# INTRODUCING A COMPREHENSIVE APPROACH TO ASSESSING PRONUNCIATION TALENT

*Matthias Jilka[1], Henrike Baumotte[1], Natalie Lewandowski[1], Susanne Reiterer[2] and Giuseppina Rota[1]*

[1]University of Stuttgart, Germany; [2]University of Tübingen, Germany
jilka@ifla.uni-stuttgart.de, {baumotte, lewandne, rota}@ims.uni-stuttgart.de,
susanne.reiterer@med.uni-tuebingen.de

## ABSTRACT

This paper introduces a comprehensive project with the objective of finding the neural correlates of pronunciation talent. It concentrates on the first part of this undertaking, describing the extensive tests necessary to measure phonetic talent in its various dimensions such as production and perception, the segmental and suprasegmental levels of speech or different utterance forms such as spontaneous speech, reading and imitation. The project also investigates psychological and behavioral influences on pronunciation performance, as well as correlations with general linguistic aptitude. Example tasks making use of the distinction between categorical and realizational differences in intonation are used to demonstrate the detailed analyses allowed by the chosen approach. The described measures allow a reliable classification of talent level to be used in the selection of subjects for the neuroimaging studies in the second part of the project. Additionally, it provides general insights into the interactions between the examined talent-related parameters.

**Keywords:** pronunciation talent, SLA, testing, $F_0$

## 1. INTRODUCTION

Individuals differ greatly from each other in their ability to acquire a foreign, or "second" language (L2). This is especially true for the acquisition of the L2 sound system. The possible explanations for these differences are numerous. They reach from special genetic equipment and particular evolved brain networks to differences in working memory, intelligence and personality factors such as motivation, extraversion or empathy.

The prevalent phonetically oriented studies of second language acquisition (SLA), on the other hand, see the differences less in inherent learner characteristics and focus on external factors such as age of learning, length of residence or amount of L1 and L2 use.

In this paper we indroduce a large-scale project that has the objective of finding the neural correlates of foreign language pronunciation talent. It is intended to contribute to the understanding of the neurobiological basis of language learning ability. In order to achieve this aim, a comprehensive test battery is created that allows the measuring of phonetic talent in its possible manifestations, offering the possibility to look at all these potentially significant factors simultaneously instead of only confirming the relevance of a particular one. The test results will be used to select groups of talented, untalented and average candidates who will perform additional tasks while being examined by means of functional magnetic resonance imaging (fMRI).

This presentation gives an overview of the structure of the project, and focuses especially on the test battery, describing the methods and techniques used to ascertain phonetic talent. The examination of abilities in the production and perception of intonation (in terms of tonal categories and/or their realizations) is used to illustrate the potential of this approach.

## 2. OBJECTIVES AND ASSUMPTIONS

It is generally assumed that language talent consists of different independent linguistic skills.

Especially the phonetic-articulatory aspect seems to be a specific sub-skill for which separate neural substrates have been assumed. A widely acknowledged and fundamental distinction has been drawn between two substrates of linguistic ability: talent for grammar vs. talent for accent [7]. The separate position of pronunciation skills is widely accepted. In fact, a special difficulty of pronunciation acquisition as opposed to that of other aspects of grammar is virtually proverbial, as evidenced by the so-called “Joseph Conrad Phenomenon” (e.g. [1]), which refers to the Polish-born novelist’s native-like abilities in English grammar (syntax, morphology), vocabulary and

style being opposed to his strongly accented pronunciation.

Many individual factors have been shown to contribute to learners' performance in a second language. This does not mean that any one of them, even age of learning onset, can be regarded as the sole determining factor of L2 ability. Within groups of learners who acquire a certain L2 at roughly the same age there will be some who perform better than others. This is most evident in the context of formal learning, e.g. in the classroom, where factors such as age or amount of L1 and L2 use are fairly well controlled. The concept of a critical period has been weakened by studies that show the existence of natural learners (immersed in the L2 culture) who do not achieve native-like competence despite having started acquisition well within the critical period [3] as well as by studies that show that it is possible for late (adult) learners to attain native-like pronunciation ability (e.g., [1]).

