen, 2008). This feature is easy to perceive visually, since in-rounding involves hidden teeth (behind the lips), while out-rounding does not. The Finnish vowel system recruits only three degrees of openness, while Swedish makes use of four degrees. The temporal aspects may be difficult to perceive visually (Wada et al. 2003). Swedish is often referred to as being a stress timed language (Engstrand, 2004): The distances between stressed syllables are kept more or less constant. In addition, Swedish make use of long and short vowels in stressed syllables. The following consonant length is complementary striving to keep stressed syllables at constant length. Finnish is often referred as a quantity language and vowels as well as consonants can be long or short regardless of stressing. Unlike Swedish, a long vowel can be followed by a long consonant. Swedish is abundant in quite complex consonant clusters. Finnish is more restrained from that point of view.

# Method

## Speech material

The study is almost a replica of that of Soto-Faraco et al. (2007). One bilingual, Finno-Swedish, male was chosen as speaker. His Swedish pronunciation was judged to be on L1-level, by the authors. His Finnish pronunciation was judged (by two finnish students) to be almost on L1-level. The speaker was videotaped while pronouncing the four swedish sentences: (i) *Fisken är huvudföda för trädlevande djur*, (ii) *Vid denna sorts trottoarer brukar det vara pölfritt,* (iii) *Denna är ej upplåten för motorfordon utan är normalt enbart avsedd för rullstolar*, (iv) *En motorväg är en sån väg som består av två körbanor med i normalfallet två körfält,* and four finnish sentences: (i) *Teiden luokitteluperusteet vaihtelevat maittain,* (ii) *Jänisten sukupuolet ovat samannäköisiä,* (iii) *Yleensä tiet ovat myös numeroitu, ja usein tieluokka voidaan päätellä numerosta,* (iv) *Kalojen tukirangan huomattavin osa muodostuu selkärangasta ja kallosta*.

## Subjects

Three groups were examined. Group 1 consisted of 22 (12 female and 10 male) L2-speakers of Swedish, aged 23-63 years (Mean= 37.5 years). They were all Spanish speaking immigrants from Latin-America. Group 2 consisted of 12 (6 male and 6 female) L1 speakers of Swedish, aged 18-53 years (Mean=38.8 years). Group 3 consisted of 10 (4 female and 6 male) L1 speakers of Spanish, aged 24-47 years (Mean=37.7 years). They were all residents of San Sebastian (Spain), with Spanish as their L1 and no knowledge in Swedish or Finnish.

## Procedure

Each group was presented 16 sentence pairs in quasi-randomized order. The subjects' task was to judge whether or not the following sentence was in the same language as the first. In group 1, information about their education in Swedish was asked for (i.e. number of semesters at the School of SFI, Swedish for immigrants) and the age when arriving in Sweden. In group 1, the subjects were asked to estimate their use of Swedish as compared with their use of Spanish on a four-degree-scale.

# Results

## Group 1 (L2 speaker of Swedish)

Group 1 achieved a result of, on average, 10.59 correct answer of 16 possible (sd = 2.17). A one-sample t-test revealed that their performance was significantly over chance level ($p<0.001$, $t_{21}=3.819$). Performance was positively correlated with use of Swedish as compared to Spanish (Pearson $\rho = 0.353$, $p = 0.107$, $N=22$). This correlation did not reach significance level. Performance was negatively correlated with age when arriving to Sweden ($\rho = -0.328$, $p = 0.136$). Performance was weakly negatively correlated with years spent in Sweden ($\rho = -0.217$, $p = 0.333$). However this correlation failed to attain significance.

### Group 2 (L1 speakers of Swedish)

Group 2 achieved a result of, on average, 11.25 correct answer of 16 possible (sd = 1.35). A one-sample t-test revealed that their performance was significantly over chance level ($p<0.001$, $t_{11}=4.437$).

### Group 3 (No knowledge in Swedish)

Group 3 achieved a result of, on average, 9.50 correct answer of 16 possible (sd = 1.96). A one-sample t-test revealed that their performance was significantly over chance level ($p<0.05$, $t_9=2.262$).

### Summary Results

An ANOVA was performed to reveal any significant differences between the groups. None of the differences reached significance, but the difference between group 2 and 3 almost reached significance ($p = 0.101$).

## Discussion

There was a negative correlation between subjects' age of arrival and performance. Studies has shown that learning L2 pronunciation is more difficult among elderly as compared with younger learners (Ellis, 1994). It is likely that age is a factor that impedes visual perception of an L2 as well.

Subjects in group 3 performed better than chance level. These results cannot be full explained, and they are not in line with the results obtained by Soto-Faraco et al. (2007), whose subjects, with no knowledge in the target languages, didn't reach a result above chance level. There are several possible explanations: The visual differences between Swedish and Finnish (produced by our speaker) could be so large, that the difference might be perceived on basis of psycho-optic signals. Another possible explanation has to do with extra- or paralinguistic visual signals that have become language dependent (for our speaker).

Group 2 achieved a higher score than group3. The difference almost reached significance, and showed that knowledge in at least one of the target languages favors visual discriminability. These tendencies are in line with Soto-Faraco et al. (2007).

In group 1, there were tendencies that the estimated amount of use of Swedish was well correlated with the performance. This factor was stronger than the number of semesters spent on learning Swedish at SFI. This is important, since the learning and establishment of new visual categories doesn't stop after the course taken. If we acknowledge the visually discriminable categories as part of the linguistic competence, it must be a goal for L2 learners to master these categories.

## Acknowledgements

We would like to thank Pilvi Mattila and Liisa Nousiainen for their evaluation of the speaker's swedish accent. Thank you Hernan Quintero, Lena Höglund Santamarta and Xila Philström for practical issues, concerning the experimental procedure in a foreign language.

## References

Amcoff S (1970) Visuell perception av talljud och avläsestöd för hörselskadade. Rapport Nr. 7, LSH Uppsala: Pedagogiska institutionen.

Ellis R. (1994) The study of second language acquisition. Oxford: University press.

Engstrand O. (2004) Fonetikens grunder. Lund: Studentlitteratur.

Erber N. (1969) Interaction of audition and vision in the recognition of oral speech stimuli. Journal of Speech and hearing Research 12, 423-425.

Johnson F.M. et al. (1988) Sex differences in lipreading. Bulletin of Psychonomic Society 26, 106-108.

Kricos P. (1996) Differences in visual intelligibility across talkers. In Stork D and Hennecke M. (eds.) Speech reading by humans and machines. Heidelberg, Germany: Springer-Verlag, 43-55.

McGurk H. & MacDonald J. (1976) Hearing lips and seeing voices. Nature 264, 746-748.

Soto-Faraco S. et al. (2007) Discriminating languages by speech-reading. Perception and Psychophysics 69, 218-231.

Sumby W.H. & Pollack I. (1954) Visual contribution to speech intelligibility in noise. Journal of Acoustical Society of America 26, 212-215.

Traunmüller H. & Öhrström N. (2007) Audiovisual perception of openness and lip rounding in front vowels. Journal of Phonetic. 35, 244-258.

Volotinen H. (2008) För öppet och tonlöst. Uttalssvårigheter i svenskan på universitetsnivån. Pro Gradu-avhandling i svenska språket, Institutionen för språk. Jyväskylä university.

Wada Y. et al. (2003) Audiovisual integration in temporal perception. international Psychophysiology 50, 117-124.
Weikum W. et al. (2007) Visual language discrimination in infancy. Science. 316. 1159.
McGurk H. MacDonald J. (1976) Hearing lips and seeing voices. Nature 264, 746-748.