

As the project in addition aims to characterize also other aspects of public speaking, a variety of representative speech samples will be collected. In analyses of this material, fluency, pausing, prominence, emphasis and voice characteristics will be central. Among the questions we seek answers to are: What types of strategies are used for holding the floor? How does speech perceived as fluent and disfluent respectively differ acoustically? How are prominence and emphasis used in speech in media? What are the prosodic characteristics of agitation? Answers to these questions, we believe, will add to our understanding of human communicative capability and will also be useful in modeling speaking style variation. Knowledge gained within the project may further be expected to be practically applicable.

Acknowledgements

This work was supported by The Swedish Research Council (VR).

References

- Braga, D. & M.A. Marques, 2004. The pragmatics of prosodic features in the political debate. *Proc. Speech Prosody 2004*, Nara, 321-324.
- Bruce, G. & P. Touati, 1992. On the analysis of prosody in spontaneous speech with exemplification from Swedish and French. *Speech Communication 11*, 453-458.
- Carlson, R., B. Granström, M. Heldner, D. House, B. Megyesi, E. Strangert & M. Swerts, 2002. Boundaries and groupings – the structuring of speech in different communicative situations: a description of the GROG project. *TMH-QPSR 44*, 65-68.
- Gussenhoven, C., 2002. Intonation and interpretation: Phonetics and phonology. *Proc. Speech Prosody 2002*, Aix-en-Provence, 11-13.
- Heldner, M. & B. Megyesi, 2003. Exploring the prosody-syntax interface in conversations. *Proc. ICPHS 2003*, Barcelona, 2501-2504.
- Hirschberg, J., 2002. The pragmatics of intonational meaning. *Proc. Speech Prosody 2002*, Aix-en-Provence, 65-68.
- Horne, M., E. Strangert & M. Heldner, 1995. Prosodic boundary strength in Swedish: final lengthening and silent interval duration. *Proc. ICPHS 1995*, Stockholm, 170-173.
- Johannesson, K., 1998/1990. *Retorik eller konsten att övertyga*. Stockholm: Norstedts Förlag.
- Liscombe, J., J. Venditti & J. Hirschberg, 2003. Classifying subject ratings of emotional speech using acoustic features. *Proc. Eurospeech 2003*, Geneva, 725-728.
- Mozziconacci, S., 2002. Prosody and emotions. *Proc. Speech Prosody 2002*, Aix-en-Provence, 1-9.
- Strangert, E., 1991. Phonetic characteristics of professional news reading. *PERILUS XII*. Institute of Linguistics, University of Stockholm, 39-42.
- Strangert, E., 1993. Speaking style and pausing. *PHONUM 2*. Reports from the Department of Phonetics, University of Umeå, 121-137.
- Strangert, E., 2004. Speech chunks in conversation: Syntactic and prosodic aspects. *Proc. Speech Prosody 2004*, Nara, 305-308.
- Strangert, E., 2005. Prosody in public speech: analyses of a news announcement and a political interview. *Proc. Interspeech 2005*, Lisboa, 3401-3404.
- Strangert, E., & R. Carlson. 2006. On modeling and synthesis of conversational speech. *Proc. Nordic Prosody IX, 2004*, Lund, 255-264.
- Touati, P., 1991. Temporal profiles and tonal configurations in French political speech. *Working Papers 38*. Department of Linguistics, Lund University, 205-219.
- Wichmann, A., 2002. Attitudinal intonation and the inferential process. *Proc. Speech Prosody 2002*, Aix-en-Provence, 11-22.

Effects of Age on VOT: Categorical Perception of Swedish Stops by Near-native L2 Speakers

Katrin Stölten

Centre for Research on Bilingualism, Stockholm University

Katrin.Stoelten@biling.su.se

Abstract

This study is concerned with effects of age of onset of L2 acquisition on categorical perception of the voicing contrast in Swedish word initial stops. 41 L1 Spanish early and late learners of L2 Swedish, who had carefully been screened for their 'nativelike' L2-proficiency, as well as 15 native speakers of Swedish participated in the study. Three voicing continua were created on the basis of naturally generated word pairs with /p t k b d g/ in initial position. Identification tests revealed an overall age effect on category boundary placement in the nativelike L2 speakers' perception of the three voicing continua. Only a small minority of the late L2 learners perceived the voicing contrast in a way comparable to native-speaker categorization. Findings concerning the early learners suggest that most, but far from all, early L2 speakers show nativelike behavior when their perception of the L2 is analyzed in detail.