## 3. BASIC METHODOLOGY

The tests presented here must be extensive in order to cover all possible expressions of talent and to look for correlations between the many individual factors that have been described as significant in previous studies. Such a large-scale overview should provide interesting results in its own right.

The testing of many “traditional” factors within one and the same experimental setup is intented to identify correlations between these factors and with the overall talent score. The experimental paradigm includes psychological and behavioral tests/questionnaires as well as tasks assessing phonetic and general linguistic aptitude. The testing procedure is expected to reliably express the talent levels of the speakers, facilitating the formation of high, low and average ability groups for the subsequent neuroimaging study.

These experiments require a certain control over the comparability of brain structure in the examined subjects. This is ensured by recruiting the majority of candidates from among university students of English at a German university, who were the same age, had started to learn English at around the age of 10 and had the same learning history (type of instruction, exposure to the L2). For especially talented individuals, however, the only requirement was that they be native speakers of German but not of English.

Indeed, pronunciation skills are assessed mainly in English, but also in German (with respect to intonation). An imitative task in Hindi as a language none of the subjects were familiar with was also included. The use of English has several advantages. First, the large number of subjects familiar with this L2 increases the likelihood of finding talented individuals with native-like pronunciation skills. Second, comparative linguistic descriptions are easily available, both for segmental and prosodic characteristics. Third, it allows the testing of prosodic phenomena, as detailed intonation descriptions and resynthesis tools for German [5] and English [4] are available.

## 4. TESTING NON-PHONETIC FACTORS

Several tests examine factors that are not directly connected to actual phonetic performance, but have been assumed to exert an influence on it. This includes a detailed introspective questionnaire that asks for basic biographical data including educational background and investigates language-related attitudes and experiences, such as general motivation, self-assessment of one’s own abilities, preferred learning styles, the extent of instruction, the extent of L2/L1 experience, use and input, and the attitude towards foreign languages. Psychological tests are carried out as well. Apart from a personality and empathy test, there are also more cognitively oriented tests of verbal and non-verbal intelligence as well as phonological working memory.

General linguistic aptitude is also investigated by means of MLAT - the Modern Language Aptitude Test [2], a test intended to predict success in foreign language learning in native speakers of English.

## 5. TESTING PHONETIC ABILITY

Our investigation of the test subjects' phonetic abilities is intended to cover skills in the segmental and suprasegmental levels to an equal extent.

Testing consists of three major analysis blocks: perception, production, and perception and production combined, i.e. imitation tasks. During the introduction of the various subtasks included in these analysis blocks, some preliminary findings with respect to intonation (within the framework of a category-based model of intonation desription) are presented in order to demonstrate the significance of differences between tone categories

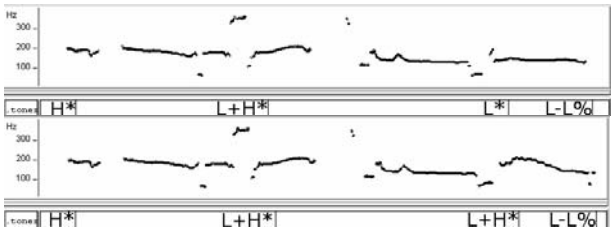


**Figure 1:** Pair comparison of two low pass filtered versions of the utterance "People will get used to the idea". The nuclear pitch accents (L* vs. L+H*) are different.

and the phonetic realizations of identical categories and how they are reflected in the subjects' performances.

### 5.1. Perception

The tasks focusing on perception abilities deal mainly with prosodic features. The only exception is a task of accent identification, in which listeners have to recognize the mother tongue of a non-native reader of the German version of "The North Wind and the Sun". This multiple choice test aims to investigate the listeners' awareness of the characteristic phonetic features of other languages.

