A Resource for Learning Swedish Oral Skills

Linda Ösp Heimisdóttir[1], Cecilia Ovesdotter Alm[2,], Ian Alden Coots[3], Kateri Krantz-Odendahl[4]

[1] Department of Linguistics, Cornell University, USA
[2] Department of German Studies, Cornell University, USA
[3] Cornell University, USA
[4] Cornell University, USA

`loh5@cornell.edu, eoa5@cornell.edu, iac9@cornell.edu, kgk35@cornell.edu`

Abstract

This paper discusses an interactive CALL (Computer-Assisted Language Learning) resource for the learning of Swedish oral skills. We present a linguistically sound model of oral skills divided into four levels and a sharable Swedish exercise database that reflects this model. We also discuss corresponding Swedish speech data that capture several dimensions of sociolinguistic speech variations (e.g., gender, native/heritage native/non-native speakership, and region). We also report on an evaluation of this new resource using a methodology that considers learning gain in terms of oral perception and listening skills, as well as four categories of user satisfaction. The paper concludes with thoughts for future directions of CALL technologies.

Index Terms: Swedish, productive and perceptive oral skills, Second Language Studies, Computer-Assisted Language Learning, CALL resource, Less Commonly Taught Language

1. Introduction

Second and foreign language acquisition is a booming industry that, for researchers, presents an opportunity to enhance the learning experiences and outcomes of children and adult learners (both within and outside traditional educational systems). Many languages, especially Less Commonly Taught Languages (LCTLs), suffer from a lack of learning tools that can contribute to enhancing various learners' oral skills in a structured manner. Such is the case for Swedish (a language spoken natively in Sweden and parts of Finland), which is taught as a second language to immigrant learners in Sweden and as a foreign language to learners abroad, for instance at around 30 institutions of higher learning in the USA, and globally at around 200 tertiary institutions in various countries, including for example Germany, China, Estonia, France, Korea, Italy, Japan, Mexico, Russia, and Ukraine [1].

We present a learner resource, the *Swedish Web Audio Lab* (SWAL), that provides structured interactive training with a range of Swedish oral skills for first-year learners of Swedish by tying learners' understanding and enhancement of Swedish oral perception to its oral production. Our main contributions in this paper consist of presenting (1) a linguistically sound four-partite model for oral skills, applicable to language learning, (2) a sharable Swedish exercise database that corresponds to this model, (3) the corresponding Swedish speech data that cover a range of sociolinguistic parameters, and (4) encouraging results for the resource from a two-fold evaluation method that considered both user learning gain and user satisfaction, respectively.

2. Previous research literature

2.1. Research on listening and oral skills in Second Language Acquisition

In a paper about the complexity and importance of listening skills for Second Language Acquisition (SLA), Rebecca L. Oxford [2] writes that "listening is more than just a perception of sounds, although perception is the foundation. Listening also includes comprehension of meaning-bearing words, phrases, clauses, sentences, and connected discourse [2]." She goes on to explain that, although many ungrammatical, reduced or incomplete forms contained in ordinary speech can affect the comprehension abilities of listeners, redundancy may allow them to pick up the meaning of what they missed, and in certain situations listeners also have the opportunity to ask speakers for clarification or repetition [2].

Building on this information, Oxford provides guidelines for creating listening activities for foreign language learners. These include that the activity must have a real, communicative purpose and use authentic language without significantly slower or simpler speech. Also, the activity must require listeners to respond in a meaningful way by saying something, following a command or request, asking a question, etc. [2]. In other words, the activities should approach realistic scenarios and they should attempt to expose students to real-world speech and interlocutor interaction.

Schierloh and Hayes-Harb [3] write that "comprehending speech in a second/foreign language [...] can be difficult due to the challenge of discriminating the phonemic contrasts of the language, lack of familiarity with vocabulary items, and non-native-like phonological, morphological, syntactic, and discourse processing [3]." They also say that a challenge facing all speakers is, e.g. the variable nature of spoken language, caused by regional or foreign accents and the different voices of speakers [3]. Moreover, some studies on foreign language speech perception indicate that talker intelligibility influences listening accuracy [4], [5]. Others have shown gender to be a variable; female talkers have been reported to have higher vowel intelligibility than male talkers [6]. Schierloh and Hayes-Harb's research on native speakers of English adapting German as a second language concludes that although talker familiarity does not seem to be a factor in their ability to comprehend a foreign language, clear articulation promotes comprehension [3]. Brown [7] and Hadley [8] have written useful and inspirational overviews on perceptive and productive oral skills materials in SLA.