1 Introduction

From extensive research on infant perception it has become a well-known fact that children during their first year of life tune in on the first language (L1) phonetic categories, leaving them insensitive to contrasts not existing in their native language (e.g. Werker & Tees, 1984).

In a study by Ruben (1997) it was found that children who had suffered from otitis media during their first year of life showed significantly less capacity for phonetic discrimination compared to children with normal hearing during infancy when they were tested at the age of nine years. Such findings do not only demonstrate the importance of early linguistic exposure, they have also been interpreted as an indication for the existence of a critical period for phonetic/phonological acquisition which may be over at the age of one year (Ruben, 1997).

In research of age effects on language acquisition one classical issue is concerned whether theories of a critical period can be applied to second language (L2) acquisition. The question is whether the capacity to acquire phonetic detail in L2 learning is weakened or lost due to lack of verbal input during a limited time frame for phonetic sensitivity, or whether a nativelike perception and an accent-free pronunciation is possible for any adult L2 learner.

The present study is part of an extensive project on early and late L2 learners of Swedish with Spanish as their L1. The subjects have been selected on the criterion that they are perceived by native listeners as mother-tongue speakers of Swedish in everyday oral communication. Thereafter, the candidates' nativelike L2 proficiency has been tested for various linguistic skills. The present study focuses on the analysis of the nativelike subjects' categorical perception of the voicing contrast in Swedish word initial stops.

Both Swedish and Spanish recognize a phonological distinction between voiced and voiceless stops in terms of voice onset time (VOT) but they differ as to where on the VOT continuum the stop categories separate. In contrary to languages like Swedish and English,

which treat short-lag stops as voiced and long-lag stops as voiceless, in Spanish short-lag stops are recognized as voiceless, while stops with voicing lead are categorized as voiced (e.g. Zampini & Green, 2001). Consequently, Spanish phoneme boundaries are perceptually located at lower VOT values than in, for example, English (Abramson & Lisker, 1973).

Since language-specific category boundaries are established at a very early stage in language development a great amount of perceptual sensibility is needed by a second language learner in order to detect the categories present in the target language. The fact that L2 learners generally show difficulties in correctly perceiving and producing these language-specific categories (e.g. Flege & Eefting, 1987) suggests that categorical perception may be considered a good device for the analysis of nativelike subjects' L2 proficiency.

For the present study the following research questions have been formulated:

- (1) Is there a general age effect on categorical perception among apparently nativelike L2 speakers of Swedish?
- (2) Are there late L2 learners who show category boundaries within the range of native-speaker categorization?
- (3) Do all (or most) early L2 learners show category boundaries within the range of native-speaker categorization?

2 Method

2.1 Subjects

A total of 41 native speakers of Spanish (age 21-52 years), who had previously been screened for their 'nativelike' proficiency of Swedish in three screening experiments (see Abrahamsson & Hyltenstam, 2006), were chosen as subjects for the study. The participants' age of onset (AO) of L2 acquisition varied between 1 and 19 years and their mean length of residence in Sweden was 24 years. The subjects had an educational level of no less than senior high school and they all had acquired the variety of Swedish spoken in the great Stockholm area.

The control subjects consisting of 15 native speakers of Swedish were carefully matched with the experimental group regarding present age, sex, educational level and variety of Swedish. All participants went through a hearing test (OSCILLA SM910 screening audiometer) in order to ensure that none of the subjects suffered from any hearing impairment.

2.2 Stimuli

The speech stimuli were prepared on the basis of naturally generated productions of three Swedish minimal word pairs: *par* /pa:r/ (pair, couple) vs. *bar* /ba:r/ 'bar, bare, carried', *tal* /ta:l/ 'number, speech' vs. *dal* /da:l/ 'valley', *kal* /ka:l/ 'naked, bald' vs. *gal* /ga:l/ 'crow(s), call(s)'. A female speaker of Stockholm Swedish with knowledge of phonetics was recorded in an anechoic chamber while reading aloud the words in isolation. The speaker was instructed to articulate the voiceless stops with an extended aspiration interval and the voiced counterparts with a clear period of voicing prior to stop release. All readings were digitized at 22 kHz with a 16-bit resolution.

For all stop productions VOT was determined by measuring the time interval between the release burst and the onset of voicing in the following vowel. Thereafter, the release bursts of the voiceless stops were equalized to 5ms. The aspiration phase was then extended to +100ms VOT by generating multiple copies of centre proportions of the voicing lag interval. The stimuli for the perception test were created by shortening the aspiration interval in 5ms-steps. Voicing lead was simulated by copying the prevoicing interval from the original production of the corresponding voiced stop and placing it prior to the burst. The prevoicing maximum was first put at -100ms and then varied in 5ms-steps. Finally, a set of 30 speech stimuli ranging from +90ms to -60ms VOT for each stop continuum was considered appropriate for the study.

