Language Talent and Brain Activity
Trends in Applied Linguistics

Editors
Ulrike Jessner
Claire Kramsch

Mouton de Gruyter
Berlin · New York
Language Talent and Brain Activity

Edited by
Grzegorz Dogil
Susanne Maria Reiterer

Mouton de Gruyter
Berlin · New York
Language talent and brain activity / edited by Grzegorz Dogil, Susanne Maria Reiterer.
p. cm. — (Trends in applied linguistics ; 1)
Includes bibliographical references and index.
ISBN 978-3-11-020518-3 (hardcover : alk. paper)
P118.2.L3648 2009
612.8’2336—dc22
2009021786

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

ISBN 978-3-11-020518-3
ISSN 1868-6362

© Copyright 2009 by Walter de Gruyter GmbH & Co. KG, 10785 Berlin
All rights reserved, including those of translation into foreign languages. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.
Cover design: Martin Zech, Bremen, using a photo by Roswitha Schacht/morguefile.com.
Typesetting: Dörlemann Satz GmbH & Co. KG, Lemförde.
Printed in Germany.
Preface

This volume reflects an interdisciplinary scientific journey into a fascinating area of language talent. Officially, the journey started in February 2006, when the German Research Council (Deutsche Forschungsgemeinschaft-DFG) approved our grant proposal for a project “Gehirn-Korrelate der phonetischen Fremdsprachenbegabung”, shortly “The Brain & Language Pronunciation Talent”. All three of us (Ackermann, Dogil & Grodd) have a research record on the biological foundations of language; however, for a very long time we approached this topic from the perspective of cortical deficiency illustrated by aphasic, dysarthric and other populations with cortical disability. Around 2004 we were confronted and challenged by Dr. Susanne Reiterer with an idea of investigating biological foundations of language faculty from the perspective of cortical efficiency exemplified by especially talented language users. We discussed this new research perspective with a number of colleagues and students and were struck by its implications and its potential. The project started at a usual scale with Dr. Reiterer as a principal co-investigator in Tübingen and Dr. Giuseppina Rota as a co-investigator in Stuttgart. However, peer-to-peer propaganda before it ever started has changed the project’s personal situation quite dramatically. Dr. Reiterer’s research idea was apparently so attractive that a whole “platoon” of young, multilingual, multicultural and exceptionally talented researchers joined in. Dr. Matthias Jilka, who established himself as a linguistic-phonetic leader of the project, was joined by Volha Anufryk, Henrike Baumotte, Natalie Lewandowski, all supported by a DFG Graduate School “Linguistic representations and their interpretation” in Stuttgart. Dr. Reiterer “recruited” Davide Nardo and Xiaochen Hu, both psychologists supported by DFG grants in Tübingen. Our extremely efficient student assistants Arev Anus and Jagoda Sieczkowska managed the pool of over 100 subjects, most of whom were phonetically recorded, psychologically tested and even brain scanned during the early phases of this project. The enthusiasm of all these people has enabled us to investigate “language talent” from several different angles. This book will show you how exceptional performance is correlated with personality, musicality, intelligence, memory span, sociolinguistic, psycholinguistic and detailed phonetic skills of more and less talented speakers. It will also give you new ideas about tests and methods of investigating language talent. We will try to convince you that
efficiency in language learning (talent) is correlated with specific cortical efficiency which can be measured by neuroimaging methods. We even provide you with an example of a “language genius” which shows that once all the factors that foster talent come into play, there is virtually no limit to the number of languages that you can learn to speak – without an accent.

Enjoy

Grzegorz Dogil, Hermann Ackermann, Wolfgang Grodd and Susanne Maria Reiterer
Stuttgart and Tübingen, in May 2009
# Table of contents

- Preface ................................................................. v
  *Grzegorz Dogil, Wolfgang Grodd, Hermann Ackermann and Susanne Maria Reiterer*

- Short portraits of the authors ................................... ix

- Talent and proficiency in language ............................. 1
  *Matthias Jilka*

- Assessment of phonetic ability ................................. 17
  *Matthias Jilka*

- Cognitive aspects of language talent .......................... 67
  *Giuseppina Rota and Susanne Maria Reiterer*

- Personality and pronunciation talent ........................ 97
  *Xiaochen Hu and Susanne Maria Reiterer*

- Functional imaging of language competent brain areas ... 131
  *Wolfgang Grodd, Dirk Wildgruber and Vinod Kumar*

- Colour figur section ............................................. 154

- Brain and Language talent: a synopsis ....................... 155
  *Susanne Maria Reiterer*

- Foreign Accent Syndrome FAS:
  An incidental “speech talent” following acquired brain damage .... 193
  *Bettina Brendel and Hermann Ackermann*

- Musicality and phonetic language aptitude .................. 213
  *Davide Nardo and Susanne Maria Reiterer*

- Sociolinguistic factors in language proficiency:
  phonetic convergence as a signature of pronunciation talent ... 257
  *Natalie Lewandowski*

- Segmental factors in language proficiency:
  coarticulatory resistance as a signature of pronunciation talent ... 279
  *Henrike Baumotte*
Table of contents

Prosodic factors in language proficiency: intonational variation as a signature of pronunciation talent . . . . 305
Volha Anufryk

Direct brain feedback and language learning from the gifted . . . 337
Giuseppina Rota

Beyond talent: a short language biography of Prof. Max Mangold 351
Grzegorz Dogil

Index ............................................................. 359
Short portraits of the authors

Ackermann, Hermann
Professor, Research Group Neurophonetics, Department of General Neurology, Hertie Institute for Clinical Brain Research – University of Tübingen, Germany. »Habilitation« in Neurology at the University of Tübingen. Research interests: cerebral correlates of speech motor control (articulation and phonation), cerebral processing of speech sounds (spatial multimodal stimulus configurations).

Anufryk, Volha
PhD student, funded by the International Graduate School 609 »Sprachliche Repräsentationen und ihre Interpretation« at the Department of Experimental Phonetics, Institute for Natural Language Processing – Universität Stuttgart, Germany. MA from Minsk State Linguistic University, Belarus. Research interests: language variation and change, multilevel processing and production of prosody, second language acquisition.

Baumotte, Henrike
PhD student, funded by the International Graduate School 609 »Sprachliche Repräsentationen und ihre Interpretation« at the Department of Experimental Phonetics, Institute for Natural Language Processing – Universität Stuttgart, Germany. Magister Artium in Clinical Linguistics from the University of Bielefeld, Germany. Research interests: phonetics, linguistics, coarticulation, neurolinguistics, brain imaging, speech therapy, aphasia, second language acquisition.

Brendel, Bettina
Researcher, Research Group Neurophonetics, Department of General Neurology, Hertie Institute for Clinical Brain Research – University of Tübingen, Germany. PhD in Neurophonetics from the Institute of Phonetics, Ludwig-Maximilians-University, Munich, Germany. Research interests: cerebral correlates of speech motor control, apraxia of speech, dysarthria.
Dogil, Grzegorz
Professor, Department of Experimental Phonetics, Head of the Institute for Natural Language Processing – Universität Stuttgart, Germany. »Habilitation« in General and Applied Linguistics at the University of Bielefeld, Germany. Research interests: phonetics, phonology, clinical linguistics, neurophonetics.

Grodd, Wolfgang
Professor, Head of the Section for Experimental MR of the CNS, Department of Diagnostic and Interventional Neuroradiology – University of Tübingen, Germany. Research interests: structural and functional imaging of the brain, functional organisation of the sensorimotor system and the cerebellum, imaging of emotional and cognitive processing.

Hu, Xiaochen
PhD student, Graduate School of Neural & Behavioral Sciences, International Max Plank Research School, Hertie Institute for Clinical Brain Research – University of Tübingen, Germany. MSc in Neuro-cognitive Psychology from the University of Munich, Germany. Research interests: the neurobiological substrates of second language pronunciation talent in combination with the regulatory effect of personality on the second language production process.

Jilka, Matthias
Assistant Professor and Researcher, Department of English Linguistics – Universität Stuttgart, Germany. PhD in Linguistics from the Institute for Natural Language Processing, Stuttgart. Research interests: second language acquisition (forms and causes of foreign accent); text-to-speech synthesis; perception of prosodic phenomena, esp. intonation contours; speech sounds in the languages of the world.

Kumar, Vinod
PhD student, Graduate School of Neural & Behavioral Sciences, International Max Plank Research School – Universitïy of Tübingen; Section of Experimental MR of the CNS, Department of Diagnostic and Interventional Neuroradiology – University of Tübingen, Germany. MSc in Bioinformatics, India. Research interests: neuroimaging, mapping the functional and connectional brain architecture, especially the corticocerebellar interactions.
Lewandowski, Natalie

PhD student, funded by the International Graduate School 609 »Sprachliche Repräsentationen und ihre Interpretation« at the Department of Experimental Phonetics, Institute for Natural Language Processing – Universität Stuttgart, Germany. MA in Applied Linguistics from the Adam Mickiewicz University, Poznañ, Poland. Research interests: phonetics, linguistics, acquisition of pronunciation in second and foreign languages, speaker alignment in dialog, phonetic convergence, English as an L2 in native speakers of German.

Nardo, Davide

Researcher, Department of Psychology »Sapienza« – University of Rome, Italy. PhD in cognitive neuroscience of music processing from the Universities of Tübingen and Rome. Research interests: cognitive, experimental and biological psychology, neural correlates of music processing, language talent and musicality, brain imaging.

Reiterer, Susanne Maria

Researcher, Principal Investigator, Department of Diagnostic and Interventional Neuroradiology; Department of General Neurology, Hertie Institute for Clinical Brain Research – University of Tübingen, Germany. Postdoctoral research fellow in the Section for Experimental MR of the CNS. MA in Linguistics, PhD in Psychology, Cognitive Neuroscience of Second Language Processing, from the University of Vienna, Austria. Research interests: neurolinguistics, second language acquisition, bilingualism and the brain, language aptitude, individual differences in expertise of language learning, brain imaging, cognitive neuroscience.

Rota, Giuseppina

Researcher, Institute for Natural Language Processing, Stuttgart – Universität Stuttgart, Germany; Institute for Behavioral Neurobiology and Medical Psychology – University of Tübingen, Germany. Ph.D. in cognitive neuroscience from the Institute for Natural Language Processing, Stuttgart, Germany. Research interests: learning and brain plasticity, language talent, brain imaging, cognitive neuroscience.
Talent and proficiency in language

Matthias Jilka

1. The significance of language and speech abilities

It is common knowledge that the way in which someone makes use of language and speech contributes significantly to our overall impression of their general personality. Besides obvious characteristics such as rhetorical abilities and the content of what is being said, such impressions are also conveyed by the manner in which one speaks. This concerns both general voice characteristics, which transport extra- and para-linguistic information, and the actual phonetic realization of speech in terms of prosodic and segmental structure. Depending on the sociolinguistic environment, a certain kind of pronunciation that adheres to a particular norm is valued and contributes to social success. The ability to switch between different sociolects or dialects may be of great advantage in the avoidance of communication difficulties, i.e. misunderstandings, misjudgments, or even the experience of prejudice and discrimination. Similarly, knowledge of a foreign or “second” language (L2) is generally beneficial.

Assuming comparable learning conditions, adult individuals vary greatly in their aptitude and speed in acquiring a different dialect or language. This is especially true for the acquisition of the L2 sound system (including all segmental and prosodic manifestations on both the phonetic and phonological level). The possible explanations for these differences are numerous. They reach from special genetic equipment to particular evolved brain networks, differences in declarative (i.e. learning and use of fact and event knowledge) and procedural (i.e. acquisition and expression of motor and cognitive skills) memory (Ullman 2007), intelligence and personality factors such as motivation, extraversion or even empathy. The prevalent phonetically oriented studies of second language acquisition (SLA), however, focus less on inherent learner characteristics, emphasizing instead external factors such as age of learning, age of arrival, length of residence or amount of L1 and L2 use. With respect to these external factors it should thus be possible to create conditions that are optimal for the acquisition of a second language.
The assumption of factors inherent to the speaker, on the other hand, leads to the conclusion that some aspects of language learning ability are immune to external influences. This is what would generally be termed “talent” or, more weakly, a disposition toward good performance in language-related activities.

2. Manifestations of talent in general and in particular

2.1. Basic conceptions

As already indicated, it is generally agreed that a distinction has to be drawn between proficiency, i.e. the overtly observable performance of a particular skill, and talent per se. Factors such as motivation, practice and experience contribute to the degree of the proficiency but are not part of the talent. The consequences of this view correspond with popular concepts that consider talent to be an innate, somewhat mysterious ability that a person either has or does not have.

It should also be noted that the acceptance of a certain skill as a talent is culture-dependent (e.g. Fein & Obler 1988). The traditional occidental idea of talent covers such classic skills as those in painting, music or writing. Accordingly, talents are perceived differently in different places. The mastery of a second language for example may be less valued in India than in a monolingual society. Also, the appreciation for a certain talent will change over time: while some talents have disappeared (e.g. good sword-fighting skills), others have come into existence (e.g. digital media skills).

The idea that a certain talent is innate and therefore reflected in a person's biological make-up is relatively straightforward with respect to purely physical talent. Relating non-physical abilities to the brain seems to be a logical extension of this line of reasoning, but is not as widespread. While there are some neurologically-oriented studies of talent, or rather genius, as typified by such efforts as the examination of sections of Einstein's brain (Diamond et al. 1985), the study of this field is not yet broadly established. However, investigations of the neurological substrates of language talent have been undertaken for example by Geschwind and Galaburda (1985) in their description of pathological language talent as being related to the increased growth of particular brain areas (triggered by the delayed growth of others).
2.2. Language talent

Influential neuropsychological models of the source and structure of talent are based on a model of distinct faculties where special abilities are adjacent to each other, and linguistic talent is comparable to musical, logical, spatial talent, etc. (Gardner 1983). In the majority of cases, the measurement of language talent simply consists of tests of general ability. However, in individual cases of exceptional, sometimes pathological, language talents (i.e. extremely fast and successful L2 learners) extensive tests have been applied in the attempt to diagnose the exact nature of the skill. Novoa et al. (1988), for example, used a test battery that examined a speaker's abilities with respect to IQ, vocabulary skill, verbal fluency, verbal memory, apprehension of abstract patterns, and learning of code systems. Just as language talent is considered one of many discrete talents within the model of distinct faculties, it is also generally assumed that language talent consists of different independent linguistic skills. These roughly reflect Chomskyian concepts of linguistic modularity and are identified both in different types of language disorders and processes of L2 acquisition. In Long's model of maturation, for example, the loss of competence in acquiring an L2 is cumulative, i.e. increases with age. The deterioration supposedly begins at 6 years of age with phonology being affected relatively early (Long 1990).

2.3. The special status of phonetic skills

The special position of phonetic skills as opposed to other linguistic abilities is widely acknowledged. Typically a fundamental distinction is drawn between two substrates of linguistic ability, described as talent for grammar vs. talent for accent (Schneiderman & Desmarais, 1988). In fact, this generally assumed special difficulty of pronunciation acquisition in contrast to other linguistic features is virtually proverbial, as exemplified by the common use of the term “Joseph Conrad Phenomenon” (e.g. Bongaerts et al. 1995, Bongaerts 1999, Guiora 1990, Abu-Rabia & Kehat 2004), which refers to the Polish-born novelist's native-like abilities in English grammar (syntax, morphology), vocabulary and style being accompanied by his strongly accented pronunciation. For identical reasons the condition is also often named “Henry Kissinger Phenomenon”.

This separation of pronunciation from other L2 skills has been confirmed in a number of experimental studies. Neufeld (1987), for example,
showed that ratings of pronunciation skills did not correlate with the results of general language aptitude tests.

In neurophysiological concepts of second language acquisition (e.g. Schneiderman & Desmarais 1988) the neurological substrates of “grammar” and “accent” are assumed to face different challenges in the acquisition of an L2: while both must display neurocognitive flexibility in order to bypass the system established for L1, in the acquisition of pronunciation there is the additional need to bypass established motor pathways in order to control articulatory movements. This additional effort is claimed to account for both the greater difficulties in acquiring the phonetic aspects of language and the differences between children and adults. Other neurophysiological approaches simply argue that pronunciation is a genetically specified lower order function (Walsh & Diller 1981, as quoted in Flege 1987).

Sociopsychological concepts of language acquisition also attempt to explain the special position of pronunciation. Guiora’s approach (Guiora et al. 1972a), for example, sets pronunciation apart from other linguistic features as speech constitutes a higher manifestation of self-representation (Guiora 1990). Contrary to other aspects of language, pronunciation ability and empathy are influenced by the permeability of ego boundaries (i.e. the enhanced flexibility of psychological processes), constituting the so-called “language ego”. The theory predicts that if the ego boundaries are weakened, pronunciation of non-native speech sounds will improve. Studies provoking an enhancement of ego-permeability by means of hypnosis (Schumann et al. 1978), alcohol (Guiora et al. 1972b) or valium (Guiora et al. 1980) claim to confirm this notion.

Non-phonetic aspects of competence in the acquisition of a second language are generally described to a considerably smaller extent. Most such studies analyze L2 speakers’ abilities in the interface of morphology, syntax and semantics. Birdsong (1992) for example tested non-native speakers’ grammaticality judgments in the L2 and found not only that a large number of individuals showed native-like competence, but also that there was no difference between features/parameters of the L2 that are marked or unmarked in universal grammar. Montrul and Slabakova (2001) come to a similar conclusion in their investigation of L2 competence in aspect interpretation (preterite vs. imperfect tense) of Spanish verbs. Compared to pronunciation, learners show a relative ease in the acquisition of such features of the L2. Indeed, adults are more likely to outperform children in grammar or semantics-related tasks of L2 acquisition. In Ioup and Tansonboon’s comparison (1987) of adults and children acquiring Thai,
for example, adults were more successful mastering Thai syntax, whereas children were better in the acquisition of its tonal system.

As a matter of fact, the phonetic subsystem is generally thought to be more difficult to acquire, as it is assumed to rely mostly on hard-wired biological processes that cannot easily be influenced by conscious learning efforts (with the possible exception of neurofeedback techniques, see also chapter 11). Accordingly, virtually everyone who acquires an L2 after a certain critical period (Lenneberg 1967) will exhibit a foreign accent. There is, however, no agreement regarding the cut-off point, i.e. the age at which accent-free mastery of the L2 on both the segmental and suprasegmental level should still be possible. Within the literature on second language acquisition research, the age limits proposed range from young infancy (Kuhl et al. 1992) over childhood (5 years, Krashen 1973) to puberty or adolescence (Johnson and Newport 1989).

Language acquisition within this critical age period should thus always be successful, leading to native speaker status. True second language acquisition would only occur after this period and as a consequence should not allow the so-called “late learner” to attain a native-like pronunciation. Indeed the Fundamental Difference Hypothesis (Bley-Vroman 1989) postulates an elementary distinction between child language development (L1) and foreign language learning (L2) which is caused by differences in the internal cognitive state of adults vs. children and a resulting change in the language faculty that denies adults access to Universal Grammar. The great majority of phonetics-oriented studies in SLA research, however, reject a strict interpretation of the Critical Period Hypothesis (CPH) and the implication that at best an extremely small number of late L2 learners, possibly pathological talents with deficits in other areas, can acquire phonetic skills that are indistinguishable from native speakers.