Another multiple choice task asks listeners to assign one of four offered interpretations to a presented recording, the only relevant parameter in decision making being prosody.

This is also the case when listeners are simply required to determine whether two presented stimuli are different from each other or not. Both rather obvious categorical and subtle realizational intonation differences are created by means of resynthesis [6]. Control pairs of identical stimuli are of course also included.

Both interpretation and difference identification tasks are administered in English and German, the latter task also using low pass filtered stimuli, such that any possible segmental influences are excluded. The difference detection task is based on the assumption that categorical differences are easier to perceive than mere phonetic differences

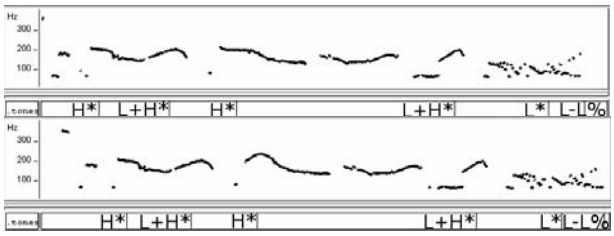


**Figure 2:** Pair comparison of two low pass filtered versions of the utterance "The state is not planning on putting more police on the road". The third pitch accent (H*) is realized with a later peak in the bottom version.

that only consist in a deviating realization of the same tone category. The results of the tests will show whether this is indeed always the case. Nevertheless talented listeners should be more successful in the more difficult tasks.

Figures 1 und 2 depict example stimulus pairs for both possibilities. In Figure 1 there are different nuclear pitch accents (L* vs. L+H*) in the two versions, whereas in Figure 2 a realizational difference in the third pitch accent, H*, (later alignment of the peak) can be observed.

### 5.2. Production

Excluding imitation tasks for the moment, this analysis block restricts itself to those tasks and elicitation techniques in which the subjects produce speech without a given acoustic model. This includes spontaneous speech, which should reflect natural overall abilities the best. Quasi-spontaneous output is also elicited by having the subjects narrate a short cartoon.

The other major technique is reading. The selected tasks concentrate on short, mid-length and long sentences, as individual words do not present enough of a challenge to most readers and longer paragraphs are notoriously difficult to evaluate. Reading tasks have the advantage of making the use of avoidance strategies with respect to problematic sounds, words or sentence structures much more difficult. They facilitate full coverage of the phoneme inventory, allophonic variations and phonotactic constellations of the L2, i.e., English. They also permit a reasonably controlled elicitation of pitch accent distribution and tunes (i.e., combinations of pitch accents and boundary constellations) associated with particular discourse situations (e.g. declaratives, Yes/No-questions, continuation rises) and to enforce specific interpretation-dependent pitch accent assignments.

Unlike pure perception tasks where automatic

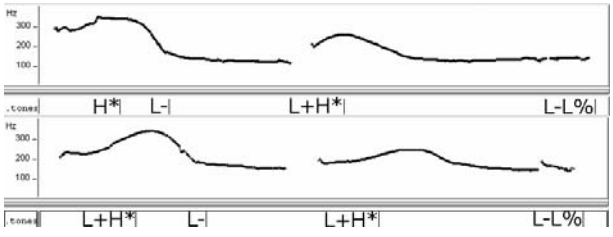


**Figure 3:** Delayed imitation of the utterance "Not again" The second rising pitch accent (L+H*) occurs distinctly later than in the original.

scoring can be used, the evaluation of production tasks requires rating by native listeners (as well as targeted acoustic analysis). An 11-point rating scale will be applied for this purpose, following the recommendation of [8] who claim that this best exploits listeners' full range of sensitivity. Ratings that are within two standard deviations of the mean ratings for native speakers will be accepted as native-like [1].