2.2. Challenges for learning oral Swedish

Andersson [9] and Engstrand [10] discuss aspects of spoken Swedish. Swedish is rich in segments (around 17 long or short vowels and 17–18 consonant phonemes, with various phonetic

realizations). The sound inventory itself poses one dimension of difficulties for learners of Swedish. For a native speaker of English, for instance, the crowded high front vowel space, which lacks equivalence in English, can present a challenge, especially the distinction between [y:] and [ʉ:] and their short counterparts [9]. As another illustration, separating the Swedish consonants [ɧ] and [ç] is particularly challenging, in part because of existing allophones and perceived perceptional proximity, but also due to varying orthographic representations, and a high degree of variation, depending on factors such as gender and region [11]. Another regional feature of Swedish is the pronunciation of /r/, which has both back and front variants [10]. Moreover, combinations of /r/ and a following dental consonant are generally contracted into a single retroflex sound [9]. These and other factors contribute to challenges for producing and perceiving sounds for learners.

At the word level, a well-known and demanding feature is the distinction between two pitch accents (generally absent in e.g. Finnish Swedish). The acute accent (often called *accent 1*) appears, e.g. in monosyllabic words and words with stress on the last syllable, while the gravis (*accent 2*) applies to, e.g. polysyllabic words with initial stress. Word accents can give different meanings to qualitatively similar words that look identical 'on paper' [9], e.g., *stycken* 'pieces' versus *stycken* 'paragraphs'. Another illustration of a word-level feature is particle stress. Verbs' meanings may differ depending on if they appear with a stressed particle or an unstressed preposition, as in *hälsa på* 'visit' versus *hälsa på* 'greet'.

As is true of other languages, Swedish has many features that have to be taken into consideration at the sentence level, e.g. sentence intonation, emphatic or contrastive stress [10], affect (including its subjective nature and internal variation) [12], coarticulation crossing phrasal boundaries, or intonation in questions versus in statements [10]. At the discourse level, some of the challenges involve understanding and using discourse strategies, engaging in dialogic interaction, mastering pragmatic functions, and comprehending or summarizing spoken texts. Essentially, the challenge posed to a learner at these levels is to develop the ability to understand and use linguistic building blocks appropriately in meaningful oral contexts.

3. SWAL tool, model and resources

This section outlines the types of resources drawn upon or created within the SWAL project.

3.1. The SWAL tool

The SWAL resource described uses the Web Audio Lab Version 1.2 for its implementation. It was developed by Slava Papernom for language instruction, and it is maintained by staff at the Language Resource Center at Cornell University (http://lrc.cornell.edu). Figure 1 has a sample screen shot of the Web Audio Lab exercise interface.

The exercise interface has an advantage over listening-only procedures in that it provides an aggregate approach to oral skills. For our purposes, it can essentially tie Swedish oral perception to Swedish oral production. Users respond orally, for instance in conversation, longer passages, sentences, single words or smaller segments. They can also respond to displayed images. Learners have the option of re-listening and revising their responses. Furthermore, model responses can be provided, allowing users to compare their own responses to models. Responses can be made accessible to the instructor in a separate, convenient online interface where the instructor can inspect work completion or comment on the students' productions either orally or in writing. Learners can

subsequently access the instructor's feedback via the student view in the online interface.

Figure 1. *SWAL task screenshot: meaningful utterance completion with family tree visual and oral fragment*
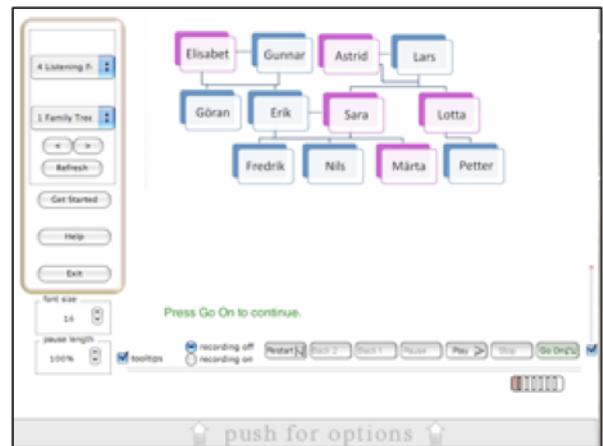
3.2. Oral skills model and exercise corpus

In accordance to the research literature on Swedish phonetics and the adaptation of oral skills (see section 2) as well as instructional experience, a content database of exercises was created to provide an overview of perceptive, productive and conversational oral skills, adapted to Swedish.