The general criticism of the CPH is summarized in Flege (1987), who argues that there is no discontinuity in neural development that coincides with a change in speech-learning abilities. Most importantly, no abrupt differences in L2 learning success can be found in studies among speakers of gradually rising age of learning (e.g. Flege 1995) as would be expected if there were a clear-cut critical period. Instead, Flege suggests the concept of a less strictly defined “Sensitive Period” that accounts for children’s greater abilities in the acquisition of non-native speech. He theorizes that children learn in an auditory rather than phonetic mode as they have less firmly established L1 categories. This would allow them to form more accurate perceptual targets. In addition, an impressive
number of cases of individuals, non-pathological experts who have achieved native-like pronunciation is described in the literature (e.g. Ioup 1995; Abu-Rabia & Kehat 2004), also contradicting the conclusions suggested by a straightforward interpretation of the CPH.

A possible way to reconcile the phenomenon of the critical period with the observations concerning second language acquisition just mentioned, may lie in insights expressed by Hensch (2005, 2008) in his work on the mechanisms underlying critical periods of brain development. It states that critical periods are essential in the consolidation of neural systems by shaping cortical areas crucial for an organism’s development and survival. According to Hensch, the primary cortex (including, for example, the auditory cortex and the motor cortex) is therefore subject to effects of a critical period, and, as a consequence, language-related abilities are affected. Once a particular function is established within the critical period it continues to organize connections to other parts of the brain, on the one hand, and, on the other hand, also inhibits the establishment of competing functions. However, it is not evident that critical periods are also as important with respect to the association cortex (e.g., the connection between motor and sensory areas) which are more directly associated with language. This distinction would allow an explanation of the situation encountered in second language acquisition, namely a clear age-dependent influence that is, however, not equally strong with respect to the different linguistic levels (e.g. phonetics vs. syntax vs. lexicon).

3. Testing language skills

3.1. First and second language skills

Individual differences in language-related performance and especially acquisition ability are typically perceived and investigated with respect to proficiency in a non-native language, although general rhetorical abilities, involving such factors as choice of words (verbal intelligence), syntactic constructions and appropriateness in the pragmatic context can also be seen as expressions of proficiency and talent in the native language.

However, general abilities in speech production and perception as needed in normal communicative situations are by definition considered equal in native speakers not suffering from any language or cognitive disorders. Similarly, all native speakers should also exhibit equal levels of pronunciation skills within their native language or dialect (again with
the exception of those with speech impairments). Exceptional control of pronunciation in the L1 may show itself in the mimicry/imitation of dialects, foreign accents or specific, characteristic voices. While a few studies see such abilities as an expression of a universal pronunciation skill and relate them to L2 pronunciation talent, most research focuses on measuring imitation skill in the L2 alone. A notable exception is, for example, Markham (1997) who investigated talented imitators of L1 dialects and of other people, and was able to show that a considerable number of talented speakers were in fact able to also reproduce L2 speech in a native-like fashion. Another interesting approach is presented by Flege and Hammond (1982), who asked native speakers of English to read English with a Spanish accent in order to test these speakers’ awareness of non-categorical, non-phonological features (such as differences in voice onset time and syllable-final lengthening) of Spanish-accented English and, therefore, by extension, of Spanish itself. The study’s focus was, however, clearly on determining the potential difference in the speakers’ perception of categorical vs. non-categorical features. It did not directly investigate the speakers’ imitative or rather accent mimicry skills. Nevertheless, Flege and Hammond’s results imply a connection between an L1-based awareness of phonetic features and L2 pronunciation skills.

In studies of phonological working memory, on the other hand, investigations of the relationship between outstanding skills in that area in the L1 and L2 language talent are a lot more common (e.g. Papagno & Vallar 1995). As a matter of fact, Service (1992), for example, extends this line of reasoning to conclude that the ability to represent unfamiliar phonological material in working memory underlies the acquisition of vocabulary items in the L2. Similarly, Adams and Gathercole (2000) find that children with better repetition skills of foreign words also exhibit better L1 skills. Also, the concept of the Linguistic Coding Differences Hypothesis (Sparks et al. 1998) maintains that L1 skills are the foundation of successful L2 learning and that overall L1 skill reflects overall L2 skill.

Neuropsychological accounts of L2 performance (e.g. Schneiderman & Desmarais 1988) assume a stronger separation between L1 and L2, postulating that L2 performance will be better if the cognitive pathways established for the L1 are avoided such that a direct interaction of the brain’s inherent language-processing skills with the L2’s properties is preferred. This would make the acquisition of the L2 more similar to that of an L1 (and thus potentially more successful). The predominant models of second language acquisition like the Speech Learning Model (Flege 1995), the Perceptual Assimilation Model (Best 1995) or the Native Lan-
guage Magnet Theory of speech perception and production (Kuhl 1991; Kuhl & Iverson 1995) also share this basic idea and consider the representations established for the L1 to be crucial for the manifestation of foreign accent in the L2, as interference is predicted especially in cases of similar phoneme categories. In contrast to this, completely new categories are claimed to be acquired with greater accuracy. Similar effects can also be shown for prosodic phenomena (e.g. Ladd & Morton 1997; Jilka 2000).

3.2. Distinguishing proficiency and talent

A speaker’s performance is of course testable in many different ways. Abilities can be examined, measured and evaluated in various task types such as reading, comprehension, speaking, grammar etc. Such tests are typically geared toward the foreign learners of a particular language, not toward native speakers (again excluding tests investigating possible language disorders).

In the case of English there are for example the “Test of English as a Foreign Language” (TOEFL) for learners of American English and the Cambridge test “English for Speakers of Other Languages” (ESOL) for learners of British English. In terms of the distinction between proficiency and talent, there is of course the complication already mentioned earlier (see 2.1.) that proficiency as evaluated in a straightforward performance test consists of both inherent (i.e. talent-like), and external factors (such as, for example, amount of L1 and L2 use). It is certainly not a trivial matter to try to separate something like inherent talent – should it exist at all – from these other factors by means of experimental design and then determine what it contributes to overall proficiency.

Accordingly, individual test tasks should be defined and constructed in such a way that the targeted abilities are indeed investigated. A general hypothesis of the major influences on performance is necessary. These influences can be summarized and classified as involving first of all developmental/neuropsychological issues such as age of learning, secondly, psychological factors like motivation and attitude towards the L2, and finally, the important question of practice and experience, i.e. amount of language use. The factor of experience can be extended to the more abstract notion of general linguistic expertise, i.e. a greater familiarity with the wide variety of possible linguistic structures (e.g. unusual types of sounds, syntactic structures etc.) even without concrete knowledge of the specific language in question.
To get at the core of “talent”, these factors would have to be controlled or – even better – completely excluded. In the case of experience this could only be achieved by using tasks where experience is by definition equally high for all test subjects because it involves the L1 or a completely unknown language, where none of the test subjects have any prior experience.

This would, however, not be guaranteed to achieve an absolute neutralization of the factors of experience and practice as the aforementioned problem of expert knowledge still remains. Also, it is only an idealized notion that all native speakers are equal in the amount of experience they have with their L1. In the selection of subjects for such an experiment it is therefore preferable to have a large homogeneous group of the same age and “learning career”, i.e. identical time and circumstances of the acquisition of the L2 (see section 2.1. of the following chapter for more details). Of course, this is just as helpful in accounting for the influence of such factors as age of learning. Nevertheless, it is essential to collect detailed information with respect to all these issues in a questionnaire administered to each of the test subjects such that possible correspondences with performance do not remain undetected.

A similar procedure is applied for the examination of the psychological factors influencing performance. A number of cognitive and socio-psychological tests and questionnaires are carried out in order to assess various aspects of the test subjects’ cognitive abilities (working memory, intelligence, etc.) and personality traits (extraversion vs. introversion, etc.) including motivation and interest specifically in acquisition of languages (see chapter by Hu and Reiterer, p. 97, for a more detailed account of the psychological testing). Clearly, the factor of motivation or lack thereof has a strong influence on performance quality during the actual carrying out of a test. Motivation and ambition, thus, have to be taken into account on several levels, not only in terms of a general personality trait that helps or hinders the overall long-term success in language acquisition. There does not seem to be an obvious way of preventing this factor from having an influence (other than very high financial rewards or death threats to ensure the highest degree of motivation in everyone). On the other hand, it might even be asked whether certain psychological features, such as motivation, that are part of the personality, should be considered part of the talent.

We have strived to take all of these considerations into account in the large-scale experiment of assessing pronunciation talent presented here. The phonetic tasks described in the following chapter are thus either
meant to simply assess general performance or attempt to limit the influence of the outside factors just discussed, while also looking at different aspects of phonetic talent ranging from prosody to segments, from perception to production etc.

3.2.1. Studies of individual factors influencing degree of performance

In the study of second language acquisition, a multitude of investigations has been conducted that have concentrated on different individual factors and shown them to have a significant influence on performance in a non-native language. It was evident that a neurologically based factor like age of learning onset (or rather age of arrival in the context of immigration) could not be the sole determining factor of L2 ability-Birdsong (2006, 2008), for example, discusses the common effect of several age-related effects that include progressive L1 entrenchment, neuro-cognitive development and other biological factors. It is quite apparent that 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 especially clear in the context of formal learning, e.g. in the classroom (e.g. Sparks et al. 1998), where factors such as age or amount of L1 and L2 use can be controlled to a considerable degree. The influential concept of the “critical” period is weakened by these different competence levels of learners, as well the lack of abrupt differences in L2 learning success in speakers of gradually rising age of learning (see also section 2.3). In addition to this, it has been shown that there are natural learners (i.e. immersed in the L2 culture) who nevertheless do not achieve native-like competence despite having started acquisition well within the critical period (Flege 1987) and, on the other hand, that it is possible for late (adult) learners to attain native-like pronunciation ability (e.g. Bongaerts et al. 1995; Amunts et al. 2004). These latter studies do take factors such as talent and other favorable prerequisites (high motivation, specific pronunciation instruction) into consideration, but restrict themselves to the examination of pronunciation performance, often in tasks with a relatively low cognitive load.

A by no means exhaustive survey of the mainly phonetics-oriented studies of second language acquisition demonstrates the variety of individual external factors concerning circumstances of language acquisition and use that have been investigated. They include, again, age of learning (e.g. Johnson & Newport 1989), length of residence/age of arrival (e.g. Flege et al. 1995) or amount of L1 / L2 use (e.g. Piske et al., 2001).
Cognitive/psychological investigations examine, e.g. working memory (e.g. Papagno & Vallar 1995), motivation (e.g. Moyer 1999) or personality factors such as empathy and extraversion (e.g. Edmondson & House 1993). Most studies have been content to demonstrate the significance of the one factor they examined. However, some correlation studies intended to identify the interaction of factors that are significantly correlated with high achievement have been carried out as well, suggesting a complex mix of confounding factors. Analyses performed by Bongaerts and colleagues (Bongaerts 1999; Bongaerts et al. 1995, 2000), for example, find that on the basis of innate talent, specific pronunciation training, high motivation, substantial L2 input, and typological proximity of the L1 are very likely to lead to native-like performance.

3.2.2. Language aptitude

Apart from pure performance tests for L2 speakers like TOEFL or ESOL and research studies investigating the significance of well-defined factors on specific aspects of language performance, there are also tests that do not attempt to assess performance in a particular L2. Instead they are designed to measure general language ability mainly on the basis of the L1, predicting potential success acquiring any L2. These so-called language aptitude tests cover such areas as vocabulary memory, syntactic structure and the coding of symbols and meanings or sounds.

A very well-known representative of this kind of test is the Modern Language Aptitude Test, short MLAT (Carroll & Sapon 1959), which was conceived for native speakers of English in order to predict success/talent at learning a second language. Other language aptitude tests for English speakers are the Oxford Language Aptitude Test and the Defense Language Aptitude Battery (DLAB) used by the American Department of Defense. The Differentielle Sprach-Diagnosticum (DSD), a similar test suite proposed for German (Acker 2001), was assembled from various individual tests on verbal intelligence, grammatical judgment, comprehension etc.
4. Conclusion

In this chapter we have attempted to introduce the overall objective of the presented research project. The notions of talent and proficiency as they are understood in the context of this project are defined. The major challenge is to make sure that the tests administered in this project (and presented in detail in the following chapter) are appropriate in addressing these goals and do not just measure proficiency in a foreign language. To emphasize awareness of this problem a brief analysis of different types of language tests as well as of the research of factors significant in second language acquisition was given. A wide variety of influences including neurophysiological, cognitive, sociopsychological or simply language use-related aspects and their possible interactions was covered. It was made clear that control of as many of these external factors as possible is essential in order to know their contribution and get to the core of the underlying talent. The different types of tasks introduced in the following chapter must therefore be defined clearly with respect to what they are expected to measure.

References

Abu-Rabia, S. & Kehat S.

Acker, B.

Adams, A. M. & Gathercole, S.

Amunts, K., Schleicher, A. & Zilles, K.

Best, C.

Birdsong, D.
Talent and proficiency in language

Birdsong, D.

Birdsong, D.

Bley-Vroman, R.

Bongaerts, T., Planken, B. & Schils E.

Bongaerts, T.

Bongaerts, T., Mennen, S. & van der Slik, F.

Carroll, J. & Sapon, S.

Diamond, M.C., Scheibel, A. B., Murphy, G. M. & Harvey, T.

Edmondson, W. J. & House, J.

Fein, D. & Obler, L.

Flege, J. & Hammond, R.

Flege, J.

Flege, J.
Flege, J., Munro, M. J. & MacKay, I.  

Gardner, H.  

Geschwind, N. & Galaburda, A. M.  

Guiora, A., Brannon, R. & Dull, C.  


Guiora, A., Acton, W., Erard, R. & Strickland, F.  

Guiora, A.  

Hensch, T.  

Hensch, T.  

Ioup, G. & Tansomboon, A.  

Ioup, G.  

Jilka, M.  

Johnson, J. & Newport, E.  
Krashen, S. D.

Kuhl, P.

Kuhl, P., Wiliams, K. & Lacerda, F.

Kuhl, P. & Iverson, P.

Ladd, D. R. & Morton, R.

Lenneberg, E.

Long, M.

Markham D.

Montrul, S. & Slabakova, R.

Moyer, A.

Neufeld, G.

Novoa, L., Fein, D. & Obler, L.

Pagagno, C. & Vallar, G.
Piske, T., MacKay, I. & Flege, J.

Schneiderman, E. & Desmarais, C.

Schumann, J., Holroyd, J., Campbell, N. & Ward, F.

Service, E.

Sparks, R., Artzer, M., Ganschow, L., Siebenhar, D. & Plageman, M.

Ullman, M.

Walsh, T. & Diller, K.
Assessment of phonetic ability

Matthias Jilka

1. Introduction

The project presented here has the objective of providing a comprehensive examination of talent in pronunciation, especially of a second language. The project investigates this ability with respect to its multiple phonetic/linguistic manifestations, but also considers psychological and behavioral influences on pronunciation performance – and, most importantly, also has the objective of finding neural correlates of pronunciation talent, i.e. differences in brain activity between talented and untalented speakers. This chapter focuses on the first part of this undertaking and describes the extensive tests necessary to assess phonetic talent in its various dimensions such as production and perception, the segmental and suprasegmental levels of speech and different utterance forms (such as spontaneous speech, reading, and imitation). The tests are intended to provide general insights into both the nature of pronunciation talent per se and the more specific interactions between the examined talent-related parameters. They are also meant to serve the purpose of allowing a reliable classification of talent and proficiency levels to be used in the selection of subjects for the subsequent neuroimaging studies.

The present description thus begins with general information on the test participants and the languages used, followed by sections detailing the investigation of non-phonetic abilities or conditions possibly correlating with phonetic talent, such as the external circumstances of second language acquisition (e.g. age of learning, amount of language use, type of instruction), the learners’ psychological characteristics (encompassing both personality traits such as degree of extraversion or motivation and cognitive abilities such as working memory or intelligence), and general language aptitude and proficiency in the examined languages. Phonetic abilities proper are then covered in subsections on production, perception and imitation (which combines the first two aspects).
2. Speakers and languages

102 native speakers of German served as test subjects. Among them a core group of 50 university students of English shared a number of key variables such as age (range: 20–23 years), age of onset of L2 English learning (10 years) and type of experience/training in this L2 (English instruction in the formal setting of the German school system, relatively low amounts of experience in English-speaking environments). The remaining test subjects belonged to a wider range of ages and backgrounds and were recruited via ads or personal contacts. Besides general interest in the project, most were chosen due to their self-professed pronunciation talent or lack thereof. 15 native speakers of English also participated in order to provide data for comparisons.

The languages used are mainly English, and to a smaller degree German and Hindi. English was chosen because the large number of learners increases the likelihood of finding individuals with native-like pronunciation skills. Furthermore, comparative (English vs. German) linguistic descriptions, both for the segmental and prosodic characteristics are easily available. The testing of prosodic phenomena in particular is facilitated due to the existence of detailed intonation descriptions in combination with resynthesis tools for both German (e.g. Mayer 1997) and English (e.g. Jilka et al. 1999). The German and Hindi tasks are designed to restrict the influence of learning experience on performance. As native speakers of German the test subjects should have very similar, if not equal amounts of experience with it, whereas Hindi is a language they are all completely unfamiliar with.

3. External circumstances of language acquisition

An extensive online questionnaire is employed in order to gather all the relevant biographical data. Following examples from Moyer (1999) or Markham (1997), the questionnaire also investigates language-related attitudes and experiences that are assumed to influence language ability. Specific questions address issues such as general motivation in language learning and a self-assessment of one’s own linguistic abilities. The test participants are also asked to describe in detail the extent of their experience with any second languages they speak. This includes information about the so-called “Age of Learning”, i.e. the age at which they first started learning, the formal (i.e. type and duration of instruction) and in-
formal (e.g. stays in an L2 environment) circumstances of learning, as well as the amount of L2 use and input in everyday life at the time of the recording (compared to L1 use).

Also included are questions designed to provide information with regard to the test subjects’ attitudes towards the respective non-native languages (the focus is of course on English) and cultures in general. This should yield an impression of the distance between German and English (culture) as perceived by the speakers.

In view of the problematic distinction between the concepts of superficial proficiency and innate talent, information about all aspects of L2 experience is, of course, especially important, as experience can be assumed to make a significant contribution to overall proficiency (see also section 3.2. of the preceding chapter). While there may not be a straightforward method to precisely correlate years of experience or amount of L2 use directly with performance, it is certainly possible to generally account for the effect of larger differences between speakers.

4. Psychological testing

A considerable number of individual psychological characteristics have been shown to correlate with L2 performance. As this project strives to provide a comprehensive picture of all the potential influences on L2 performance, a variety of established psychological tests was administered to the subjects in order to replicate these results and possibly relate them to each other.

Two major types of tests can be distinguished: those examining psychological aspects of the test subjects’ personalities and those focusing on cognitive abilities.

This section presents only a brief overview; a more detailed description can be found in the chapters specifically dealing with personality factors and (cognitive) psychological aspects respectively.