2.3 Testing procedure

A forced-choice identification task designed and run in E-Prime v1.0 (Schneider, Eschman & Zuccolotto, 2002) was performed by each subject individually in a sound treated room. The three voicing continua were tested separately. The speech stimuli were preceded by the carrier phrase *Nu hör du* 'Now you will here' and randomly presented through headphones (KOSS KTX/PRO) one at a time. For each stimulus the listeners were told to decide whether they heard the word containing a voiced or voiceless stop and confirm their answer by pressing a corresponding button on the keyboard. The experimenter was a male native speaker of Stockholm Swedish.

3 Results

Stop category boundaries (in ms VOT) were calculated for each subject and plotted against their age of onset. Since category boundary locations vary with place of articulation (see, e.g. Abramson & Lisker, 1973) the stop pairs were analyzed separately. Due to extreme VOT values the results from one subject (AO 5) had to be discarded from further analysis.

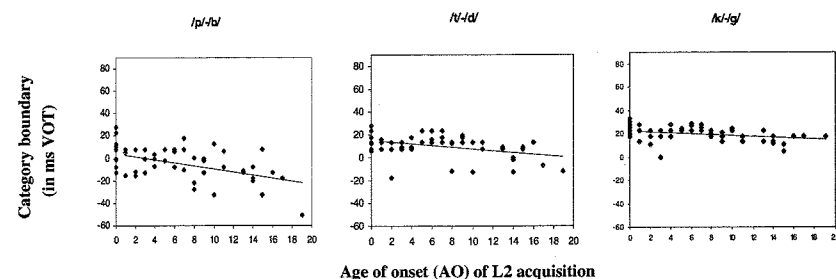


Figure 1. Age of onset (AO) in relation to mean category boundaries (in ms VOT) for the bilabial, dental and velar stop continuum; the values at AO 0 represent category boundaries of the 15 native speakers.

As can be seen in Figure 1 correlations between AO on perceived category boundary are existent for both the /p/-/b/ ($r = -.468$, $p < .01$) and the /t/-/d/ ($r = -.340$, $p < .05$) contrast, whereas the correlation for the /k/-/g/ contrast did not reach significance ($r = -.291$, $p < .069$).

In order to compare the nativelike candidates in a more systematic way, the subjects were divided into a group of early (AO 1-11) and late learners (AO 13-19). Group comparisons revealed that the control subjects change phoneme categories at the longest mean VOTs (+7.23ms for /p/-/b/; +15.34ms for /t/-/d/; +24.62ms for /k/-/g/), while the late L2 listeners show the shortest category crossover points (-17.57ms for /p/-/b/; +1.28ms for /t/-/d/; +14.86ms for /k/-/g/). The group of early learners changes category boundaries at VOTs somewhere in between the late learners and the controls (-2.93ms for /p/-/b/; +10.74 for /t/-/d/; +20.17ms for /k/-/g/). An ANOVA confirmed that these group differences were highly significant for both the bilabial ($F(2,52) = 11.807$, $p < .001$), the dental ($F(2,52) = 7.847$, $p < .001$) and the velar stop contrast ($F(2,52) = 8.815$, $p < .001$). Post-hoc comparisons (Fisher's LSD) showed that except for comparisons between the native speakers and the early learners in case of the dental stop contrast all remaining group differences were significant.

However, as can be seen in Figure 1 most of the nativelike candidates change categories at estimated VOTs within the range of native speaker categorization. This holds for both the bilabial (30 subjects), the dental (32 subjects) and the velar (29 subjects) voicing contrast.

Whereas most of the early learners (21 out of 30) perceive category boundaries within the range of native-speaker categorization for all three places of articulation this applies to only two of the ten late learners (AO 13 and 16). In the group of early learners nine subjects show category boundaries within the range of native-speaker categorization for either one or two of the Swedish minimal pairs. At the same time no early learner was found who exhibits non-nativelike category crossover points for all three places of stop articulation. Finally, the analysis of the group of late L2 learners shows that seven individuals change phoneme category within the range of native-speaker categorization for either one or two of the three places of articulation. In contrast, only one subject (AO 14) does not exhibit category boundaries within the range of native-speaker categorization for any of the stops.