### 5.3. Perception and production combined

Imitation tasks demand both perception and production abilities, with the former being a prerequisite for the latter. They are especially suitable to the testing of complex tonal constellations and difficult segments as well as of subtle phonetic variation on both the segmental and suprasegmental level.

Prosodic variations are created by means of resynthesis and based on detailed decriptions of the intonation systems of German [5] and English [4]. Two types of imitation task are performed: direct and delayed imitation.

Direct imitation requires the test subjects to simply reproduce a stimulus immediately after having perceived it. However, this approach is often criticized as not being a reflection of true linguistic skill, producing behavior that temporarily exceeds actual competence.

In order to prevent direct imitations from sensory memory, the method of delayed imitation is used as well. This elicitation technique follows the template of a short dialog between a speaker A and a speaker B. After speaker A's utterance is repeated, the test subject is required to imitate speaker B. This procedure is argued to avoid the bypassing of linguistic abilties while at the same time retaining the advantage of challenging the speaker/listener to a phonetically exact reproduction of the original.

Both elicitation techniques are applied to English and German stimuli. Their effects in investigating speaker's talents in imitating prosody are presented in Figure 3. It shows the delayed imitation of the exclamation "Not again" which is not completely successful as evidenced by the later alignment of the rising pitch accent on "again".

## 6. SUMMARY

This paper introduces a large-scale project that pursues the objective of finding the neural correlates of pronunciation talent. It describes the structure of the fundamental first part of the project which covers the extensive testing necessary in order to measure phonetic talent in its various manifestations. The study presents a number of tasks that investigate the test subjects' abilities in production and perception, differentiating between the segmental and suprasegmental levels of speech and different utterance forms such as spontaneous speech, reading and imitation. Psychological and behavioral influences on pronunciation performance, as well as correlations with general linguistic aptitude are also considered. Example analyses making use of the distinction between categorical and realizational differences in intonation are used to demonstrate the detailed investigations made possible by the chosen approach.

While the project is well under way, its extent does not permit quick results. The assessment of pronunciation talent primarily serves the purpose of selecting candidates for the later neuroimaging studies. The study does, however, indicate the project's potential in providing general insights into the interactions between the examined talent-related parameters.

## 7. REFERENCES

[1] Bongaerts, T., Planken, B., Schils, E. 1995. Can late starters attain a native accent in foreign language? A test of the critical period hypothesis. In: Singleton, D., Lengyel, Z. (eds), *The Age Factor in Second Language Acquisition.* Clevedon: Multilingual Matters, 30-50.

[2] Carroll, J., Sapon, S. 1959. *Modern Language Aptitude Test (MLAT): Manual.* New York. The Psychological Corporation

[3] Flege, J., Munro, M., McKay. I. 1995. Effects of age of second language learning on the production of English consonants. *Speech Communication* 16:1-26.

[4] Jilka, M., Möhler, G., Dogil, G. 1999. Rules for the Generation of ToBI-Based American English Intonation. *Speech Communication* 28: 83-108.

[5] Mayer. J. 1997. Intonation und Bedeutung. PhD Dissertation. *Arbeiten des Instituts für Maschinelle Sprachverarbeitung* 3(4). University of Stuttgart

[6] Möhler, G. 1998. Theoriebasierte Modellierung der Deutschen Intonation für die Sprachsynthese. PhD Dissertation. *Arbeiten des Instituts für Maschinelle Sprachverarbeitung* 4(2), University of Stuttgart

[7] Schneiderman, E., Desmarais, C. 1988. A Neuropsychological Substrate for Talent in Second-Language Acquisition. In: Obler, L., Fine, D. (eds), *The Exceptional Brain: The Neuropsychology of Talent and Special Abilities*. New York: Guildford, 103-126.

[8] Southwood, M., Flege, J. 1999. Scaling Foreign Accent: Direct Magnitude Estimation versus Interval Scaling. *Clinical Linguistics and Phonetics* 13: 335-349.