4.1. Personality questionnaire

According to Trait Theory, personality reflects a human being’s individual consistency in behavior across time and situations (Allport 1961). We applied several questionnaires that aim to determine whether certain personality characteristics correlate with linguistic abilities.
The Neuroticism Extraversion Openness Personality Inventory assesses personality traits according to five major factors, namely neuroticism, openness, extraversion, agreeableness and conscientiousness (Costa & McGrae 1992). The BIS/BAS questionnaire (Carver & White 1994), on the other hand, investigates whether a person's actions are driven more by the positive motivation toward something desired (Behavioral Approach System – BAS) or the desire to avoid a negative experience (Behavioral Inhibition System – BIS). A questionnaire assessing empathy, E-Skala (Leibetsderer et al. 2001), is administered to examine if this characteristic is associated with a greater readiness to adapt to unfamiliar phonetic features.

4.2. Cognitive tests

While certain cognitive abilities such as working memory or musicality are often assumed to correlate with linguistic performance, this is not the case for other cognitive skill types, e.g. intelligence. Tests assessing both forms of cognitive abilities were conducted to validate these insights.

As far as intelligence tests are concerned, the Raven Advanced Progressive Matrices Test (Raven et al. 1998), for example, addresses the capacity of non-verbal logical reasoning and recognition. A different type of intelligence assessment, namely verbal intelligence, is performed by the Mehrfachwahl-Wortschatz-Intelligenztest (MWT), which focuses on extended vocabulary knowledge in the native language (Lehrl 2005). The latter test would be expected to show a stronger connection to linguistic, though not necessarily phonetic, performance.

We are also administering standard tests of phonological working memory (Gathercole et al. 1994), as it has repeatedly been argued that “phonological short term memory” can predict success in foreign as well as native language learning (e.g. Baddeley 2003). Many studies have reported such correlations, but their exact nature is still unclear. Research with children has demonstrated that the ability to learn new words is greatly affected by working memory span, a correlation that appears to continue into adolescence (Gathercole et al. 1994).

In addition, a task evaluating musicality by asking the listener to recognize tonal and rhythmic changes in pairs of sound files (Gordon 1989) is also included, as is a test measuring mental flexibility (Simon 1990), defined as the ability to handle different situations in different ways and thus easily adapt to changes, learn from mistakes and be creative in solving problems.
5. General language abilities

There are, of course, already existing and established tests for assessing general linguistic aptitude and quite a large number of tests for simply evaluating overall proficiency in a particular second language. Results from such test batteries should serve as useful comparisons to the investigations in this project, which focus on phonetic/phonological abilities.

5.1. Language aptitude

Tests of language aptitude are designed to provide standardized procedures measuring general language learning abilities, i.e. predicting success at acquiring a second language.

The Modern Language Aptitude Test, short MLAT (Carroll & Sapon 1959), was, for example, conceived to achieve this for native speakers of English. A shortened version of the MLAT is applied to the native speakers of German participating in the experiment. As there is a strong component of English proficiency to be taken into account, the results cannot be interpreted to reflect an inherent aptitude like they would for actual native speakers of English. Nevertheless, they allow for comparisons of the test subjects’ general language abilities such as handling of syntactic structure, vocabulary and the relationship between sounds and letters.

Of the three task types taken from the MLAT test, “Spelling Cues” is the one that includes the phonological component just mentioned. The use of an alternative spelling system is meant to test the ability to associate sounds and symbols (and, in the case of non-native speakers of English, also vocabulary knowledge). The example shown below demonstrates that an alternatively spelled word (“srtn”) must be associated with one of four possible synonyms or otherwise related terms:

srtn   A. foolish     B. honest     C. complaint   D. sure     E. fatal

Also included is a task that tests comprehension of grammatical structure by requiring a comparison of the syntactic functions of elements in different sentences (“Words in Sentences”), as well as a test of verbal memory (“Paired Associates”), in which test participants must memorize 24 Kurdish vocabulary items within two minutes and then pick the correct item out of five choices offered.
5.2. English proficiency

We also applied a small portion of the Test of English as a Foreign Language (TOEFL) in order to get a clearer impression of our German test subjects’ overall English proficiency.

The tasks are, however, limited to a listening comprehension test (participants listen to a lecture of approximately two minutes in duration and then answer questions about the content of what they heard in a multiple choice scenario) and a so-called “structure” test in which both grammaticality judgments and vocabulary knowledge are required.

6. Testing phonetic ability

The tasks intended to assess the test participants’ phonetic abilities were performed in an approximately 90 minute session in the anechoic chamber of the Institute for Natural Language Processing at the Universität Stuttgart (45 minutes of production/imitation tasks + 45 minutes of perception tasks). In the production tasks, the speakers are free to use any native accent of English (General American, Received Pronunciation, Australian English, etc.) they believe they can master. The imitation and perception tasks are evenly divided between GA and RP stimuli.

6.1. Speech production

6.1.1. Types of tasks

In order to get a complete picture of the test speakers’ abilities, various elicitation techniques are applied, following the example of studies like Oyama (1976), Bongaerts et al. (1995), Markham (1997) or Flege and Hillenbrand (1987). This is done primarily to be able to elicit differing types of intonational configurations, speaking rates and degrees of fluency. The degree of foreign accent can therefore be expected to vary as a function of the type of elicitation, thus revealing differences between the speakers according to the complexity of the task.

Assessments of segmental production can of course be made in all controlled environments, though opportunities for comparisons are obviously limited in (quasi-)spontaneous speech where it is only possible to analyze what the speakers happen to say.

In both the reading and imitation tasks, all phonemes, important allophonic variations and phonotactic constellations of the L2, i.e. English,
are covered. The known problematic areas for native speakers of German are investigated in detail. We consider problems like the production of uvular /ʁ/, the lack of dental fricatives (and their replacement by /s/ or /d/, e.g. “the threat” as /dɔ sʁɛt/) or the use of clear /l/ in all positions including the syllable coda to be indicative of a heavy German accent. Other characteristic deviations like the final devoicing of voiced obstruents (e.g. /bɪɡ/ pronounced as /bɪɾk/) or the insertion of glottal stops before morphemes that start in a vowel (e.g. “Ann ate an egg” pronounced as /ʔən ʔɛɾt ʔən ʔεɡ/) and the raising of /æ/ to /ɛ/ are seen as slightly less serious (if only because they are so frequent). Finally, subtle differences in vowel coloring or vowel lengthening before voiced final consonants are at worst interpreted as a weak foreign accent and are often beyond the threshold of awareness of other German speakers.

The tasks intended to assess prosody production use especially imitation and reading tasks to elicit tunes (i.e. combinations of pitch accents and boundary constellations) associated with particular discourse situations (e.g. declaratives, Yes/No-questions, continuation rises). Insights gained in our previous work (e.g. Jilka 2000) facilitate the identification of possible deviations due to foreign accent such as category transfer in the realization of continuation rises or subtle variation in the phonetic realization of equivalent categories. Very few comparative studies actually comment on the relationship between segmental and prosodic skills. There are a number of studies that include prosodic features in evaluations of foreign accent, but only rarely are the results compared with those for segmental features. Ioup and Tansonboom (1987), for example, in their experiment on the acquisition of Thai, observe that children are more successful in the acquisition of tone, whereas adults are more successful at reproducing aspiration features. The great majority of studies that address suprasegmental features in relation to segmental features do so by offering general observations with respect to the contribution of prosody to overall L2 quality. As far as issues of teaching and intelligibility are concerned, Derwing et al. (1998), for example, find that a prosody-oriented global teaching approach yields better results in speech production than an exclusively segment-oriented approach. Similarly, Fraser (2000) argues that, for L2 learners’ productions, word and sentence stress are as important to intelligibility as segmental aspects. In summary, it can be stated that while there are some investigations that describe both the general relevance of prosody in the perception of foreign accent, as well as the significance of very specific instances of intonational deviations in the speech of L2 learners (e.g. Jilka 2000 or Mennen
1999), there are as of yet no insights as to correlations between segmental and suprasegmental talent within L2 learners’ productions.

6.1.2. Quasi-spontaneous speech

Spontaneous speech is by definition the most natural form of human communication and certainly reflects overall abilities the best, allowing especially representative impressions of fluency, speaking rate, segmental realizations and choice of prosodic patterns. At the same time, it is very difficult to elicit specifically desired prosodic or segmental realizations. We attempt to control spontaneous output to a certain degree by having the subjects narrate a short Gary Larson cartoon (Larson 1992), see also Figure 1.

![Cartoon](image)

*Figure 1.* Cartoon to be narrated in English by the native German test subjects (words like “cow”, “doorbell”, “farmer”, “grass” are likely to occur)
which to a small extent suggests vocabulary choices (and thus segmental realizations), and also triggers certain prosodic constellations like continuation rises.

At the start of each recording session, speakers are also asked to introduce themselves and talk about their experiences with the English language/culture. This monologue is then extended into a short conversation with the instructor, resulting in about 4 minutes of quasi-spontaneous speech altogether (the cartoon narration takes up approximately 30 to 90 seconds depending on the speaker). Due to the laboratory situation, the speech production can, of course, not be called completely spontaneous, therefore we have termed these tasks “acted speech”.

6.1.3. Read speech

Reading tasks are useful in that they facilitate a controlled coverage of the phoneme inventory and specific phonotactic constellations, as well as a reasonably controlled elicitation of pitch accent distribution and tunes associated with specific discourse situations. They also prevent the use of avoidance strategies with respect to problematic sounds, words or sentence structures, thus reducing the potential influence of syntactic/morphological errors on evaluation. It has to be taken into account, however, that the produced speech is restricted to the prosodic patterns specific to this style of speech and may also elicit a more careful segmental articulation than “regular” speech.

For the English reading tasks, we use the standard IPA text for segmental coverage “The Northwind and the Sun” and a page-long extract from the Mark Twain short story “Mrs. McWilliams and the Lightning” (Twain 1968). The latter text contains less familiar vocabulary which is also more difficult to pronounce (e.g. words like “unlocatable” or “deliberately”) as well as various dialog situations that elicit question or exclamation intonation, etc. Additional reading tasks consist of individual sentences intended to enforce the production of specific segmental and/or prosodic constellations. An example is the reading of the short phrase “She's teething” in such a way that it reflects different emotional states like surprise, happiness, pity or anger (the pronunciation of the voiced dental fricative also proved to be challenging to a lot of speakers). Such a task obviously emphasizes the importance of overall prosodic choices which may be assessed by ear and not necessarily by means of an instrumental analysis. Figure 2 depicts examples of more (top contour) or less (bottom contour) successful attempts to express surprise.
Equivalent sentence-reading tasks are also applied in German, with a larger focus on the appropriate prosodic choices and realizations, e.g. reading the sentence “Alle Politiker sind nicht korrupt” in such a way that it means either that all politicians without exception are not corrupt or that a lot of politicians are indeed corrupt, but not all of them (the second interpretation requires the use of a hat pattern, see, e.g. ’t Hart et al. (1990) or Féry (1993).

Additionally, the speakers are asked to imitate an English accent while reading “Der Nordwind und die Sonne” in German. This task was inspired by a procedure introduced in Flege and Hammond (1982) and constitutes a form of delayed mimicry that is meant to test the speakers’ awareness of major phonetic and phonological features of other languages and their ability to realize them.

6.2. Speech perception

A speaker’s phonetic ability does not necessarily manifest itself only in the form of speech production. Corresponding skills in speech perception must be considered equally important.

As far as the perception of segmental features is concerned, an investigation of performance in the native language or a relatively familiar sec-
Assessment of phonetic ability

Second language like English is unlikely to yield interesting insights. Discrimination or identification tasks would either be rather simple (as in the case of phonological differences) or irrelevant with respect to everyday language use, i.e. comprehension (as in the case of allophonic variation). Only the identification of foreign or regional accents regularly involves this ability (see also section 6.2.3).

The speech perception tasks included in this project thus concentrate mainly on the detection, comprehension and interpretation of suprasegmental features that may carry semantic, pragmatic and possibly extra-linguistic information. Pitch, in particular, but also duration and loudness contribute to overall phrase intonation, which encodes discourse structure and meaning (see Dogil 2003). In this investigation, intonation description and analysis are based on the framework of the Tone Sequence Model (e.g. Pierrehumbert 1980) as represented by ToBI (Silverman et al. 1992). Intonational differences are therefore interpreted as either categorical (different tone category, e.g. pitch accent) or realisational (same category, different phonetic realization). The existence of detailed intonation “grammars” for English (Jilka et al. 1999) and German (e.g. Féry 1993; Mayer 1997 or Grice & Baumann 2002) as well as a resynthesis tool (Möhler 1998) allows for deliberate phonological and phonetic modifications of the original models without significant distortion effects.

6.2.1. Pair comparisons

The pair comparison tasks are designed to test the listeners’ ability to recognize prosodic differences between otherwise identical utterances. The test participants are simply required to decide whether a presented pair of stimuli consists of two identical recordings or not. The pair comparisons are carried out both in German and English in order to investigate awareness of native and non-native prosody. Tonal differences are created artificially by using resynthesis to modify an original recording. As already indicated above, such intonational differences may be categorical in nature, i.e. a pitch accent or boundary tone may be affected in its entirety, e.g. added, deleted or replaced by another one. Figure 3 shows such a case in the phrase “People will get used to the idea”, where the nuclear pitch accent L* on “idea” is replaced by a rising L+H* accent, creating an additional rise-fall movement. The category distinction should be relatively easy to distinguish. While in both versions “idea” refers to an entity mentioned earlier in the context, the higher promi-
nence triggered by the use of the $L+H^*$ accent in the second version highlights the fact that the possibility of designating this referent an “idea” is something unexpected or unforeseeable by the hearer. It is this insight that represents new information rather than the referent of the expression itself. While in Riester’s information structure annotation scheme (Riester 2008) both versions would be classified as instances of the information status category “d-given” (discourse given), a further subclassification is conceivable.

Tonal differences may also concern only the phonetic realization of tonal categories. Parameters such as, for example, the timing of peaks relative to syllable duration or pitch height might deviate from each other, resulting in an audible difference. In the example in Figure 4, the $H^*$ pitch accent on “planning” in the utterance “The State is not planning on putting more police on the road” occurs relatively later and higher (see bottom contour) than in the original (top contour).

As categorical differences are more clearly associated with changes in interpretation/discourse meaning, it is generally expected that they are
perceived more easily than the more subtle realizational differences (which are harder to distinguish from actually identical pairs) and that more talented listeners would be able to show their greater sensitivity especially with respect to the latter type of difference.

The pair comparison tasks are also applied using low-pass filtered speech, i.e. speech in which the higher frequency components carrying the segmental information in the speech signal are blocked. The listener thus only has access to the prosodic information from the $F_0$ contour. Language-specific differences between German and English are then also limited to prosodic content. It has previously been shown that, if long enough stretches of speech are provided, the given languages can indeed be discriminated or even identified in low-pass filtered speech (see, e.g. Ramus & Mehler 1999; Jilka 2000) such that differences in performance depending on the underlying language are a possibility that can be inves-

**Figure 4.** Example for a pair comparison with a realizational difference between two versions of the utterance “The State is not planning on putting more police on the road”. Both versions have a high ($H^*$) pitch accent on “PLAnning”, but in the bottom version it occurs later and higher, causing a slightly more pronounced up-and-down movement of the contour.
tigated. The main objective is, however, to assess the test subjects’ ability to handle this particular type of strictly prosodic information and recognize differences (additional tests generally examining language identification in low-pass filtered speech are carried out as well).

6.2.2. **Prosody interpretation**

Another aspect of speech perception is tested in a task that is intended to assess experience and awareness in the ability to draw semantic/pragmatic information from the prosodic structure of an utterance without help from a surrounding general context. The listeners need not only be capable of perceiving the phonetic detail inherent to the speech signal, but must also interpret it and associate it with the appropriate discourse meaning. There are two types of task presentation. In the first scenario the listeners hear an utterance, for example “Du hast dafür 35 Euro bekommen” (You received 35 Euros for that), and are then asked to decide which one of four suggested situations/interpretations (e.g. the speaker wants to remind her forgetful friend, thinks this is way too much, thinks this is very funny, or cannot believe it) matches it best. The alternative mode of presentation is more restricted with respect to multiple choice answers, as the test participants listen to two prosodically different versions of an utterance and are required to answer a question concerning their interpretation, e.g. “In which version does the speaker sound angrier?”.

As in the pair comparison, results in the German version of this task might be expected to reflect a greater sensitivity than those in the subjects’ non-native language English, with differences among speakers in German more likely to be indicative of differences in general language awareness.

6.2.3. **Accent identification**

This is the only speech perception task not dealing predominantly with prosodic features. The test participants listen to various versions of “Der Nordwind und die Sonne” as read by non-native speakers of German and are asked to identify the speakers’ respective countries of origin. Four possible answers are offered in a multiple-choice scenario. The task is meant to assess general awareness of and experience with characteristic segmental and prosodic features of languages. The potential answers are chosen in such a way that some should be considerably less likely than others. This is also reflected in the evaluation (see also section 7.5).
6.3. Perception and production combined

Imitation tasks combine the abilities addressed in the two preceding subsections: the correct perception of a model and the ability to reproduce a correct representation of it. The obvious advantage of such tasks is that they allow a maximum amount of direct control over what a speaker will say, thus facilitating the direct testing of complex tonal constellations and difficult segments as well as of subtle phonetic variation on both the segmental and suprasegmental levels.

Besides the expected German and English tasks, there is an additional challenge involving Hindi as a language that none of the test participants were familiar with. The Hindi imitations of words and short phrases focus mainly on segmental aspects, especially the perception of those sounds and phonotactic structures not present in the native language and the ability to reproduce the perceived acoustic patterns.

The German and English imitations emphasize prosody to a much larger extent. They require the speakers to perceive, understand and mimic appropriate prosodic patterns, both in categorical and realizational dimensions. Prosodic constellations of interest (e.g. focus, particular F0 movements, certain types of discourse situations) were created directly by a model speaker or using resynthesis. Prosody imitation is the main challenge in the German tasks, even though many of our native German speakers (those who are actually speakers of the Swabian dialect) also did not find it easy to reproduce the standard High German pronunciation of the model. The English imitation tasks are also designed to test the reproduction of prosodic phenomena. The model utterances reflected the segmental characteristics of the pronunciation of either standard British English (Received Pronunciation) or General American, creating an additional challenge for speakers less familiar with one or both of these accents. It must also be pointed out that general English competence plays a role as comprehension problems may influence the performance.

6.3.1. Direct imitation

The approach of directly imitating a given model is sometimes criticized as not being a reflection of linguistic skill at all since it may produce behavior that temporarily exceeds actual competence (Weismer & Cariski 1985) and has no “carry-over into the post-imitative tasks” (Barry 1989, p. 167). Others, however, are more supportive of the approach. Markham (1997), for example, argues for its use because acquisition itself is a
strongly imitative phenomenon and Kuhl and Meltzoff (1995) show that imitation can result in nonidentical, but functionally equivalent, reproductions of the modelled behavior, a fact that speaks against the view that direct imitation bypasses all levels of linguistic processing. Thus, a number of studies (e.g. Markham 1997; Neufeld 1987) use direct imitation in their evaluations of language skills.