The database was categorized into a model that breaks up oral skills into four linguistic levels, as shown in Figure 2: (1) segment/sound level, (2) word level, (3) sentence level, and (4) discourse level. When developing the exercise database, special attention was given to properties of oral Swedish that are difficult for English-speaking learners. We also attempted to present variety in exercises to avoid user fatigue, which is an important factor to consider when users interact with computational tools [12].

Segment/Sound Level, e.g.,	Word Level, e.g.,
• Alphabet • Vowels • Consonants • Length	• Reductions, exceptions • Names, stress • Accents • Compound words
Sentence Level, e.g.	**Discourse Level**, e.g.
• Oral grammar • Filler words • Emotion/emphatic speech • Questions and statements • Sentence intonation, stress	• Dialogue/ask for clarification • Maps • Retelling • Short text comprehension
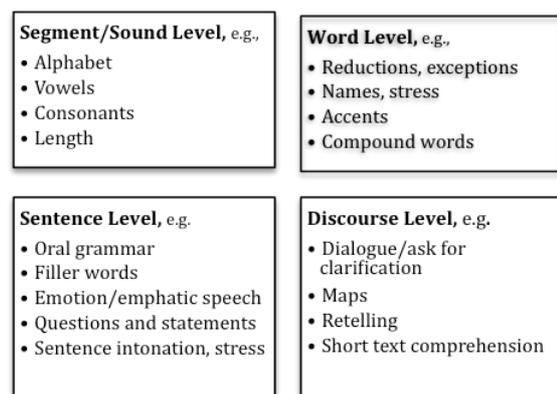
Figure 2: *Model of four oral skills levels with illustrations of trained forms and functions*

The segment/sound level trains, directly or indirectly (e.g. as embedded in words), Swedish phonemes and sound-grapheme correspondences, segment length, and known difficult segments. The word level focuses on word stress and pitch accents, common reductions and exceptions, the composition of words into compounds, stress patterns, and intonation patterns of proper names. At the sentence level, students are trained to listen for lexical, grammatical, phonetic and other cues to apply meaning to sentences. They are also asked to

form sentences of their own, bearing rules of stress and intonation in mind (e.g. with appropriate use of sentence fillers, forming questions or lexical lists, the use of emphatic and contrastive stress, oral grammar comprehension, and emotional speech). Finally, at the discourse level students are asked to tackle various situations, understand and give directions, hold up a conversation, listen for information in short narrations, perform oral retelling, engage with oral trouble-shooting strategies, and stylistic aspects. The exercises also allow learners to engage with level-appropriate semantic fields, e.g., the relationship between weather, clothing, transaction, travel, everyday technological applications, etc.

The following illustrates how users train the different levels by providing a few examples of exercises at each level:

- *Segment/sound level:* Tasks involve, for instance, spelling out spoken words or honing the distinction between proximate vowels by interacting with vowel contrast in words.
- *Word level:* Some of the tasks require users to produce and identify the correct pitch accents after exposure, or practicing how to pronounce proper names of people with correct stress placement.
- *Sentence level:* Tasks at this level include, for example, asking learners to rephrase statements into questions in Swedish after paying attention to how end-tone may set Swedish questions apart from English ones; or sensitizing learners to the dimension of affect in speech, emotion categories, and affective lexical and intonation cues.
- *Discourse level:* tasks may include participating in a conversation that requires a user to ask for clarification, or require users to orally select an image based on a description, or retell information by summarizing a short oral passage.

Naturally, the four levels interact and overlap to some degree. For instance, an electronics sale announcement activity embeds lexical meaning in a meaningful discourse context. This task engages the learner in selective listening with a specified lexical field in mind and with an expressive marketing speech style 'radio' commercial. In addition, users pronounce their answer with select words using list intonation.

3.3. Swedish speech data

The speech data collected to correspond with the exercise corpus was recorded in a studio with expert technical staff. Nine Swedish speakers were chosen for the task and allotted exercises to ensure exposure to variety in terms of speakers' sex, region, speakership (native, heritage native and non-native[1]), and speaker non-familiarity vs. familiarity (instructing voice) to maximize speech diversity. Table 1 provides an overview of speaker characteristics. Speakers were provided texts in advance. They were instructed to read at a natural pace, and the recordings aimed at collecting authentic speech.[2] The recorded speech of the last two speakers (8 and 9) was used for evaluation purposes (see Section 4). All in all, four males and five females participated

as speakers. Of these, five were native speakers, two were non-native speakers, and two were heritage natives, who had grown up with Swedish as a home language in an English-dominant setting. Some of the speakers were project participants, who also participated in the resource development at large.