4 Summary and conclusions

The present study has shown that age of onset has an effect on apparently nativelike L2 speakers' categorical perception of the voicing contrast in Swedish word initial stops. In addition to negative correlations between AO and perceived category boundaries, significant group differences were found. The late L2 learners change phoneme category at the shortest crossover points, thereby deviating the most from the Swedish controls. In short, the data confirm that there is a general age effect on categorical perception even among L2 speakers who seem to have attained a nativelike L2-proficiency (Research Question 1).

Among the late L2 learners only two subjects (AO 13 and 16) change stop category within the range of native-speaker categorization regarding all three places of articulation. Thus, only a small minority of late, apparently nativelike L2 speakers show actual nativelike behavior concerning the categorical perception of the voicing contrast (Research Question 2).

Most of the early L2 learners change category for the three stop continua at VOTs within the range of native-speaker categorization. On the contrary, no subject with an early AO was identified who showed non-nativelike category boundaries for all three stop continua. Thus, most, but far from all, early learners show nativelike behavior when their perception of the L2 is analyzed in detail (Research Question 3).

References

- Abrahamsson, N. & K. Hyltenstam, 2006. Inlärningsålder och uppfattad inföddhet i andraspråket – lyssnarexperiment med avancerade L2-talare av svenska. *Nordisk tidskrift för andrespråksforskning* 1:1, 9-36.
- Abramson, A. & L. Lisker, 1973. Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics* 1, 1-8.
- Flege, J.E. & W. Eefting, 1987. Production and perception of English stops by native Spanish speakers. *Journal of Phonetics* 15, 67-83.
- Ruben, R.J., 1997. A time frame of critical/sensitive periods of language development. *Acta Otolaryngologica* 117, 202-205.
- Schneider, W., A. Eschman & A. Zuccolotto, 2002. *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools, Inc.
- Werker, J.F. & R.C. Tees, 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development* 7, 49-63.
- Zampini, M.L. & K.P. Green, 2001. The voicing contrast in English and Spanish: the relationship between perception and production. In J.L. Nicol (ed.), *One Mind, Two Languages*. Oxford: Blackwell.

Stress, Accent and Vowel Durations in Finnish

Kari Suomi

Department of Finnish, Information Studies and Logopedics, Oulu University

kari.suomi@oulu.fi

Abstract

The paper summarises recent research on the interaction of prominence and vowel durations in Finnish, a language with fixed initial stress and a quantity opposition in both vowels and consonants; to be more accurate, the research has been conducted on Northern Finnish. It is shown that, in one-foot words, there are four statistically distinct, non-contrastive duration degrees for phonologically single vowels, and three such degrees for phonologically double vowels. It is shown that the distributions of these duration degrees are crucially determined by moraic structure. Also sentence accent has a moraic alignment, with a tonal rise occurring on the word's first mora and a fall on the second mora. It is argued that the durational alternations are motivated by the particular way in which accent is realised.

1 Introduction

In Finnish word stress is invariably associated with the initial syllable, and there is a binary quantity opposition in both vowels and consonants, independent of stress, effectively signalled by only durational differences. There are very good grounds for interpreting the quantity oppositions syntagmatically, as distinctions between a single phoneme and a double phoneme, i.e. a sequence of two identical phonemes (Karlsson, 1969). This interpretation is also reflected in the orthography, and thus there are written words like *taka*, *taaka*, *takka*, *taakka*, *takaa*, *taakaa*, *takkaa*, *taakkaa*. However, the orthography only indicates the contrastive, phonemic quantity distinctions and, beyond this, it does not in any way reflect the actual durations of phonetic segments. Thus, for example, the orthography or a phonemic transcription do not in any way express the fact that, in e.g. the dialect discussed in this paper, the second-syllable single vowel in *taka* has a duration that is almost twice as long as that in *taaka*, *takka* and *taakka*. This paper summarises recent research on such non-contrastive vowel duration alternations, and suggests their motivations. The paper only looks at vowel durations, and only in words that consist of just one, primary-stressed foot, and thus the effect of secondary stress on vowel durations, which has not been systematically examined, is excluded.

As will be seen below, the mora is an important unit in Finnish prosody. The morae of a syllable are counted as follows: the first vowel phoneme – the syllable nucleus – is the first mora, and every phoneme segment following in the same syllable counts as an additional mora. Below, reference will be made to a word's morae, and e.g. the words *taka*, *taaka* and *taakka* have the moraic structures CM_1CM_2 , $CM_1M_2CM_3$ and $CM_1M_2M_3CM_4$, respectively (where M_n refers to the word's n^{th} mora, and C is a non-moraic consonant).

2 Vowel duration patterns

Suomi & Ylitalo (2004) investigated segment durations in unaccented, trisyllabic nonsense words that consist of one foot each. The segmental composition of the nonsense words was