Direct imitation tasks are also part of our test battery. Stimulus presentation is quite straightforward as the test subjects are simply required to imitate an utterance they are presented with. An adjustment is made for the Hindi test in that the stimuli are presented three times in a row before imitation so the speakers have the chance to familiarize themselves a little more with the unknown sounds and sound combinations.

6.3.2. Delayed imitation

Delayed imitation tasks strive to preserve the advantage of all imitations, namely the high degree of control and precision, while avoiding the potential disadvantage of allowing direct imitations from sensory memory that bypass actual linguistic abilities. For this reason, delayed imitation is argued by some to be the best possible elicitation technique (e.g. Piske et al. 2001; Flege 1995; Flege et al. 1999). Delayed imitation tasks typically take the form of a presentation of a question, an answer and then the repeating of the question, after which the listener/speaker is required to reproduce, i.e. imitate, the given answer. The procedure thus has somewhat greater demands on memory capacity than direct imitation does.

Figures 5 and 6 show examples that reflect success or the lack thereof with respect to the prosodic aspects of an imitation, be it direct or delayed.

Figure 5 depicts an imitation where the speaker succeeded (both acoustically and visually) in reproducing the incredulous question “Wo wohnt deine Oma?” (Where does your grandma live?), which manifests the unusual contour shape of a steady rise from the start to the end. The intonation shape is triggered by the unusual preceding context “Meine Oma muss sich immer vor den Gorillas in Acht nehmen” (My grandma must always beware of the gorillas).

Figure 6, on the other hand, shows a failed imitation of the half-irritated, half-exhausted exclamation “Not again” in answer to the information that “The neighbors are having another party”. The imitator (bottom contour) manages to reproduce the low back vowel /ɑ/ in “not”, but has problems with the intonation contour. The second F0 peak
Figure 5. Successful delayed imitation of the incredulous question “Wo wohnt deine Oma?” (Where does your grandma live?) – top contour: original, bottom contour: imitation

Figure 6. Failed delayed imitation of the exclamation “Not again!” – top contour: original, bottom contour: imitation
(on “aGAIN”), especially, occurs too late and is not high enough, and the anticipation of the first peak (on “NOT”) is also not reflected in this German speaker’s version. For this reason, she produces a steeper rise toward it (which may be reflected by an L+H* label instead of just H* as in the original).

7. Evaluation methods

7.1. Overview

This extensive experiment uses many different task types and creates a virtual deluge of data. Accordingly, results also manifest themselves in multiple forms, thus requiring an elaborate effort with respect to evaluation and analysis.

For pure production tasks involving reading, narration and acted speech, only a subjective perceptual assessment by a sufficiently large number of raters is feasible. For those tasks with more narrowly defined objectives, such as accent imitation and interpretative reading, it is somewhat easier to perform an expert analysis by referring to an expected model. This applies to an even larger degree to imitation tasks where instrumental analysis can support the comparison with the respective originals. In perception tasks, on the other hand, evaluation is achieved through an automatic scoring system which allows only a limited number of answers. In this case the main analytical power lies in the construction of the task itself, as well as in the selection and editing of the test stimuli.

As a matter of fact, it is a major challenge to the project as a whole to devise tasks in such a way that results actually reflect insights with respect to the targeted phenomena. This can be relatively straightforward for concrete abilities like reading, but becomes more complex for less well-defined areas where it is more difficult to exclude additional contributing factors. The main problems include the distinction between prosodic and segmental abilities, production and perception, as well as, of course, the differentiation between talent and mere proficiency. To some degree, the efficiency, i.e. the potential for insight offered by the tasks, but also the individual stimuli, cannot be known beforehand, but may only become clear in the analysis of the data. Nevertheless, a general knowledge of the contribution of particular influences to a certain task should allow for a sufficiently precise targeting of a certain object of investigation.
This problem of the relationship between task planning and evaluation and the actual nature and significance of results is addressed in more detail in the conclusion of this section. It is preceded by an account of the evaluation procedure with respect to the concrete distinction of proficiency and talent, as well as a description of the assessment methods, namely perceptual evaluation by raters, expert analysis and automatic scoring.

7.2. Talent vs. proficiency revisited

As talent is just one of several factors contributing to a speaker’s actually perceptible performance, achieving a distinction between superficial proficiency and inherent talent is far from being trivial.

It is possible to use the information gathered in the questionnaires addressing biographical data (see section 3) and personality factors (see section 4.1.) to better assess the influence of experience and practice (e.g. acquisition history and extent of language use today), as well as the degree of motivation. While there can be no straightforward quantification of the relation between this type of information and actual performance, it does allow a measured consideration of differences, for example between students of English and test subjects with little or no academic experience, or speakers with the long-term experience of living in an English-speaking country and others who only spent a two-week vacation in that environment.

It is certainly a more promising approach to isolating “talent”, when tasks are designed in such a way that the contribution of the confounding factors of experience and practice is neutralized as much as possible. These two factors may also play a diminished role with respect to intricate phonetic characteristics that are generally beyond the threshold of the speaker’s conscious awareness. The detailed investigations by Anufryk, Baumotte and Lewandowski described in the subsequent chapters of this volume deal with such phenomena, namely variation in intonation, coarticulatory resistance, and convergence, respectively.

As native speakers of German, the test participants should have very similar, if not equal, amounts of experience in that language such that this factor can reasonably be expected to play a very reduced role in the tasks involving German (interpretative reading, pair comparison, prosodic interpretation, imitation). In the Hindi imitation task, on the other hand, none of the speakers had any experience at all. As none of them were professional linguists, major differences in theoretical knowledge with respect to phonetics are also extremely improbable.
In the English tasks, the possible effect of practice and experience can, of course, not be excluded even though it is expected to be smaller in tasks focusing on prosody, as learners are less likely to have acquired any conscious or even formal knowledge about this aspect of speech.

7.3. Perceptual evaluation

Within the context of such a large examination of phonetic ability an overall assessment of the general quality of performance in the investigated L2, i.e. English (and to a much smaller extent Hindi), should of course have top priority.

A basic perceptual evaluation is necessary in order to obtain a reliable score of the test subjects’ abilities in relation to each other and also in absolute terms in order to determine how good the good speakers actually are in comparison with native speakers.

In order to reach a large number of native English-speaking raters (of several accents of English) and of native speakers of Hindi (for the assessment of the Hindi imitation task) a web-based evaluation design is used. The English evaluation focuses on two production tasks, namely the narration of the Gary Larson cartoon and the reading of “The North Wind and the Sun”. 117 speakers (102 German test subjects + 15 native speakers of English with various accents) must be rated. The analysis scheme for the actual assessment calls for the use of 200 raters overall. The raters are divided into four groups and each group rates 60 speakers with an overlap of 30 speakers per group so that each speaker will be rated by 100 raters. The stimuli are presented in a random order, and the two tasks (cartoon narration and reading) are completely separate. See Figure 7 for an example of a web page prepared for the assessment of one speaker’s narration of the cartoon.

Following the recommendation of Piske et al. (2001), a representative sample of raters is used, although there may not be a significant difference between expert and naïve raters (e.g. Bongaerts et al. 1995 and Bongaerts 1999). The stimuli are rated on a visual analogue scale (Wewers & Lowe 1990), which consists simply in a straight line either with no labels at all or, as in the scale used in our study, with only the two extreme boundaries labeled. The scale is unipolar, denoting the absence of a phenomenon, in this case foreign accent, at one end and its maximum intensity at the other end. Any click on the scale in between these extremes is assigned to a corresponding numerical value between 0 and 10. The fact that there are no further marks on the scale may allow the rater less possi-
Assessment of phonetic ability 37

Abilities of orientation but, on the other hand, leads to a more uniform distribution of scores along the scale’s entire length (Scott & Huskisson 1976). Generally, visual analogue scales are claimed to be easily understood, quickly scored and able to reflect fine discriminations.

A comparison with the ratings for the recordings of the 15 native speakers of English also distributed throughout the stimulus presentation also allows an assessment of the speakers’ performance in absolute terms, the generally accepted criterion for determining near-nativeness being mean ratings within two standard deviations of the mean ratings for native speakers (e.g. Flege 1995; Bongaerts et al. 1995; Piske et al. 2001).

7.4. Expert analysis

The assessments obtained by raters in the perceptual evaluation just described reflect an overall impression shaped by the more or less conscious awareness of individual phenomena in the speakers’ productions. However, it is also essential to conduct a more specific investigation of pro-

Figure 7. Example of the web-based evaluation (here: rating of performance in narration of cartoon)
duction performance that helps understand which individual linguistic criteria have contributed to this impression.

For this reason, three expert raters perform a very thorough analysis and evaluation of every utterance produced by every single test participant. Specific evaluation schemes have been developed for every task. Individual criteria are defined and weighted. In the tasks using quasi-s spontaneous speech, for example, aspects such as fluency (20%), vocabulary (20%) and grammar (20%) are taken into consideration besides mere pronunciation skills (40%). In other task types corresponding criteria are deleted or added, e.g. maintaining consistency (i.e. avoiding slips into normal German pronunciation) in the accent mimicry task.

A more focused type of analysis is performed in those tasks in which the speaker’s production consists of only one utterance, i.e. interpretative reading and all imitations, as such tasks are designed to target specific segmental and/or prosodic phenomena. Evidently, the subsequent analysis concentrates on the realization of these selected phenomena. The assessment of segmental production aims to cover the pronunciation problems affecting individual speech sounds (see section 6.1.1.) of English. For prosody, on the other hand, there are two different kinds of challenges. In imitation tasks the speaker’s production is compared to a given model, i.e. the goal is to achieve an exact phonetic realization of a specific tonal event. In the reading tasks, however, the challenge is to choose tonal categories that are appropriate in a certain discourse situation or when trying to express a particular emotion. A good performance reproducing dialogs, stories or individual utterances therefore also reflects a general awareness and understanding of the relationship between sentence meaning and prosody.

An expert analysis is also performed for the Hindi imitation task in addition to the native speaker judgments in the web-based evaluation. Even though our “experts” are not native speakers of Hindi, a reasonably objective assessment is achieved via the use of weighted criteria starting with the basic requirement of reproducing the same number of syllables as in the original, the same phonemes (e.g. some kind of /r/-sound), the correct allophonic realization (e.g. an alveolar trilled /r/ and not a uvular one) and the correct stress assignment.

Within the context of a performance assessment, an instrumental analysis is not necessarily of primary importance, as, by definition, only perceptible phenomena are relevant. It is, of course, especially helpful in the examination/visualization of $F_0$ contours. More importantly, however, instrumental analysis is crucial in studies that investigate specific
phonetic phenomena assumed to reflect different talent levels (again, see the investigations by Anufryk, Baumotte and Lewandowski alluded to earlier).

7.5. Scoring of perception tests

The perception tests are designed in such a way that the listeners provide clearly defined results, either deciding whether a difference was perceived or not (pair comparison) or selecting an answer in a multiple choice scenario (interpretation, accent identification). The results can therefore be processed automatically and a general assessment of the performance is achieved in a relatively straightforward fashion on that basis. As results in the form of “yes/no” or “1, 2, 3 or 4” do not provide much information about the listeners’ actual abilities, the main burden in this type of task must of course be shouldered by the construction and selection of the individual tasks, which must be designed in such a way that only one strictly defined problem is addressed at a time. In a pair comparison, i.e. a discrimination task, this means, for example, that the recordings to be compared must be completely identical with the exception of the targeted phenomenon (see section 6.2. for a more detailed description of the actual tasks). By means of such a concise definition of the respective area of interest, the perception of prosodic phenomena ranging from paralinguistic information pertaining to emotions (e.g. sadness, anger, sarcasm) to major discourse elements related to specific tonal constellations (e.g. incredulous questions, continuation rises, etc.) to tonal categories and subtle realizational differences can be tested effectively. Additionally, the individual stimuli that make up a particular task are weighted according to their degree of difficulty, and in multiple choice questions, wrong answers may be evaluated differently depending on how plausible they are in comparison to the correct answer. The example of a stimulus in the accent identification task (non-native speakers of German read a short story and listeners must identify the speakers’ country of origin) may demonstrate this. Assuming that, in a particular example, the answer “Italy” were correct, the answer “Spain” would be rated better than the answer “Japan”, as a Spanish accent is assumed to be more similar to an Italian accent than a Japanese one.
7.6. Overall analysis:
looking for interactions, correspondences and interpretations

102 test candidates participated in the multitude of different tasks described above, producing large quantities of data to be analyzed and assessed.

All production tasks and the individual stimuli contained within them receive ratings on a scale from 0 to 10. However, as already mentioned, some tasks and individual stimuli are weighted in order to take into account their presumed greater or lesser significance. The perception tasks including the individual answers in the multiple choice scenario are also weighted and processed to yield results on a 0 to 10 scale.

The appendix to this chapter includes the example of a scorecard of the expert ratings for a single test subject. It lists all the individual scores and the major criteria for each of the tasks described above. The appendix also contains an overview of the ratings for all 102 test participants.

The strategy of assigning weights to individual stimuli before the actual analysis of the results is of course equivalent to a prediction of their likely significance. It is not at all certain which of the tasks will actually turn out to be especially useful in evaluating phonetic talent. The detection of interesting correlations and interactions between task types (e.g. production vs. perception, prosody vs. segmental, etc.) or stimuli happens, for the most part, in the post-analysis of the given data. This is also the case with respect to relating the phonetic tasks to the results of the psychological tests and personal data collected in the questionnaire.

Due to the automatic processing of scores in the perception tasks and some of the psychological tests, a quicker preliminary analysis of these data is possible. Early results show, for example, a relationship between better performances in the interpretation tasks and a higher empathy score. Performance in the German interpretation tasks correlates especially strongly with accent identification, possibly indicating a heightened awareness of the more subtle characteristics of the native language. Perhaps not surprisingly, performances in all the pair comparisons (German, English, low-pass filtered) correlate with each other, which may mean that linguistic knowledge/awareness does not play as large a role as purely perceptual abilities. This view may be supported by the fact that better performance in the low-pass filter-tasks correlates with the personality factors of less inhibition and more conscientiousness, which could be interpreted as an expression of greater effort and motivation during the test. These results provide a first impression of the entire data and the possible interdepend-
encies between all the different task types. They also indicate to what extent the tasks actually fulfill the intended functions.

The challenge of distinguishing the concepts of talent and proficiency is of course the most important issue of the analysis. The results of both those tasks designed to focus on talent and those simply measuring general proficiency in a particular area must also be related to information about the speakers’ experience and degree of motivation as determined in the questionnaires and the psychological tests. While the factor of experience can be controlled to a certain extent (see the discussion in section 7.2), this is not possible for motivational factors. Motivation is important in two different ways. Primarily, it plays an important long-term role during actual language acquisition, but it also affects the execution of the experiment itself by influencing the effort and commitment made in the immediate performance of a test. There is also the question of whether at least some aspects of motivation may be inherent to personality and thus not completely separable from talent anyway.

In summary, it must therefore be stated that there is no experimental method that directly assesses exclusively phonetic talent, but that it must be approximated via the combination of many different tests. A comprehensive analysis of all the data generated is important for the objective of obtaining general insights into the nature of speech talent and its sub-components. It also provides the best possible ranking/classification of the speakers necessary for the subsequent neuroimaging studies of talent and proficiency. Finally, the assessment of the experiment as a whole and of all the individual tasks and stimuli is also important. Once this is achieved, the most informative task types and stimuli are determined and a more compact and efficient test could be developed to routinely identify/assess phonetic talent in native speakers of German.

References

Allport, G. W.

Baddeley, A.

Barry, W.
Bongaerts, T., Planken, B. & Schils, E.

Bongaerts, T.

Carroll, J. & Sapon, S.

Carver, C. S. & White, T.L.

Costa, P. T. & McCrae, R. R.
1992 *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory. Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Derwing, T., Munro, M. & Wiebe, G.

Dogil, G.

Féry, C.
1993 *German Intonational Patterns*. Tübingen: Max Niemeyer Verlag

Flege, J. & Hammond, R.

Flege, J. & Hillenbrand, J.

Flege, J.

Flege, J., Yeni-Komshian, G. & Liu, H.
Fraser, H.

Gathercole, S. E., Willis, C. S., Baddeley, A. D. & Emslie, H.

Gordon, E.

Grice, M. & Baumann, S.

Ioup, G. & Tansomboon, A.

Jilka, M., Möhler, G. & Dogil, G.

Jilka, M.

Kuhl, P. & Meltzoff, A.

Larson, G.

Lehrl, S.

Leibetseder, M., Lairreiter, A. R., Riepler, A. & Köller, T.

Markham, D.

Mayer, J.

Mennen, I.
Möhler, G.

Moyer, A.

Neufeld, G.

Oyama, S.

Pierrehumbert, J.

Piske, T., MacKay, I. & Flege, J.

Ramus, F. & Mehler, J.

Raven, J., Raven, J. C. & Court, J. H.

Riester, A.
2008 The components of focus and their use in annotating information structure. PhD Dissertation, Universität Stuttgart

Riester, A.

Scott, J. & Huskisson, E. C.


Simon, J. R.

’t Hart, J., Collier, R. & Collier, R.
Twain, M.

Weismer, G. & Cariski, D.

Wewers, M. E. & Lowe, N. K.