Speaker	Sex	Nationality	SE area
Sp. 1	M	Native	South
Sp. 2	F	Native	South
Sp. 3	F	Heritage Native	Central
Sp. 4	M	Native	Central
Sp. 5	F	Non-Native	abroad
Sp. 6	M	Non-Native	abroad
Sp. 7	F	Native (instruct. voice)	Central/ North
Sp. 8*	F	Native	Central
Sp. 9*	M	Heritage Native	Central

Table 1. *Characterization of speakers (SE = Swedish). Speech marked by stars was used for evaluation tests.*

4. The usage and evaluation phase

During a period of around three weeks, four second semester Swedish learners (university students) focused on Swedish oral skills, theoretically and practically, with the SWAL as a primary learning tool. The SWAL training took place in 50-75 minute lessons. Students were given handouts on daily topics (http://tinyurl.com/swalwsmaterial) and the training involved brisk-paced group discussion and minor group practice, followed by individual practice with the SWAL resource. Remaining daily SWAL exercises were completed as homework. Feedback was provided, as deemed relevant by the instructor. As part of this evaluation phase, we evaluated SWAL in terms of both learning gain and user satisfaction.

4.1. Learning gain evaluation

To have an understanding of short-term as well as long-term learning gain of the SWAL training, we adminstered a pre-test just before the training initiated, a post-test immediately after it concluded, and a prolonged post-test around 6 weeks later.

The test evaluated learners' performance on most oral skills that the tool was designed to train. It consisted of 50 questions, 14 relating to each of the first three linguistic levels, 7 relating to the discourse level and 1 dictation question (the results of the dictation question are not included with the overall results since dictation could be regarded as a special ability that was never addressed in the SWAL training; rather the evaluation of that item pertains to a separate study).

Test questions focused on the perceptive dimension of the oral skills, which allowed a time and resource efficient evaluation of the tool as well as ensured objective interpretation of test results. Participants provided answers in written form (short answers or multiple choice) that were scored systematically for full or partial accuracy (http://tinyurl.com/swaltestscoresheet). As noted above, students were tested on different voices than those heard in the SWAL resource (see Table 1). To ensure that tests were equivalent in difficulty for comparisons across oral skills, we assigned the same test three times (see with user/password: *enkat* at http://tinyurl.com/swalprepost).[3]

---

[1] Including non-native speech is important for at least two reasons. First, Sweden is an immigration nation and learners will realistically encounter non-native speakers. Second, non-native models may potentially allow students to connect with learning materials and decrease possible speaking anxiety.

[2] Yet, clearer articulation is likely, given prepared texts and the studio setting. Minor retakes occurred. Speakers received a small remuneration for their contributions.

---

[3] This may raise a question of students' familiarity with the test when retaking it, but we estimate that students were unlikely to remember individual words and sentences, in particular given the time between tests.

The results showed an average improvement of around 15% from pre-test to post-test, which were administered immediately prior to and after the SWAL training workshop period. In addition, the results of the prolonged post-test indicated that the learning gain retention was high; on average students scored just over 1% lower on the prolonged post-test than they did on the post-test. Results are shown in Figure 3, indicating the scored performance for the four participants. The results must be interpreted cautiously, given the small sample (after all, Swedish is a LCTL). We do not know if results generalize to a large sample.
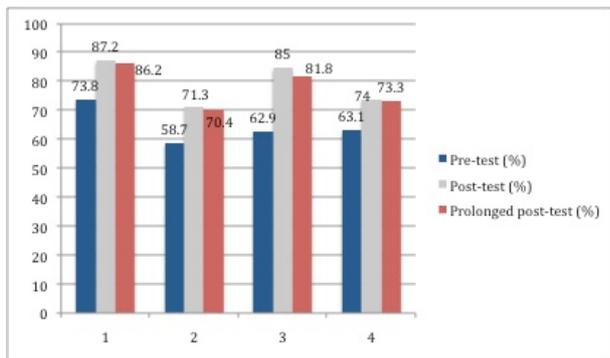
Figure 3. *Learning gain results (max. =98). The average improvement from pre-test to post-test was 14.8% and 13.3% from pre-test to prolonged post-test.*