Appendix

1. Scorecard

(Transcriptions are in SAM-PA)

**Production**

**Pro1** Introduction/Dialogue

Speaker (Number): 86

Overall score 0–10: 7.8

general impressions of

- fluency (20%) 8.5
- pronunciation (40%) 7
- vocabulary (20%) 8.5
- grammar (20%) 8

**Pro2** Cow Cartoon (*narration*)

Speaker (Number): 86

Overall score 0–10: 6.9

- fluency (20%) 7
- pronunciation (40%) 7
- vocabulary (20%) 6.5
- grammar (20%) 7
Pro3  The Northwind and the Sun *(reading)*

Speaker (Number): 86

Overall score 0–10: 7.1

<table>
<thead>
<tr>
<th></th>
<th>(10%)</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>fluency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pronunciation</td>
<td>(90%)</td>
<td>7</td>
</tr>
</tbody>
</table>

pay special attention to production of:
- continuation rises at “stronger”, “blew”, “warmly”
- pitch accents at “first succeeded”, “at last”, “and so”, “stronger (of the two)”
- phrasing (breaks) between “along – wrapped”, “cloak off – should”, “he blew – the more”, “warmly – and”

Pro4  Der Nordwind und die Sonne 
*(reading in German with faked English accent)*

Speaker (Number): 86

Overall score 0–10: 6.5

<table>
<thead>
<tr>
<th></th>
<th>(20%)</th>
<th>6.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>phrasing/intonation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>German outliers</td>
<td>(20%)</td>
<td>6.5</td>
</tr>
<tr>
<td>consistency</td>
<td>(10%)</td>
<td>6.5</td>
</tr>
<tr>
<td>segmental features</td>
<td>(50%)</td>
<td>6.5</td>
</tr>
</tbody>
</table>

pay special attention to production of:
- general vowel coloring
- pronunciation of /r/ as postalveolar (or lack of syllable-final /r/ in British English imitation)
- diphthongization of long monophthongs (esp. /e:/ and /o:/)
- initial /z/ pronounced /s/
- mispronunciation of [x] as [k] and [C] as [S] or [k]
- grapheme <w> as /w/
- rounded front vowel pronounced back, e.g. /Y/ as /U/
- affricate /ts/ (grapheme <z> pronounced /z/
- lower pronunciation of /@/
- graphemes <st> pronounced as /st/ when syllable-initial
- overall intonation, especially final tunes

Pro5 Mrs. McWilliams *(reading)*

Speaker (Number): 86

Overall score 0–10: 6.6

<table>
<thead>
<tr>
<th>fluency</th>
<th>(10%)</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>pronunciation</td>
<td>(90%)</td>
<td>6.5</td>
</tr>
</tbody>
</table>

pay special attention to production of:
- intonational patterns in “It is mostly confined to women”,
  “the devil himself – or a mouse –”,
- exclamation “Mortimer, Mortimer”
- Y/N-Question “Evangeline, is that you calling?”
- Wh-Questions “What is the matter? Where are you?”
- phrasing structure in “You never try, Mortimer – you know very well you never try.”
- general dialogue intonation performance
- segmental pronunciation of “infirmities”, “dog”,
  “flash of lightning”, “smothered and unlocatable”,
  “Mortimer”, “scrape my faculties”, “Evangeline”,
  “closet”, “awful storm”, “unreasonable, “muffled sobs”, “deliberately”
- fluency in phrases: “I should think you would take some little care of your life, for my sake and the children’s, if you will not for your own.”, “You know there is no place so dangerous as a bed in such a thunderstorm as this”, “Yet there you would lie and deliberately throw away your life!”
Pro6 Individual Sentence Accent Fake (reading)

Speaker (Number): 86

Overall score 0–10: 5.3

general impression English 7
general impression French 4.5
general impression Italian 4.5

major criteria:
- identify speakers who have no idea what to do at all
- attempt to reproduce French prosody
- production of any recognizable Italian features
- general impression

------------------------------------------------------------------------------------------

Pro7 Individual Sentence Reading

Speaker (Number): 86

Overall score 0–10: 6.6

Pro7_a question tag intonation, “hurry”, palatalization in “weren’t you” 6
Pro7_b “available”, palatalization in “that you”, not too monotonous intonation 8
Pro7_c “hood”, “jeep”, “hot”, reduction in “was” 6
Pro7_d “desires”, “informative”, “subject”, phrasing 7
Pro7_e overall intonation, “feather”, “happy” 5.5
Pro7_f stress on “so” 7

------------------------------------------------------------------------------------------
**Pro8**  **Interpretation Deutsch (reading)**

Speaker (Number): 86  
Overall score 0–10: 8.0

Pro8.a  H*L on “PoLiTiker”, H*(L) on “korr upt”  
(variations allowed)  10  
Pro8.b  hat pattern: L*H on “Alle”, H*L on “NICHT”  1  
Pro8.c  “Japanisch”-phrase – Ausdruck: keine Lust  8.5  
Pro8.d  “Japanisch”-phrase – Ausdruck: Schadenfreude  8.5  
Pro8.e  “Japanisch”-phrase – Ausdruck: Freude  8  
Pro8.f  “Japanisch”-phrase – Ausdruck: Überraschung  9.5  
Pro8.g  “Japanisch”-phrase – Ausdruck: Ärger  9  
Pro8.h  “Japanisch”-phrase – Ausdruck: bizarre Idee  9  
Pro8.i  “Japanisch”-phrase – Ausdruck: Kontrast (nicht Chinesisch)  9  
Pro8.j  “Japanisch”-phrase – Ausdruck: Versicherung, Bestätigung  7.5  

---

**Pro9**  **Interpretation English (reading)**

Speaker (Number): 86  
Overall score 0–10: 6.6

Pro9.a  intonation wh-question, stress on “score” not “we”,  
intonation of “I don’t care”  7.5  
Pro9.b  exclamation/calling contour (high plateau on “lunch”)  7.5  
Pro9.c  “teething”-phrase – express surprise (voiced /D/!!)  8  
Pro9.d  “teething”-phrase – express repetition  5.5  
Pro9.e  “teething”-phrase – express happiness  6  
Pro9.f  “teething”-phrase – express pity  7  
Pro9.g  “teething”-phrase – express anger  5  
Pro9.h  contrastive pitch accents on “shrimp” and “lobster”  5.5  
Pro9.i  contrastive pitch accents on “apple juice”  
and “orange juice”  7
Imitation

**Imi1**  Unmittelbare Imitation Deutsch

Speaker (Number): 86

Overall score 0–10: 8.5

- Imi1_a haferflo_HLH: fall-rise on “Haferflocken” 8.5
- Imi1_b haferflo_LHH: straight rise from very low accent on “Haferflocken” 8.5
- Imi1_c express satisfaction about going to Brazil and the fact that the listener will not 8
- Imi1_d express surprise at statement which is deemed to be unrealistic and bizarre 8.5
- Imi1_e confirming the fact to someone who at first did not believe it 9
- Imi1_f express lack of enthusiasm about going to Brazil 9
- Imi1_g hat pattern from “Benjamin” (L*H) to “gleich” (H*L) 9
- Imi1_h shorter/more extreme hat pattern on “wünscht” & “gleich” (F₀ still high, fall later than in Imi_7) 8
- Imi1_i rise on “wünscht”, H* on “angeschafft”, then declarative fall 8

**Imi2**  Direct Imitation English

Speaker (Number): 86

Overall score 0–10: 6.9

- Imi2_a (RP): L* on “go”, rise on “MANchester”, diphthong in “go”, lowered /a/ in “Manchester” 4.5
- Imi2_b (GA): rise on “looks”, weak rise-fall on “watch” 8
- Imi2_c (RP): H* on “looks” and “watch” (creak!), vowel in “watch” 8
- Imi2_d (GA): same as Imi2_2, but with “watch” shortened (thus rise-fall only a truncated weak rise) 7
| Imi2_e  | (RP): H* on “supPOSE”, L* on “thing”; H* on “Later”, diphthong in “suppose”, “to” with schwa, strong /t/ in later, no rhoticity | 7.5 |
| Imi2_f  | (GA): H* on “I”, L* L-H% on “good”, tap at end of “thought” | 7.5 |
| Imi2_g  | (RP): high plateau, stress on “Evening”, followed by late, weakened fall, diphthong on “going” | 5 |
| Imi2_h  | (GA): H* on “most”, then fall, rise on “INmates”, low ip- boundary, rise on VIolent”, /A/ on “not” | 4 |
| Imi2_i  | (GA): additional L+H* on “crimes” | 6.5 |
| Imi2_j  | (GA): “violent” shortened, leads to compressed rise and fall toward IP boundary | 8 |
| Imi2_k  | (RP): H* on “wouldn’t”, “DIfficult”, late rise on “out”, strong /t/ in “out” | 6.5 |
| Imi2_l  | (GA): high start on “two”, L* on “BOston” + H-ip-boundary, H* on “one” and “town”, tap on “out”, /A/ on “Boston”, strongly reduced vowel on “are” | 7 |
| Imi2_m  | (GA): additional rise on “out” | 6 |
| Imi2_n  | (GA): “town” lengthened, leads to longer, less steep fall | 7 |
| Imi2_o  | (RP): H* on “there’s”, “dog”, “GARden” (L+H*), “aGAIN” (with short fall), r-lessness in “there’s”, “garden”, aspirated /t/ in “little”, /a/ in “black”, /Q/ in “dog” | 7 |
| Imi2_p  | wellHL (RP): final fall | 7.5 |
| Imi2_q  | wellHLH (RP): fall-rise | 6.5 |
| Imi2_r  | wellH (RP): high level | 8 |
| Imi2_s  | wellLHL (RP): rise-fall | 7.5 |
| Imi2_t  | wellLH (RP): final rise | 9 |
Imi3  Verzögerte Imitation Deutsch

Speaker (Number): 86

Overall score 0–10: 8.1

Imi3_a  comic: contrastive focus on “Thomas” und “Markus” (L*H), also on “gelesen” and “zerrissen” 7.5
Imi3_b  geschlafen: old information focus on “geschlafen” also signalled by syntactic position (“remnant movement”): low in “geSCHLAFen”+ rise to “hat” (L*H), fall (H*L) on “keiner” 8.5
Imi3_c  lang: continuation rise on “Lang”; compressed pitch range in subordinate clause, also with continuation rise on “Regisseure” (+ L*H on “andere”); hat pattern from “Mensch” (L*H) to “Ekel” (H*L) 7
Imi3_d  marianne: calling contours (H*M %) on “Marianne” and “Telefon” 9
Imi3_e  incredulous question (rise despite Wh-question), L*H on “wo”, high level at “wohnt”, final rise on “Oma” 9
Imi3_f  petersburg: L*H on “DIESmal”, high level, high rise on “PEtersburg” (contrast focus) 8.5
Imi3_g  wehe: admonition, L*H on “wehe”, then fall to end (!H* on “kaPUTT”) 9
Imi3_h  fenster: Yes/No-question as Indirect Speech Act, unusual phrase break (L-) after “das” (also lengthened), then steep L*H on “FENster” 8
Imi3_i  ernst: expressing indignation, small rise on “ist”, fall (!H*L) on “Ernst” 7
Imi3_j  durcheinander: contrastive focus on “schwarze RÖcke” (rise-fall L*HL), L* on “Alles”, rise on “bringst” 7.5
Imi4  Delayed Imitation English

Speaker (Number): 86

Overall score 0–10: 6.9

Imi4_a (GA): H* L+H* L-L% – low /A/ and tap on “not” 7
Imi4_b (RP): H* on “Hi”, L*+H on “Tony”, high ip-boundary, high level with fall from “DOing”, /@U/ and final /I/ on “Tony”, reduced “are” 6.5
Imi4_c (GA): Three IPs; H* L-L% on “sure”; L+H* L-L% on “good”; Wh-question with L+H* on “in”; “copy with /A/, palatalization in “don’t you” 6.5
Imi4_d (RP): rise-fall on “was” (L+H* L-), H* on “tell”, reduced vowel in “us”, full vowel in “was” 7.5
Imi4_e (GA): H* on “used”, weak rise-fall (L+!H* L-L%) on “now”, diphthongs /aI/ and /aU/ 9
Imi4_f (RP): L+H* in “ESPeically” and “GRAduating”, L+H* L-L% on “MAY”; /a/ and palatalization in “graduating”, “we’ll be” reduced to /wIbI/ 3.5
Imi4_g (GA): L+H* on “FINally” stays high for remainder of word; “did it” with tap and unreleased /t/ 7.5
Imi4_h (RP): hat pattern from “seems” to “pro-”, no rhoticity, aspiration throughout “tutor” 8.5
Imi4_i (GA): incredulous question; L+H* L-L%; breaking 6.5
Imi4_j (GA): fall on “really” (L+H* L-) with slight uptick not labelled; L+H* rises on “sign”, “door” and “think” leading to up-down-up-down-up movement; /A/ on “not” 6.5
Imi4_k (RP): H* on “not”, fall-rise on “aGAIN”, /Q/ and strong /t/ in “not” 6.5
**Imi5   Direct Imitation Hindi**

Speaker (Number): 86

Overall score 0–10: 7.7

Hierarchy of criteria: Level 1: number of syllables (4 points)
Level 2: stress assignment (2 points)
Level 3: phonemes (2 points)
Level 4: phonetic realization (2 points)

Imi5_a ['rastrUp@ti] 8.5
Imi5_b [dr'trA:st] 8.5
Imi5_c [,anUsan’thA:n,gEndr] 8.5
Imi5_d ['vIsUvI,dallE] 9
Imi5_e ['mE: par’dA:n@m@n,tri: ‘kabbanUN,gi:] 6.5
Imi5_f ['mE: e: ‘ktsIkI,tsIkI pas’ge:a] 4
Imi5_g ['me:rA abna’de:s a’tSahE] 8.5
Imi5_h ['Wo: ‘zIlmE ‘tErhaHE] 8

(transcriptions are approximate)

**Perception**

**Per1   Paarvergleich**

Speaker (Number): 86

Overall score 0–10: 13 (16) = 8.1

i = files are identical; r = realizational difference;
k = categorical difference

Wer hat dieses Essen gekocht?
Per1_a  k   essen_fall_216 – essen_fall_216_mod2 (additional H*L on “gekocht”) (1) 1
Per1_b  i   essen_fall_216 – essen_fall_216 (1) 1
Per1_c  r   essen_risefall – essen_risefall_mod6 (peak on “essen” later) (3) 0
| Per1_d | k | leo_Hut1 – leo_Hut3 | (1) | 1 |
| Per1_e | i | leo_Hut1 – leo_Hut1 | (1) | 1 |
| Per1_f | k | brasilien_freude1 – brasilien_aerger | (1) | 1 |
| Per1_g | i | brasilien_stimmt2 – brasilien_stimmt2 | (1) | 1 |
| Per1_h | r | brasilien_stimmt2 – brasilien_stimmt3 | (3) | 3 |

| Per1_i | i | essen_fall_216 – essen_fall_216 | (1) | 1 |
| Per1_j | r | brasilien_stimmt2 – brasilien_stimmt4 | (3) | 3 |

### Per2 Pair Comparison

**Speaker (Number):** 86

**Overall score 0–10:** 10 (25) = 4.0

- **i =** files are identical; **r =** realizational difference; **c =** categorical difference

**Would you like some cream?**

Per2_a | r | cream – cream_mod3 | (“cream” louder) | (0) | 0 |

**People will get used to the idea**

Per2_b | c | idea – idea_mod2 | (L* on “used” instead of H*) | (2) | 2 |

**They’re in jail for such things as bad checks or stealing**

Per2_c | i | badchecks – badchecks | (1) | 1 |

**People will get used to the idea**

Per2_d | r | ideaBRNC – ideaBRNC_mod2 | (early peak on “idea”, subtle) | (4) | 0 |
Most county jail inmates did not commit violent crimes

Would you like some cream?

They’re in jail for such things as bad checks or stealing

It looks like a watch

There’s a spoon in here

Once upon a time there were three bears

It looks like a watch

Haven’t you asked the boss for more?
**Per3  Pair Comparison ToBI**

Speaker (Number): 86

Overall score 0–10: 7 (14) = 5.0

\[ i = \text{files are identical}; \ r = \text{realizational difference}; \ c = \text{categorical difference} \]

Marianna made the marmalade

| Per3_a | c | made144_pt3 – made144_pt2 | (1) 1 |
| Per3_b | i | made144_pt1 – made144_pt1 | (1) 1 |
| Per3_c | r | made344_pt3 – made444_pt1 | (4) 0 |
| Per3_d | c | made144_pt1 – made144_pt2 | (1) 1 |
| Per3_e | r | made344_pt3 – made444_pt2 | (3) 0 |
| Per3_f | c | made244_pt2 – made244_pt3 | (1) 1 |
| Per3_g | i | made144_pt3 – made144_pt3 | (1) 1 |
| Per3_h | r | made444_pt1 – made444_pt2 | (2) 2 |

**Per4  Interpretation Deutsch**

Speaker (Number): 86

Overall score 0–10: 6 (11) = 5.5

| Per4_a | bedauert1 vs. bedauernicht1 | Answer: bedauernicht1 / a. | (1) 0 |
| Per4_b | euro_vergesslich2 | Answer: a. | (2) 0 |
| Per4_c | euro_viel1 | Answer: c. | (2) 2 |
| Per4_d | euro_vergewissern2 | Answer: b. (accept d. for 1p.) | (2) 1 |
| Per4_e | euro_glaube1 | Answer: d. (accept b for 2p.) | (3) 3 |
| Per4_f | pille_nicht2 vs. pille2 | Answer: pille2 / b. | (1) 0 |
**Per5  Interpretation English**

Speaker (Number): 86

Overall score 0–10: 12 (14) = 8.6

Per5_a  good2
    Answer: d. (1) 0

Per5_b  spoon2
    Answer: a. (2) 2

Per5_c  won1 – won2
    Answer: won1 / a. (1) 0

Per5_d  bloomingdales1 – bloomingdales2
    Answer: bloomingdales1 / a. (3) 3

Per5_e  marianna1 – marianna2
    Answer: marianna1 / a. (1) 1

Per5_f  name
    Answer: c. (2) 2

Per5_g  oughtto1 – oughtto2
    Answer: oughtto1 (a.) (1) 1

Per5_h  alright
    Answer: b. (3) 3

------------------------------------------------------------------------------------------

**Per6  Low Pass Paarvergleich Deutsch**

Speaker (Number): 86

Overall score 0–10: 10 (18) = 5.6

Kanzler Schröder hatte drei Köche und einen Gärtner

6_a  i  kanzler_drei – kanzler_drei (1) 0

Wer hat dieses Essen gekocht?

6_b  k  essen_fall_216 – essen_fall_216_mod4
    (L*HL on “essen”) (1) 1

Per6_c  r  essen_risefall – essen_risefall_mod8
    (peak on “essen” much later) (1) 1

Per6_d  k  essen_risfallris – essen_risfallris_mod6
    (no accent on “essen”) (1) 1

Per6_e  i  essen_fall16 – essen_fall16 (1) 1
Sie werden von der regulären Armee Kroatiens unterstützt
6_f r kroatien – kroatien_mod5
("regulären" lengthened) (3) 3

Kanzler Schröder hatte drei Köche und einen Gärtner
6_g k kanzler_drei – kanzler_drei_mod5
(peak on “Gärtner”) (1) 0
Per6_h i kanzler_drei – kanzler_drei (1) 1
Per6_i r kanzler_drei – kanzler_drei_mod1
(peak on “drei” earlier) (3) 0

Die Konvergenzkriterien müssten unbedingt eingehalten werden
6_j k konvergenz – konvergenz_mod2
(L*H on “eingehalten”) (3) 0

Wer hat dieses Essen gekocht?
Per6_k r essen_fall16 – essen_fall16_mod3
(question-like rise at the end) (1) 1
Per6_l i essen_risfallris – essen_risfallris (1) 1

Per7 Low Pass Pair Comparison

Speaker (Number): 86

Overall score 0–10: 7 (18) = 3.9

The State is not planning on putting more police on the road
Per7_a r road – road_mod3
(higher and later peak on “planning”) (3) 0

It looks like a watch
Per7_b i watch – watch (1) 0

I thought it was good
Per7_c c good2 – good2_mod6
(L+H* L-L% as final tune) (1) 1
People will get used to the idea

Per7_d  c  idea – idea_mod3
          (L+H* on “idea”)
(1)  1

Per7_e  c  ideaBRNC – ideaBRNC_mod4
          (final rise)
(2)  0

I thought it was good

Per7_f  i  good2 – good2
(1)  1

Most county jail inmates did not commit violent crimes

Per7_g  c  inmates – inmates_mod1
          (L+H* on “crimes”)
(1)  0

The State is not planning on putting more police on the road

Per7_h  i  road – road
(1)  1

There’s a spoon in here

Per7_i  r  spoon2 – spoon_mod1
          (final !H-L% as H-L% – higher plateau)
(2)  2

That’s what I thought

Per7_j  c  thought80 – thought80_mod1
          (L* on “thought” instead of H*)
(1)  1

I’d like to see the cheapest flight from Atlanta to Baltimore

Per7_k  r  cheapest80 -cheapest80_mod1
          (final fall on Baltimore higher on first syllable)
(2)  0

It looks like a watch

Per7_l  c  watch – watch_mod8
          (Y/N question)
(2)  0

--------------------------------------------------------------------------
Per8  Low Pass Language Identification

Speaker (Number): 86

Overall score 0–10: 4.0

(correct answer in capital letters, given answer underlined)

<table>
<thead>
<tr>
<th>Per8_a</th>
<th>Choices: a. German</th>
<th>b. ENGLISH</th>
<th>c. Japanese</th>
<th>d. Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per8_b</td>
<td>Choices: a. JAPANESE</td>
<td>b. German</td>
<td>c. Arabic</td>
<td>d. French</td>
</tr>
<tr>
<td>Per8_c</td>
<td>Choices: a. French</td>
<td>b. English</td>
<td>c. GERMAN</td>
<td>d. Arabic</td>
</tr>
</tbody>
</table>

Per9  Accent Identification

Speaker (Number): 86

Overall score 0–10: 9.5 (20) = 4.8

(correct answer in capital letters, given answer underlined)

<table>
<thead>
<tr>
<th>Per9_a</th>
<th>Choices: a. Italian</th>
<th>b. English</th>
<th>c. Russian</th>
<th>d. FRENCH</th>
<th>score (1)</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per9_b</td>
<td>Choices: a. ITALIAN</td>
<td>b. Dutch</td>
<td>c. Japanese</td>
<td>d. Spanish</td>
<td>score (2)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Spanish 0.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per9_c</td>
<td>Choices: a. Spanish</td>
<td>b. Swedish</td>
<td>c. American</td>
<td>d. GREEK</td>
<td>score (3)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Spanish 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per9_d</td>
<td>Choices: a. Norwegian</td>
<td>b. AMERICAN</td>
<td>c. French</td>
<td>d. Finnish</td>
<td>score (2)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Norwegian 0.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per9_e</td>
<td>Choices: a. Greek</td>
<td>b. Turkish</td>
<td>c. ARABIC</td>
<td>d. Japanese</td>
<td>score (3)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Turkish 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per9_f</td>
<td>Choices: a. Swedish</td>
<td>b. Chinese</td>
<td>c. Italian</td>
<td>d. RUSSIAN</td>
<td>score (3)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Swedish 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per9_g</td>
<td>Choices: a. Austrian</td>
<td>b. Danish</td>
<td>c. DUTCH</td>
<td>d. Polish</td>
<td>score (1)</td>
<td>1</td>
</tr>
<tr>
<td>Per9_h</td>
<td>Choices: a. Dutch</td>
<td>b. American</td>
<td>c. ENGLISH</td>
<td>d. Swedish</td>
<td>score (2)</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(American 0.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per9_i</td>
<td>Choices: a. AUSTRIAN</td>
<td>b. Dutch</td>
<td>c. Swiss</td>
<td>d. Danish</td>
<td>score (1)</td>
<td>1</td>
</tr>
<tr>
<td>Per9_j</td>
<td>Choices: a. Russian</td>
<td>b. Portuguese</td>
<td>c. ICELANDIC</td>
<td>d. Hungarian</td>
<td>score (2)</td>
<td>0</td>
</tr>
</tbody>
</table>
Summary:

Pro1  Intro  7.8
Pro2  Cartoon  6.9
Pro3  Northwind  7.1
Pro4  Nordwind_Acc  6.5
Pro5  McWilliams  6.6
Pro6  Accent Fake  5.3
Pro7  Sentence Reading  6.6
Pro8  Inter_Read_D  8.0
Pro9  Inter_Read_E  6.6

Free Speech (Pro1,2): 7.4
Reading (Pro3,5,7,8,9): 7.0
Accent (Pro4,6): 5.9
Interpretation (Pro8,9): 7.3
English (Pro1,2,3,5,7,9): 6.9
German (Pro4,6,8): 6.6
Overall Production: 6.82

Imi1  Unm_D  8.5
Imi2  Dir_E  6.9
Imi3  Verz_D  8.1
Imi4  Del_E  6.9
Imi5  Hindi  7.7

Direct Imitation (Imi1,2,5): 7.7
Delayed Imitation (Imi3,4): 7.5
German (Imi1,3): 8.3
English (Imi2,4): 6.9
Overall Imitation: 7.62
<table>
<thead>
<tr>
<th>Per1</th>
<th>Paar_D</th>
<th>8.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per2</td>
<td>Pair_E</td>
<td>4.0</td>
</tr>
<tr>
<td>Per3</td>
<td>Pair_ToBI</td>
<td>5.0</td>
</tr>
<tr>
<td>Per4</td>
<td>Int_D</td>
<td>5.5</td>
</tr>
<tr>
<td>Per5</td>
<td>Int_E</td>
<td>8.6</td>
</tr>
<tr>
<td>Per6</td>
<td>LP_D</td>
<td>5.6</td>
</tr>
<tr>
<td>Per7</td>
<td>LP_E</td>
<td>3.9</td>
</tr>
<tr>
<td>Per8</td>
<td>LP_ID</td>
<td>4.0</td>
</tr>
<tr>
<td>Per9</td>
<td>Acc_ID</td>
<td>4.8</td>
</tr>
</tbody>
</table>

Pair reg speech (Per1,2,3): 5.7
Interpretation (Per4,5): 7.1
Pair LP (Per6,7): 4.8
Identification (Per8,9): 4.4
German (Per1,4,6,9): 6.0
English (Per2,3,5,7): 5.4
LP (Per6,7,8): 4.5
NO LP (Per1,2,3,4,5,9): 6.0
Overall Perception: 5.50

Overall: 6.478 (without Per8 LP_ID: 6.591)

English Competence (Pro1,2,3,5,7,9, Imi2,4, Per2,5) ⇒ 6.80
Talent/Awareness (Pro4,6,8, Imi3,4,5, Per1,6,7,9) ⇒ 6.49
2. Ratings overview

(missing speaker numbers were assigned to control native speakers of English)

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Overall</th>
<th>Competence</th>
<th>Talent</th>
<th>Production</th>
<th>Imitation</th>
<th>Perception</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.291</td>
<td>7.930</td>
<td>6.780</td>
<td>7.17</td>
<td>8.64</td>
<td>6.08</td>
</tr>
<tr>
<td>2</td>
<td>7.205</td>
<td>7.600</td>
<td>7.000</td>
<td>7.48</td>
<td>7.96</td>
<td>5.93</td>
</tr>
<tr>
<td>3</td>
<td>6.982</td>
<td>7.040</td>
<td>6.950</td>
<td>6.94</td>
<td>7.76</td>
<td>6.03</td>
</tr>
<tr>
<td>5</td>
<td>7.732</td>
<td>8.120</td>
<td>7.700</td>
<td>8.24</td>
<td>8.58</td>
<td>6.33</td>
</tr>
<tr>
<td>6</td>
<td>7.236</td>
<td>7.270</td>
<td>7.430</td>
<td>6.88</td>
<td>7.80</td>
<td>6.92</td>
</tr>
<tr>
<td>8</td>
<td>7.123</td>
<td>7.580</td>
<td>6.610</td>
<td>6.87</td>
<td>8.28</td>
<td>6.17</td>
</tr>
<tr>
<td>9</td>
<td>7.573</td>
<td>7.210</td>
<td>8.040</td>
<td>7.08</td>
<td>7.66</td>
<td>7.84</td>
</tr>
<tr>
<td>10</td>
<td>6.696</td>
<td>7.000</td>
<td>6.290</td>
<td>7.00</td>
<td>7.68</td>
<td>5.54</td>
</tr>
<tr>
<td>11</td>
<td>7.850</td>
<td>7.520</td>
<td>8.070</td>
<td>7.70</td>
<td>8.34</td>
<td>6.86</td>
</tr>
<tr>
<td>12</td>
<td>8.050</td>
<td>7.870</td>
<td>8.060</td>
<td>8.24</td>
<td>8.42</td>
<td>6.98</td>
</tr>
<tr>
<td>13</td>
<td>7.941</td>
<td>7.990</td>
<td>8.060</td>
<td>8.63</td>
<td>8.54</td>
<td>6.26</td>
</tr>
<tr>
<td>14</td>
<td>7.255</td>
<td>7.270</td>
<td>7.070</td>
<td>7.33</td>
<td>7.82</td>
<td>6.06</td>
</tr>
<tr>
<td>15</td>
<td>6.573</td>
<td>6.990</td>
<td>6.040</td>
<td>6.52</td>
<td>7.88</td>
<td>5.61</td>
</tr>
<tr>
<td>16</td>
<td>7.064</td>
<td>7.060</td>
<td>6.790</td>
<td>6.78</td>
<td>7.86</td>
<td>6.57</td>
</tr>
<tr>
<td>17</td>
<td>6.973</td>
<td>7.410</td>
<td>6.890</td>
<td>7.56</td>
<td>8.00</td>
<td>5.27</td>
</tr>
<tr>
<td>18</td>
<td>6.646</td>
<td>6.540</td>
<td>6.680</td>
<td>6.62</td>
<td>7.42</td>
<td>5.50</td>
</tr>
<tr>
<td>19</td>
<td>6.759</td>
<td>6.810</td>
<td>6.640</td>
<td>6.11</td>
<td>7.94</td>
<td>6.00</td>
</tr>
<tr>
<td>20</td>
<td>6.596</td>
<td>6.570</td>
<td>6.210</td>
<td>5.61</td>
<td>7.48</td>
<td>6.58</td>
</tr>
<tr>
<td>21</td>
<td>6.891</td>
<td>6.250</td>
<td>7.430</td>
<td>6.42</td>
<td>7.52</td>
<td>6.91</td>
</tr>
<tr>
<td>22</td>
<td>5.959</td>
<td>6.940</td>
<td>5.400</td>
<td>5.52</td>
<td>7.12</td>
<td>5.09</td>
</tr>
<tr>
<td>23</td>
<td>7.450</td>
<td>7.860</td>
<td>7.430</td>
<td>7.02</td>
<td>8.20</td>
<td>6.63</td>
</tr>
<tr>
<td>24</td>
<td>6.246</td>
<td>6.920</td>
<td>5.330</td>
<td>6.47</td>
<td>6.86</td>
<td>5.21</td>
</tr>
<tr>
<td>26</td>
<td>7.350</td>
<td>7.960</td>
<td>6.810</td>
<td>6.96</td>
<td>8.26</td>
<td>6.64</td>
</tr>
<tr>
<td>27</td>
<td>6.650</td>
<td>7.400</td>
<td>6.000</td>
<td>6.17</td>
<td>7.88</td>
<td>5.93</td>
</tr>
<tr>
<td>28</td>
<td>7.341</td>
<td>7.600</td>
<td>7.350</td>
<td>7.70</td>
<td>7.94</td>
<td>6.28</td>
</tr>
<tr>
<td>29</td>
<td>7.000</td>
<td>6.930</td>
<td>7.390</td>
<td>6.61</td>
<td>7.74</td>
<td>5.68</td>
</tr>
<tr>
<td>30</td>
<td>7.741</td>
<td>6.950</td>
<td>7.710</td>
<td>7.33</td>
<td>8.02</td>
<td>7.80</td>
</tr>
<tr>
<td>31</td>
<td>7.186</td>
<td>7.640</td>
<td>6.650</td>
<td>7.41</td>
<td>7.70</td>
<td>6.32</td>
</tr>
<tr>
<td>32</td>
<td>5.441</td>
<td>4.830</td>
<td>5.940</td>
<td>4.77</td>
<td>6.00</td>
<td>5.42</td>
</tr>
<tr>
<td>33</td>
<td>6.864</td>
<td>7.230</td>
<td>6.440</td>
<td>7.07</td>
<td>8.00</td>
<td>5.49</td>
</tr>
<tr>
<td>34</td>
<td>6.318</td>
<td>6.260</td>
<td>6.200</td>
<td>6.13</td>
<td>7.16</td>
<td>5.33</td>
</tr>
<tr>
<td>Speaker</td>
<td>Overall</td>
<td>Competence</td>
<td>Talent</td>
<td>Production</td>
<td>Imitation</td>
<td>Perception</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>------------</td>
<td>--------</td>
<td>------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>35</td>
<td>7.109</td>
<td>7.050</td>
<td>7.230</td>
<td>6.91</td>
<td>7.72</td>
<td>6.40</td>
</tr>
<tr>
<td>36</td>
<td>6.486</td>
<td>6.730</td>
<td>6.000</td>
<td>6.80</td>
<td>7.54</td>
<td>5.09</td>
</tr>
<tr>
<td>37</td>
<td>6.446</td>
<td>6.050</td>
<td>7.020</td>
<td>6.40</td>
<td>7.74</td>
<td>5.72</td>
</tr>
<tr>
<td>38</td>
<td>6.855</td>
<td>6.970</td>
<td>7.050</td>
<td>7.09</td>
<td>7.74</td>
<td>5.59</td>
</tr>
<tr>
<td>39</td>
<td>7.159</td>
<td>7.640</td>
<td>6.780</td>
<td>8.00</td>
<td>7.64</td>
<td>5.70</td>
</tr>
<tr>
<td>40</td>
<td>7.668</td>
<td>7.800</td>
<td>7.570</td>
<td>7.73</td>
<td>8.16</td>
<td>6.70</td>
</tr>
<tr>
<td>41</td>
<td>8.059</td>
<td>9.010</td>
<td>7.160</td>
<td>8.88</td>
<td>8.28</td>
<td>7.11</td>
</tr>
<tr>
<td>42</td>
<td>6.382</td>
<td>6.330</td>
<td>6.440</td>
<td>6.82</td>
<td>7.70</td>
<td>4.72</td>
</tr>
<tr>
<td>43</td>
<td>6.732</td>
<td>7.320</td>
<td>6.260</td>
<td>6.80</td>
<td>7.62</td>
<td>5.64</td>
</tr>
<tr>
<td>44</td>
<td>7.591</td>
<td>7.690</td>
<td>7.550</td>
<td>8.09</td>
<td>8.42</td>
<td>5.79</td>
</tr>
<tr>
<td>45</td>
<td>7.836</td>
<td>8.410</td>
<td>7.410</td>
<td>8.74</td>
<td>8.60</td>
<td>5.86</td>
</tr>
<tr>
<td>47</td>
<td>7.691</td>
<td>8.070</td>
<td>7.490</td>
<td>8.14</td>
<td>8.60</td>
<td>6.32</td>
</tr>
<tr>
<td>48</td>
<td>7.891</td>
<td>7.460</td>
<td>8.250</td>
<td>7.20</td>
<td>8.98</td>
<td>7.77</td>
</tr>
<tr>
<td>49</td>
<td>7.132</td>
<td>7.150</td>
<td>6.990</td>
<td>7.01</td>
<td>8.06</td>
<td>6.61</td>
</tr>
<tr>
<td>51</td>
<td>7.518</td>
<td>7.350</td>
<td>7.980</td>
<td>7.33</td>
<td>8.34</td>
<td>6.41</td>
</tr>
<tr>
<td>52</td>
<td>7.686</td>
<td>7.880</td>
<td>7.710</td>
<td>8.20</td>
<td>8.42</td>
<td>6.13</td>
</tr>
<tr>
<td>54</td>
<td>6.532</td>
<td>6.760</td>
<td>5.980</td>
<td>5.80</td>
<td>7.86</td>
<td>6.47</td>
</tr>
<tr>
<td>55</td>
<td>7.114</td>
<td>6.910</td>
<td>7.120</td>
<td>6.73</td>
<td>7.92</td>
<td>6.48</td>
</tr>
<tr>
<td>56</td>
<td>7.836</td>
<td>7.740</td>
<td>7.780</td>
<td>7.86</td>
<td>7.80</td>
<td>7.19</td>
</tr>
<tr>
<td>57</td>
<td>7.522</td>
<td>7.920</td>
<td>7.140</td>
<td>8.12</td>
<td>8.12</td>
<td>6.20</td>
</tr>
<tr>
<td>58</td>
<td>6.846</td>
<td>6.470</td>
<td>6.880</td>
<td>6.00</td>
<td>7.84</td>
<td>6.60</td>
</tr>
<tr>
<td>59</td>
<td>7.500</td>
<td>7.600</td>
<td>7.990</td>
<td>7.29</td>
<td>7.78</td>
<td>6.94</td>
</tr>
<tr>
<td>60</td>
<td>8.532</td>
<td>9.100</td>
<td>8.190</td>
<td>9.09</td>
<td>8.74</td>
<td>6.91</td>
</tr>
<tr>
<td>61</td>
<td>7.714</td>
<td>7.380</td>
<td>8.000</td>
<td>7.99</td>
<td>8.44</td>
<td>6.40</td>
</tr>
<tr>
<td>62</td>
<td>7.873</td>
<td>8.160</td>
<td>7.730</td>
<td>8.74</td>
<td>8.30</td>
<td>6.33</td>
</tr>
<tr>
<td>63</td>
<td>8.468</td>
<td>9.120</td>
<td>7.920</td>
<td>9.72</td>
<td>8.68</td>
<td>6.82</td>
</tr>
<tr>
<td>64</td>
<td>7.846</td>
<td>8.280</td>
<td>7.860</td>
<td>8.77</td>
<td>8.70</td>
<td>5.80</td>
</tr>
<tr>
<td>66</td>
<td>7.618</td>
<td>8.430</td>
<td>6.910</td>
<td>8.21</td>
<td>7.60</td>
<td>6.41</td>
</tr>
<tr>
<td>68</td>
<td>6.845</td>
<td>6.550</td>
<td>7.140</td>
<td>6.87</td>
<td>7.28</td>
<td>6.27</td>
</tr>
<tr>
<td>69</td>
<td>8.227</td>
<td>8.370</td>
<td>8.200</td>
<td>8.73</td>
<td>8.44</td>
<td>7.36</td>
</tr>
<tr>
<td>76</td>
<td>5.768</td>
<td>6.850</td>
<td>4.940</td>
<td>5.54</td>
<td>7.26</td>
<td>4.74</td>
</tr>
<tr>
<td>77</td>
<td>7.786</td>
<td>8.860</td>
<td>6.980</td>
<td>8.56</td>
<td>7.92</td>
<td>6.08</td>
</tr>
<tr>
<td>78</td>
<td>7.727</td>
<td>8.030</td>
<td>7.390</td>
<td>8.08</td>
<td>8.16</td>
<td>6.50</td>
</tr>
<tr>
<td>79</td>
<td>7.205</td>
<td>7.740</td>
<td>7.220</td>
<td>7.46</td>
<td>7.50</td>
<td>6.43</td>
</tr>
<tr>
<td>80</td>
<td>7.755</td>
<td>7.830</td>
<td>7.680</td>
<td>8.03</td>
<td>8.48</td>
<td>6.66</td>
</tr>
<tr>
<td>82</td>
<td>8.218</td>
<td>9.150</td>
<td>7.500</td>
<td>9.52</td>
<td>8.78</td>
<td>5.69</td>
</tr>
<tr>
<td>83</td>
<td>7.618</td>
<td>8.200</td>
<td>7.060</td>
<td>7.76</td>
<td>7.92</td>
<td>6.69</td>
</tr>
<tr>
<td>Speaker</td>
<td>Overall</td>
<td>Competence</td>
<td>Talent</td>
<td>Production</td>
<td>Imitation</td>
<td>Perception</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>------------</td>
<td>--------</td>
<td>------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>84</td>
<td>8.236</td>
<td>8.470</td>
<td>7.980</td>
<td>9.26</td>
<td>8.66</td>
<td>6.07</td>
</tr>
<tr>
<td>86</td>
<td>6.591</td>
<td>6.800</td>
<td>6.490</td>
<td>6.82</td>
<td>7.62</td>
<td>5.50</td>
</tr>
<tr>
<td>87</td>
<td>6.727</td>
<td>7.050</td>
<td>6.370</td>
<td>7.68</td>
<td>7.80</td>
<td>4.66</td>
</tr>
<tr>
<td>88</td>
<td>7.545</td>
<td>7.620</td>
<td>7.300</td>
<td>7.83</td>
<td>7.94</td>
<td>6.20</td>
</tr>
<tr>
<td>89</td>
<td>7.745</td>
<td>8.330</td>
<td>7.290</td>
<td>8.39</td>
<td>8.20</td>
<td>5.99</td>
</tr>
<tr>
<td>91</td>
<td>7.436</td>
<td>6.830</td>
<td>7.780</td>
<td>7.44</td>
<td>8.00</td>
<td>6.73</td>
</tr>
<tr>
<td>93</td>
<td>7.945</td>
<td>8.780</td>
<td>7.510</td>
<td>9.03</td>
<td>9.42</td>
<td>5.16</td>
</tr>
<tr>
<td>94</td>
<td>6.764</td>
<td>7.360</td>
<td>6.330</td>
<td>7.46</td>
<td>7.48</td>
<td>5.14</td>
</tr>
<tr>
<td>95</td>
<td>6.323</td>
<td>7.600</td>
<td>5.210</td>
<td>6.68</td>
<td>7.74</td>
<td>4.70</td>
</tr>
<tr>
<td>96</td>
<td>7.118</td>
<td>7.560</td>
<td>6.930</td>
<td>7.88</td>
<td>8.14</td>
<td>5.00</td>
</tr>
<tr>
<td>97</td>
<td>8.195</td>
<td>8.190</td>
<td>8.250</td>
<td>8.69</td>
<td>8.36</td>
<td>7.14</td>
</tr>
<tr>
<td>98</td>
<td>3.836</td>
<td>3.100</td>
<td>4.240</td>
<td>2.52</td>
<td>4.36</td>
<td>5.10</td>
</tr>
<tr>
<td>99</td>
<td>7.223</td>
<td>7.240</td>
<td>7.580</td>
<td>8.42</td>
<td>7.46</td>
<td>5.53</td>
</tr>
<tr>
<td>100</td>
<td>6.032</td>
<td>6.310</td>
<td>5.880</td>
<td>6.24</td>
<td>7.02</td>
<td>4.82</td>
</tr>
<tr>
<td>101</td>
<td>7.591</td>
<td>7.050</td>
<td>7.670</td>
<td>7.32</td>
<td>8.26</td>
<td>6.87</td>
</tr>
<tr>
<td>102</td>
<td>6.273</td>
<td>6.600</td>
<td>5.650</td>
<td>5.39</td>
<td>7.00</td>
<td>6.06</td>
</tr>
<tr>
<td>103</td>
<td>6.800</td>
<td>8.020</td>
<td>5.950</td>
<td>7.40</td>
<td>7.58</td>
<td>5.46</td>
</tr>
<tr>
<td>104</td>
<td>7.164</td>
<td>7.520</td>
<td>6.850</td>
<td>7.47</td>
<td>7.98</td>
<td>5.83</td>
</tr>
<tr>
<td>105</td>
<td>5.010</td>
<td>5.120</td>
<td>4.410</td>
<td>4.73</td>
<td>5.28</td>
<td>4.79</td>
</tr>
<tr>
<td>106</td>
<td>6.559</td>
<td>6.200</td>
<td>6.770</td>
<td>6.00</td>
<td>8.00</td>
<td>6.00</td>
</tr>
<tr>
<td>108</td>
<td>7.268</td>
<td>7.390</td>
<td>7.010</td>
<td>7.17</td>
<td>7.82</td>
<td>6.70</td>
</tr>
<tr>
<td>109</td>
<td>6.177</td>
<td>5.830</td>
<td>6.140</td>
<td>5.58</td>
<td>6.82</td>
<td>6.18</td>
</tr>
<tr>
<td>111</td>
<td>6.064</td>
<td>5.880</td>
<td>5.970</td>
<td>4.96</td>
<td>7.90</td>
<td>5.70</td>
</tr>
<tr>
<td>112</td>
<td>7.259</td>
<td>6.710</td>
<td>7.770</td>
<td>7.21</td>
<td>8.00</td>
<td>6.76</td>
</tr>
<tr>
<td>113</td>
<td>6.632</td>
<td>7.110</td>
<td>6.480</td>
<td>7.00</td>
<td>8.60</td>
<td>4.60</td>
</tr>
<tr>
<td>114</td>
<td>6.414</td>
<td>5.990</td>
<td>6.310</td>
<td>5.47</td>
<td>7.38</td>
<td>6.11</td>
</tr>
<tr>
<td>115</td>
<td>5.582</td>
<td>5.850</td>
<td>5.100</td>
<td>5.07</td>
<td>6.60</td>
<td>5.36</td>
</tr>
<tr>
<td>116</td>
<td>7.605</td>
<td>7.940</td>
<td>7.470</td>
<td>7.97</td>
<td>8.10</td>
<td>6.34</td>
</tr>
</tbody>
</table>
Cognitive aspects of pronunciation talent