4.2. User satisfaction evaluation

In addition to the learning gain test, participants completed a satisfaction survey after the SWAL training (http://tinyurl.com/swalsurvey). Users rated SWAL on four dimensions: *usefulness* (including boosting of confidence), *enjoyment*, *fatigue*, and *interest*. In addition, students were given the opportunity to comment on aspects that could be improved or which they were particularly satisfied with. The average satisfaction results can be seen in Figure 4 below.
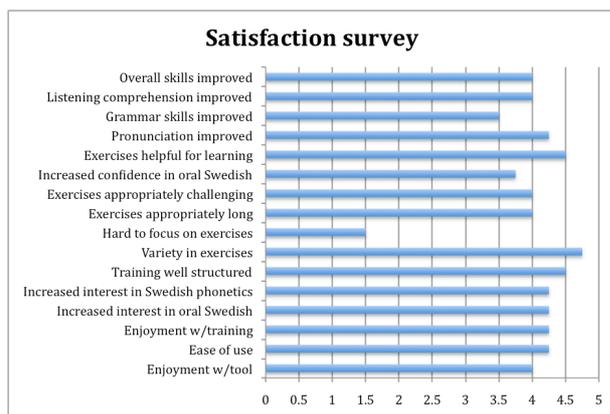

Figure 4. *Average results of user satisfaction survey. Ratings scale: 1=strongly disagree, 5=strongly agree.*

The survey replies indicated that users were in general satisfied with the resource. It scored highest in terms of exercise variety, structure of training, and overall helpfulness for learning and lowest, although still positively, in terms of improvement of grammar skills and increased confidence in oral Swedish. In the open-ended questions, students commented that they liked the exercises that combined perceptive and productive oral skills and that they found the

resource especially useful for enhancing oral comprehension and pronunciation. They also remarked that the resource would benefit from a larger set of exercises, in particular targeting conversational interaction. These comments could facilitate improvements in potential future versions.

5. Conclusion

In this paper, we have introduced a resource that provides structured interactive learning for the enhancement of Swedish oral skills. We have discussed motivations for the project, previous work, the development process, evaluation methods and results. The resource will be made available more widely after minor revisions. A natural step for taking the described resource, or other CALL applications, further would be to customize learning experiences by adding continuous adaptive techniques to the system. This could allow it to flexibly adapt to individual learners' needs and workflows, e.g. by focusing on reinforcing each individual student's weaker areas. Interactive evolutionary computation, which has previously been successfully applied to speech experiments, see e.g. [12], represents a promising area for pursuing such future directions for CALL applications.

6. Acknowledgements

7. References

[1]   Online: http://www.si.se/Svenska-spraket/Svenska/Svenska-spraket/Universitet-i-utlandet/   [Swedish   learning   abroad], accessed on 13 May 2010.

[2]   Oxford, Rebecca L. 1993. Research Update on Teaching L2 Listening. System, 21.2: 205–211.

[3]   Schierloh, Maren and Rachel Hayes-Herb. 2008. Research on Listening Comprehension: The Contributions of Talker Familiarity and Individual Talker Characteristics. UP, 41.2: 171–185.

[4]   Nygaard, Lynne C., Mitchell S. Sommers and David B. Pisoni. 1994. Speech Perception as a Talker-Contingent Process. Psychological Science, 5: 42–46.

[5]   Newman, Rochelle S. and Shannon Evers. 2007. The Effect of Talker Familiarity on Stream Segregation. Journal of Phonetics, 35: 85–103.

[6]   Ferguson, Sarah H. 2004. Talker Differences in Clear and Conversational Speech: Vowel Intelligibility for Normal Hearing Listeners. Journal of the Acoustic Society of America, 116: 2365–73.

[7]   Brown, Douglas. 2007. Teaching by Principles: An Interactive Approach to Language Pedagogy. Pearson Longman ESL, White Plains, NY.

[8]   Hadley, Alice O. 1986. Teaching Language in Context. Heinle & Heinle, Boston, Mass.

[9]   Andersson, Erik. 1994. Swedish. The Germanic Languages. Routledge, London: 271–312.

[10] Engstrand, Olle. 2004. Fonetikens grunder. Studentlitteratur, Lund [selections].

[11] Shosted, R. 2008. Acoustic characteristics of Swedish dorsal fricatives. The Journal of the Acoustical Society of America, 123: 3888.

[12] Alm, Cecilia Ovesdotter. 2009. Affect in Text and Speech. VDM Verlag, Saarbrücken.