Giuseppina Rota and Susanne Maria Reiterer

1. The significance of language and speech abilities

Pierre Fouché, a renowned professor of phonetics at Sorbonne, allegedly told his students who could imitate a number of different pronunciations, that “a strong personality requires just one accent”. In his case it was evidently true as he spoke lots of languages with a strong Catalan accent.

Are there really personality factors that prohibit from acquiring new languages (L2) and speaking them with a native-like capacity? Imagine two hypothetical individuals who are living in the same foreign environment. Presumably, a highly motivated subject will try to take advantage of every given circumstance to improve his pronunciation skills, while a less motivated person will devote fewer resources to this task. In such a case, are there psychological and cognitive characteristics that might help the reluctant learner and give him an advantage over one who is more motivated? Is there an innate or acquired set of personal skills, which even though apparently unrelated to speech production, might facilitate the process of building up linguistic competences, thus paving the way to a native-like pronunciation of L2?

The answers to these questions have crucial pedagogical implications. Usually, grammar notions and phonetic abilities are used as selection criteria to enroll students into language classes. However, tailoring selection criteria, as well as teaching techniques, to the abilities and the needs of the students would allow them to benefit more from the courses. The importance of individual differences, and the way they impact the acquisition of L2 has long been overlooked. Similarly, the question as to why some people have less difficulties learning L2, and why some develop a finer phonetic proficiency than others remains unanswered.

Stimulated by this lack of evidence, we investigated crucial cognitive factors and tested their relevance in groups of subjects classified on the basis of their proficiency and talent in L2 pronunciation. In the present chapter, we address the results of this research with respect to empathic skills, mental flexibility, working memory abilities and intelligence.
2. Empathy

Essentially, to learn a second language is to take on a new identity

Guiora, 1972

Are empathetic capacities important for the acquisition of certain cognitive skills and for the ability to correctly pronounce L2? The debate on the nature of empathy first came to light in the last century. Robert Vischer coined the concept of *Einfühlung, in-feeling* or *feeling-into* in 1873, and the German philosopher Theodor Lipps firstly applied it in psychology. Lipps identified in empathy a crucial component for the fruition of aesthetic experiences in art. Nowadays, the term *empathy* describes the mechanisms that form the ability to perceive and feel the emotional states of others, the *emotional resonance* that takes place in an individual when he/she engages in an affective communication or in an emotional exchange. The concept of empathy accounts for the emotional and cognitive reactions elicited in a person when he/she observes another person’s emotional expressions and/or behaviour. In this sense, Mehrabian and Epstein (1972) suggested that empathy is a *vicarious emotional reaction*, and they proposed an *emotional contagion* hypothesis to define its nature.

An alternative perspective emphasizes the cognitive components of empathy and defines it as a cognitive understanding of the other’s situation, based on the capacity of assuming his/her point of view on things (the *perspective taking* hypothesis, Hogan 1969).

Salovey and Mayer (1990) stressed the importance of empathetic abilities for an effective and satisfying interaction and coined the term *emotional intelligence*. According to the authors, individuals who possess an accurate ability to perceive and appraise others’ emotions are able to respond in a more flexible way to changes happening in their social environment, and to build up and maintain supportive social networks over time (Salovey et al. 1999). The concept of emotional intelligence is relatively new. However, the importance of non-intellective abilities such as sensitivity and emotional skills for predicting individual success in life has been recognized for years. Rosenthal and colleagues discovered that people who could better identify others’ emotions are more successful in their private lives, as well as in their jobs (Rosenthal 1977). An empathic person is prone to support others and is likely to get more support back (Håkansson and Montgomery 2003). According to the *empathy-
altruism-hypothesis (Batson 1990), empathy nurtures the willingness to engage in pro-social behaviors, and to act altruistically. With this regard, empathy might be seen as a core component of social intelligence (Amelang et al. 1989), and presumably sustains the recognition of relevant social cues, integrating this knowledge into social interaction.

Some researchers showed that the importance of empathic abilities is not limited to real-life situations, but extends to fictitious circumstances as well (Davis 1983; Stotland et al. 1978). In particular, the Don Quixote effect (Shapiro and Rucker 2004) creates the basis for highly empathic reactions to films and movies. This phenomenon works by making the viewers experience fictitious idealized schemas where emotional contents can become even more intense than in real-life contexts.

In the 70’s, Guiora and his colleagues proposed the idea that interpersonal competence plays a role during the acquisition of phonetic skills, thus favouring L2 pronunciation. Guiora postulated that individual differences in L2 pronunciation reflect the degree of empathic capacities that people possess. These researchers proposed the concept of language ego (Guiora et al. 1972a), which is a representation of the self defined by physical and psychological boundaries. According to this conceptualization, the development of native speech abilities takes place during a state of ego-permeability, which is prevalent in early development stages, and progressively decreases with age. For a child the assimilation of L2 represents a natural process. However, this process is progressively inhibited in adulthood, and may be the reason for maintaining a foreign accent even when L2 is spoken fluently. Kris (1952) anticipated this concept by identifying in ego-relaxation processes the mechanisms that promote creative acts. During their creation, verbal-logical operations are abandoned and a primary thinking mode, typical of the childhood, emerges. According to Guiora and co-workers (1972a), the acquisition of L2 poses a challenge to the integrity of basic identifications, because it requires the individual to step into a new world and take a new identity. In their definition, empathy requires a process in which a temporary fusion of object-self boundaries takes place, thus opening up to an immediate emotional apprehension of the affective experience of the others. This process is a crucial component for the development of native-like pronunciation skills because L2 learners need to acquire a new identity to achieve new linguistic competences. Ego-permeability enhances the speaker’s sensitivity to social interaction. It forms the basis of enhanced abilities in pronunciation because it allows speakers to discern subtle speech nuances and to learn to reproduce them.
Guiora and colleagues tested their model in a number of experiments. In one of the first studies (Guiora et al. 1967), they observed a correlation between accuracy in French pronunciation and scores on empathy skills obtained with the Micro-Momentary Expression Device. In this experiment, Guiora and colleagues assessed empathy by addressing the ability to detect subtle changes in facial expressions. This indirect measure of empathetic skills failed to correlate with pronunciation abilities in further investigations (Taylor et al. 1970). The researchers recruited 28 college students who participated in a short course in Japanese, and observed a negative correlation between their pronunciation performances and the MME scores. In a further study, 401 students belonging to the Defense Language Institute were administered the MME and other tests. This research led to a positive correlation between pronunciation scores for several languages (Japanese, Spanish, and Russian) and the MME scores. The existence of controversial evidence in these results might be attributed to the fact that the MME does not address the totality of the empathy phenomenon in its multiple sub-components such as body posture, voice intonation etc. Being able to decipher facial expressions is an important component of emotional intelligence, but it certainly does not embody the totality of it.

Guiora identified different methods to operationally induce ego-permeability. He viewed alcohol consumption as a useful means to lower subjective inhibitions and feelings of separateness of identity, and tested it in a series of experiments (Guiora et al. 1972a, 1972b). Guiora and colleagues investigated the pronunciation of Thai sentences and suggested that a temporary improvement of L2 pronunciation abilities follows the ingestion of small doses of alcohol (1 to 1½ ounces of alcohol, not on an empty stomach). In his explanation of the phenomenon, Guiora describes an induced flexible psychic state (Guiora et al. 1975), with temporary lower inhibitions and presumably heightened empathy. He hypothesized that an early positive stage of intoxication is optimal for L2 pronunciation skills. Although this claim might sound very appealing, it is clear that the recourse to alcohol cannot be considered an acceptable pedagogical method.

Encouraged by those results, Guiora and colleagues extended their research to the use of benzodiazepine (valium) and hypnosis. Schumann et al. (1978) reported a significant improvement of L2 pronunciation for deeply hypnotized subjects, thus providing support for the claim that a letting go state of feelings facilitates performances. This intuition failed to find confirmation in further studies that addressed the role of benzo-
diazepine (Guiora et al. 1980). While it seems that relaxation and low anxiety levels have a positive effect upon subjective abilities, no causal correlation between ingestion of small doses of benzodiazepine and L2 pronunciation was observed.

2.1. Assessing empathy

On the basis of this pioneering work, we decided to consider empathetic skills among those psychological factors that might correlate with talent in L2 pronunciation and, therefore, to test them. We examined empathetic skills using the questionnaire E-Scale, a psychometrically sound instrument devised by Leibetseder and colleagues (2001). The authors adopted the definition of empathy as *the effort to identify with persons in fictitious or real-life situations* (Leibetseder et al. 2007). The questionnaire explores two different dimensions: *sensitivity*, and *concern*. The dimension *sensitivity* addresses empathetic abilities elicited towards fictitious imagined situations. The dimension *concern* assesses empathy experienced when imagining real-life contexts. The 26 items that constitute the questionnaire investigate a total of four factors: cognitive-sensitivity, emotional sensitivity, emotional and cognitive concern. The questionnaire addresses cognitive sensitivity by means of items that describe fictive but concrete social situations where empathic reactions are to be elicited. Emotional sensitivity is investigated by questions that induce the empathizer to put himself into the emotional situation of the actor and deal with fictitious situations. Items address emotional concern by investigating reactions to situations characterized by a high degree of reality. Cognitive concern is studied by means of items that stimulate empathetic feelings to concrete situations of real-life, which require a cognitive analysis. The general score on empathy obtained by means of this questionnaire gives an IQ on empathy (E-IQ), which mirrors the general ability of the individual to experience empathetic feelings in both fictional and real-like life situations.

Our research group investigated with a large-scale assessment the question whether phonetic abilities in L2 pronunciation correlate with empathetic openness to others, as suggested by Guiora and his co-workers. German subjects underwent multiple phonetic assessments (for details, refer to Jilka in this volume), where we investigated phonetic proficiency in both L2 pronunciation and perception.

Subjects were classified into three groups: highly talented subjects, average subjects, and anti-talented individuals. Mean levels of E-IQ cal-
Giuseppina Rota and Susanne Maria Reiterer

culated taking into account the scores for all the items of the E-Scale questionnaire were 47.6 for the talented subjects (N= 20), 45.2 for the average talent group (N = 20), and 42.2 for the anti-talent subjects (N= 20). Correlation analyses were calculated for proficiency, talent, scores on the Modern Language Aptitude Test (MLAT), Hindi (as an unknown language) imitation and E-IQs scores. Furthermore, we correlated the E-IQ scores with responses on the following three introspective questions: 1. How much do you like imitating voices? (scale 1–5 points, 5=max), 2. How much do you like imitating accents? (1–5) and 3. How many instruments do you play?

The results were several significant correlations, which were in the weak-middle range only for the first subscale “empathic readiness” (not for the second subscale of the E-Scale “social concern”). E-IQ (empathic readiness) scores correlated positively and significantly with: talent of pronunciation (according to our talent score (see chapter M Jilka)) (r = 0.26, P = 0.024*, N = 60), with MLAT3 – phonetic coding ability – r = 0.2, P = 0.04*, N = 60 and with the total MLAT score (comprising the three sub-scales: phonetic coding ability, grammatical sensitivity and vocabulary learning), r = 0.2, P = 0.047, N = 60). Furthermore, it correlated with performance (proficiency in L2 pronunciation) r = 0.29, P = 0.013*, N = 60; imitation capacity for unknown Hindi words: r = 0.16, P = 0.048, N = 60, as well as with the three introspective questions: 1. “Do you like imitating voices?” r = 0.23, P = 0.005**, N = 60; 2. “Do you like imitating accents?” r = 0.31, P = 0.001**, N = 60; 3. “Number of musical instruments you play” r = 0.24, P = 0.004**, N = 60. Thus, the higher the subjects were rated on the “talent” score in pronunciation, the higher their score on the Modern Language Aptitude Test for “phonetic coding ability”, the higher their actual performance or proficiency in pronouncing L2 texts, the higher their scores given by Hindi native speakers on the Hindi word imitation task, the higher the Empathy score on the subscale “Empathic readiness” were.

When asked to self rate (on a scale from 1–5), their liking for imitating voices and accents the correlation was even higher and more significant. What can be seen as an interesting collateral result is that their empathic readiness correlated even with the number of instruments they played (the more instruments, the more empathic).

These correlations between empathic readiness and musical instruments point to the intrinsic and often hypothesized interwoveness between phonetic language talent and music (for this aspect see Chapter by Nardo and Reiterer, pp. 213). However, the correlations of Empathy to
our various measures of phonetic pronunciation talent show that the theory of Empathy (in connection also with personality, see also Chapter 4) in speaking and learning a foreign accent is perhaps not a myth, but a must to be further elucidated.

3. Mental flexibility

Do we need a certain degree of mental flexibility to learn L2 and to be able to pronounce phonetic features that greatly differ from the ones we are used to in our native tongue? Does speaking a second language, or multiple languages, affect brain plasticity? Can we maintain a high cognitive functioning if we continue to train our brain as time passes? Questions such as these animate scientists as they contend about the role of nature and nurture in shaping individual predispositions and abilities. A child possesses an extraordinary brain plasticity that allows him to learn every language existing on earth. As he grows up, his linguistic skills specialize for his mother tongue and the ability to assimilate new languages gradually diminishes. As he becomes older, he is still able to learn foreign languages, but he is likely to retain an accent that traces back to his native pronunciation. Are there mechanisms that shape this process? If this is the case, can we identify them and possibly reverse their effect?

Studies on genetics have just begun to explore how the brain develops and functions. The human brain is a dynamic and constantly reorganizing system that benefits from stimulating environments at every age: its plasticity is a lifelong phenomenon. Since the 50s, numerous studies conducted on animals have attempted to determine whether the brain develops differently in response to a variety of environmental conditions. Researchers devised standardized paradigms, typically raising two groups of rodents in enriched and/or un-enriched environments (Rosenzweig and Bennett 1972; Rosenzweig et al. 1969). The experimenters kept the animals of the first group in isolation, a condition of poor sensory, cognitive and motor stimulation. They provided the animals of the second one with increased social interaction, and great exposure to stimulation (e.g. many animals were typically housed together in big cages containing a great variety of toys with different shapes and color). When examining the brains of both groups, they discovered significant differences in the growth of the animals’ cells: the rodents raised in the enriched environment had a larger cortex and a higher cortical density due to an increased
amount of brain cells and cellular connections. Is there an analogous effect in children who grow up in a bilingual or multilingual – highly stimulating – environment? Does a linguistically stimulating environment affect other cognitive domains as well?

The claim about the positive effects of bilingualism on cognitive performances was first made in 1962. Peal and Lambert (1962) studied French Canadian children and observed a greater mental flexibility and verbal fluency for this bilingual group. They proposed that the bicultural environment in which the children grew up enriched both verbal and non-verbal skills. Since this study was published, great interest has arisen regarding the cognitive advantages that may derive from exposure to multiple languages from childhood onwards. A further study conducted on Finnish students suggested that bilingualism might influence intelligence: an average of higher school grades for bilinguals was observed when their performance was compared to monolingual subjects (Sundman 1994). Existing evidence suggests that bilingual children often perform better than monolingual children in tasks that require cognitive control (Bialystok 2001). From an early age, bilingual children are exposed to multiple languages and trained to process information on different registers. As a result of this process, they may develop enhanced attentional capacities. While speaking, fluent bilinguals need to suppress irrelevant intrusions of the language not currently in use. Usually, interlocutors do not perceive this suppression and they are not aware of the effort that sustains such a process. However, psycholinguistic studies conducted on bilinguals indicated that even if the speaker engages in a communication speaking one language only, the representations of both his native languages become simultaneously active (Kroll and Dijkstra 2002; Gollan and Kroll 2001; Francis 1999). According to Bialystok (2006), bilinguals are constantly trained to direct selective attention to relevant information and to ignore irrelevant but highly distracting representations. As documented by Bialystok and Codd (1997), 4- and 6-year-old bilinguals perform better than comparable monolinguals during counting tasks challenged by contradictory information. Similarly, an analogous performance is observed when bilinguals engage in sorting tasks where the label on the sorting compartment misleads the decision task (Bialystok and Martin 2004).

The same response-trend has been observed for the Simon task. Martin and Bialystok (2003) evaluated the performances of bilingual preschool children and showed that these subjects had faster reaction times when compared with the monolingual peers. This evidence was found not only
for incongruent trials, but for congruent ones as well. This result was replicated in further investigations that recruited subjects whose age varied between 30 and 80 years. Older bilingual volunteers were faster and more accurate than younger ones (Bialystok et al. 2004).

The Simon Task (Simon and Rudell 1967) was originally devised as a tool to test the effects of handedness and gender on selective attention. So far, this test has been widely explored in a number of paradigms that range from executive functions to cognitive control (for review refer to Lu and Proctor 1995). When engaging in this task, participants need to simultaneously process stimuli (e.g. colored circles that appear on a display) that have relevant features (e.g. the colour) and irrelevant highly distracting ones (e.g. the position on a screen, such as left or right). Participants are required to perform speeded button-press responses when they detect target features of the stimuli, while ignoring irrelevant misleading information. Subjects are faster when responding to congruent stimuli (e.g. red stimuli appear on the left side of the screen and they have to press a key on the left for red) and slower when responding to incongruent ones (e.g. blue stimuli appear on the left side of the screen and they have to press right for blue). Subjects need to inhibit automatic response processes to respond correctly to incongruent stimuli. Response-inhibition is a time-consuming process and typically leads to lower accuracy levels and slower reaction times, e.g. around 20–30 milliseconds (Ridderinkhof 2002; Gratton et al. 1988; Simon 1969; Fitts and Seeger 1953). This phenomenon is known as the Simon effect and it has been observed for a number of relevant features such as pitch and form (Simon and Berbaum 1990; Craft and Simon 1970; Simon 1969).

To explain why the Simon effect arises, researchers have referred to a dual-routes processing model (Wiegand and Wascher 2005; Ridderinkhof 2002; Eimer 1995). According to this model, there is an unconditional route in which responses are performed in a quick and semi-automatic way, and a conditional route, in which appropriate responses need a relatively slow intentional process in order to be performed. In the Simon task, the stimulus location induces a priming effect on the ongoing response decision by automatically activating the corresponding response. Interference arises during incongruent trials because an irrelevant feature of the stimulus activates the associated response in an automatic fashion. As the two responses enter in conflict with one another, the stimulus-response mapping process needs to override the incorrect response activation (Coles et al. 1985). This model has received an important validation from studies conducted on the lateralized readiness potential (LRP). The
LRP can be recorded by measuring event-related potentials of electrodes placed on the left and right motor cortices separately for left and right hand responses (de Jong et al. 1988; Gratton et al. 1988). The LRP indicates the preparation to move one hand to respond and its course over time. Numerous experiments have shown that incongruent trials immediately activate the incorrect response and are followed by a later activation of the correct one.

Many studies suggest that the neuronal network that is recruited during the Simon task sustains a number of other attention-demanding tasks as well (Fan et al. 2003a; Fan et al. 2003b; Peterson et al. 2002). Comparing competitive alternatives that generate conflictual tendencies and the suppression of interference seem to rely on a cortical network that encompasses the anterior cingulate cortex (ACC), the left prefrontal cortex (PFC), and the dorsolateral prefrontal cortex (DLPFC) (Kondo et al. 2004a, b). An enhanced functioning of this network presumably affects all attentional performances that rely on these brain areas. The hypothesis that these attentional processes are regulated by cortical networks that may be modified by experience was suggested by Merzenich and Jenkins (1993). Training and experience seem to play a crucial role in shaping cortical plasticity and to be crucial for the specialization of brain networks (Fan et al. 2003b). Following these lines, prolonged experiences that rely on attentional control, as it happens in bilingualism, presumably modify its neuronal bases, and strengthen their functioning.

3.1. Testing mental flexibility

Bialystok suggests that bilingualism creates cognitive advantages, and argues that bilingualism is a sort of mental fitness training, a protective factor against the decline of cognitive control and attention switching abilities, which increases with age.

In our study, we tested the hypothesis that the acquisition of L2 after puberty influences selective attention and task switching capacities. We used a classical version of the Simon task and ran it on a laptop (HP, ADM Athlon 3200) with a 12-in monitor. We programmed and ran the task using the E-Prime software (Schneider et al. 2004). The task was constituted of 4 runs, presented in sequence, each of which lasting 47.4s. Each run started with a fixation cross (+) presented in the center of the screen for 3s and followed by a blank interval of 1s. We presented either a red or a blue circle on the left \((x = 0.02^\circ, y = 0.36^\circ)\) or on the right \((x = 0.82^\circ, y = 0.36^\circ)\) of the laptop screen. Each stimulus lasted 1s.
instructed the subjects to respond to the presentation of a red circle by pressing a key on the left side of the keyboard (i.e. the letter s), and to respond to the presentation of a blue cycle by pressing a key on the right of the keyboard (i.e. the letter l), independently of the circles’ position on the screen. Circles (24 for each run, randomly presented) were separated by an interval of 900ms during which subjects could view a black screen only. For half of the cases, subjects were confronted with items that appeared on the same side of the screen as the key that they had to press (congruent condition), and for the other half to items that appeared on the opposite side (incongruent condition) of the screen. Before engaging in the Simon task, subjects underwent a practice run during which they were instructed to respond to 24 items (12 congruent and 12 incongruent).

We tested the performances (i.e. accuracy and reaction times) of our volunteers, classified in three groups (high, average and low proficiency talent in L2). We expected to find a significant difference in the performances in the subjects who were classified as talented versus the subjects who were classifies as untalented. In line with the hypotheses by Bialystok, we expected that talented individuals would show higher accuracy and reduced reaction times than controls (i.e. non-talented subjects). Talented subjects (N = 23) performed with an accuracy of 96% ± 4% (mean ± SD) during the congruent condition and with an accuracy of 83% ± 7% (mean ± SD) during the incongruent condition. Their reaction times were 385.6 ms ± 40.7 (mean ± SD) when responding to the congruent condition and 412.3% ± 44.6 (mean ± SD) when responding to the incongruent condition. Untalented subjects (N = 23) had an accuracy of 96% ± 4% (mean ± SD) for congruent trials and an accuracy of 92% ± 8% (mean ± SD) for incongruent trials. Their reaction times were 413.6 ms ± 71.4 (mean ± SD) for the congruent condition and 435.4% ± 64.08 (mean ± SD) for the incongruent condition.

We computed a linear regression analysis for the proficiency and accuracy scores of our subjects. We observed no significant correlation for L2 proficiency and accuracy in the congruent (one-tailed linear regression, r = –0.124, P = 0.311, N = 69), and incongruent conditions (one-tailed linear regression, r = –0.231, P = 0.056, N = 69). Reaction times negatively correlated with levels of proficiency (congruent condition: r = –0.268, P = 0.026*, N = 69; incongruent condition: r = –0.274, P = 0.023*, N = 69).

The results of our investigation indicate that no correlation between accuracy in the Simon task and L2 proficiency exists when L2 is ac-
quired after puberty. The data suggest that mental flexibility as measured by the Simon task is not influenced by the subject’s proficiency. In a series of experiments, Bialystok and colleagues (2004) showed that bilinguals perform better in the Simon task when compared to monolinguals. Bilingualism affects cognitive control as subjects need to train in the management of multiple languages since early age. The results of our study indicate that learning L2 after puberty does not affect brain plasticity as it does in early age. Recent findings show that bilingualism influences cortical plasticity enhancing the density of grey matter in left inferior parietal cortex (Mechelli et al. 2004). This effect correlates with the level of proficiency in L2 and its age of acquisition: brain structural changes are more evident when L2 is acquired before 5 years of age. This evidence and our findings lead us to suggest that the acquisition of L2 after puberty may have a reduced impact on brain plasticity, thus diminishing the effect of linguistic management on cognitive control.

4. Working memory

Working memory (WM) can be defined as a temporary retention of recently acquired information. While kept in memory, information can be actively manipulated and used for a variety of cognitive operations and goal-directed behaviors. The ability of holding information and manipulating it forms the substrate of multiple cognitive skills and is essential to operations such as reasoning, planning, the processing of speech and spatial inputs.

Baddeley and Hitch (1974) proposed a cognitive model of WM that has become highly influential over the years stimulating a great amount of research both in the fields of cognitive psychology and imaging neuroscience. These researchers coined the term working memory to indicate short-term storage of information and emphasized the key role of this system in cognitive processing and manipulation of data. They observed that individual performances deteriorate with the increase of concurrent memory load. WM refers to the total resources that the individual can rely on for simultaneous processing and storage. This system is an interface between perception and action and is the basis for a wide range of cognitive activities: the pieces of information acquired continue to be at the disposal of the cognitive system and is used for action and interaction.
The WM system comprises distinct components that sustain cognitive control and active maintenance of information: a central executive and two storage systems that retain verbal and spatial data, the phonological loop and the visuospatial sketchpad (Burgess et al. 2007; Stuss and Alexander 2007; Baddeley 1986). The central executive is a controller system that is responsible for the allocation of attentional resources and for the coordination of information within both verbal and spatial storage buffers. It is an interface between the two storage buffers and long-term memory that coordinates information originating from different sources. The central executive system activates when incoming cues force ongoing automatic behavior to be suspended, attentional control engaged, and action adjusted to meet new environmental demands. Evidence for the existence of the visuospatial sketchpad, which is involved in the manipulation of information on objects and spaces, originally came from clinical reports on soldiers who were wounded with gunshots during the World War I, and whose capacity to encode visuo-spatial information was impaired (Holmes 1919). A number of paradigms rely on this component, such as the Corsi’s Block Tapping Test (Milner 1971). This test is constituted of an array of blocks positioned in a semi-random fashion. The experimenter shows a sequence of blocks and the subject has to reproduce it correctly for a number of times. The phonological loop is engaged in the maintenance of linguistic information. This sub-system plays an important role for the acquisition of L2, and will be addressed with more details in the next chapter.

Empirical studies of WM have relied on behavioral measures and neuro-imaging techniques. These methods have enriched our knowledge on how this system works and have provided important insights into the neural networks that sustain its functioning. Already in the 50’s, empirical studies showed that small amounts of information are rapidly forgotten unless actively rehearsed (Brown 1958). The very first reports to this regard came from the work by Miller (1956), who quantified the limit of the short-term memory capacity as being around the magical number seven. Miller demonstrated that memory span could be measured and proposed that young adults are able to remember a range of seven plus/minus two elements, regardless of whether the elements are words, letters, digits, or other units. Further observations showed that the ability to encode information depends on a number of features concerning the material to be remembered. For instance, WM span is usually lower for long words than for short words. When verbal material such as digits, letters, words, etc. need to be encoded and recalled, an important role is played
by the time necessary to repeat those contents aloud and by the lexical status of the materials (Hulme et al. 1995). Cowan (2001) has proposed that young adults have a WM capacity of about four chunks, children and older adults a much poorer one. A number of other factors affect the WM span and so far the exact extent of the WM capacity across individuals remains an unsolved issue.

WM capacities have been recognized as a key factor for scholastic success. According to Gathercole and colleagues (2004), the expansion of WM span and the improvement of memory performances during development are due to an increase of rehearsal strategies that become available to the children. A number of studies show that WM skills predict children’s academic achievements in a variety of sectors such as mathematics and reading (Swanson 2006; Swanson et al. 1996; Gathercole et al. 2004). On the other hand, WM skills are among the executive functions that are typically compromised in children with high levels of impulsivity and inattention who are affected by the attention deficit/hyperactivity disorder (ADHD, Barkley 1997; Castellanos et al. 2006; Martinussen et al. 2005). Martinussen and Tannock (2006) documented poor performances in ADHD children who engage in complex memory tasks involving the storage and processing of both verbal and visuo-spatial data.

The importance of WM has been observed for disparate sectors. We address here the question whether WM capacities play an important role while learning new languages and whether the extent of the WM span is a good predictor of the ability to pronounce L2 with a native-like accent. Could we assume that by training those skills, people would need less effort to learn L2, and/or would improve their pronunciation skills? Could this method be an indirect but effective way for stimulating and strengthening linguistic abilities?

4.1. The phonological loop

In the model proposed by Hitch and Baddeley, the WM system relies on a linguistic sub-component, called the phonological loop, consisting of a phonological store, which functions through an articulatory rehearsal process (Baddeley 2003). According to Atkins and Baddeley (1998), this system is the structure that allows for language learning. The phonological loop forms the basis of the ability to immediately repeat novel phonological strings or non-words, and more importantly, of the ability to memorize this material. The phonological loop is believed to hold memory traces of linguistic data for a short time, and the amount of material
that it is able to store reflects its encoding capacity. Articulatory rehearsal processes allow for longer maintenance of the information. Active rehearsal of verbal material consists of its repetition, a process during which attention is repetitively directed to the items to remember, and subvocal articulation takes place (Baddeley 1986).

Empirical evidence shows that the suppression of articulation by means of tasks that engender cognitive interference (e.g. uttering nonsensical sequences of words during a detention interval) degrades retrieval (Murray 1968). On the contrary, sub-vocal repetition makes the information fade slower, and performed over time facilitates long-term encoding. The permanence of mnestic traces increases the likelihood that the information will pass from short-term memory to long-term memory. Re-articulation is not the only condition that facilitates encoding. Phonological characteristics of the material play an important role in preventing the loss of memory traces: when encoding unrelated letters, listeners recall sequences of dissimilar sounding letters (e.g. X, K, R, Y, Q) better than sequences of similar ones (e.g. P, B, T, etc.). While memorizing meaningless material, listeners need to rely on its phonological characteristics (Baddeley 2003). As for letters, the similarity of sounds is crucial for unrelated words. This process is analogous to what happens when listeners try to learn new languages, and they are confronted with phonological features that differ from their native tongue’s ones.

Baddeley (2003) proposed that the phonological loop played a crucial role in the evolution of the human species because it permitted the acquisition and the preservation of language across generations. This claim was corroborated by the clinical observation of a patient, whose pure phonological loop deficit correlated with his inability to acquire the vocabulary of the L2 (Baddeley 2003). Moreover, this claim was sustained by the observation that critical factors such as long world-length and articulatory suppression, which impede the functioning of the phonological loop, do interfere with the acquisition of new languages (Papagno and Vallar 1992; Papagno et al. 1991).

It seems that phonological loop capacity is a good predictor for the ability to acquire L2 in developmental and adult ages (Service 1992; Atkins and Baddeley 1998). Gathercole and Adams (1994) investigated the ability of children to correctly repeat unfamiliar non-words. They observed a correlation between the performance in this task and the children's ability to acquire their native language. Conversely, linguistic disability seems to be a valid predictor for poor performances during non-word repetition tasks, even though the level of general intelligence is
normal and no impairment in hearing and articulation can be reported (Gathercole and Baddeley 1990).

How can temporary storage capacities facilitate the process of learning new words? Baddeley proposes that the phonological loop temporarily stores the representations of new linguistic material and structures, and through reiteration impedes shading, thus allowing for a fixation of the mnestic traces to take place. This process does not require long-lasting exposure to L2 if the sequences are phonotactically regular, though it may need longer training if they are irregular (Baddeley 2003).

Neuroimaging studies have attempted to locate the cortical networks that house WM capacities and their sub-components. Studies on patients with brain injury and imaging investigations converge, suggesting that the different sub-components of the WM system are segregated in different parts of the brain, linked by reciprocal interaction (Fuster 2003). Brain regions that exhibit persistent neural activity during active maintenance of task-relevant information are good candidates for the neuronal networks that sustain WM. Left temporoparietal areas are engaged in phonological tasks that require WM (Paulesu et al. 1993; Vallar et al. 1997; Warrington et al. 1971). The supramarginal gyrus is where the phonological storage seems to be located in the brain. The ventrolateral frontal cortex (i.e. Broca’s area) is often activated when subvocal rehearsal is the main strategy for encoding (Bench et al. 1993; Awh et al. 1996). Verbal tasks that require WM typically engage the left inferior parietal lobe, premotor cortices and the cerebellum (Awh et al. 1996; Paulesu et al. 1993). The right hemisphere plays an important function by primarily housing the visuospatial component of the WM system (Smith et al. 1996). However, the rehearsal of task-relevant representations relies on the interactions between networks of brain regions in both hemispheres (Fuster 2003).

4.2. Testing working memory span

Our group assessed the digit span and word span of all subjects who participated in our psychological tests. We used the Wechsler Digit Span test, a subtest of the revised version of the Wechsler Adult Intelligence Scale (Wechsler 1939), to measure digit span. In this test, subjects are required to perform forward and backward repetition of strings of numbers, which contain an increasing amount of items (from 3 to 9 numbers for forward repetition, and from 2 to 8 numbers for backward repetition).
To assess word span we devised a further test by creating strings that contain an increasing number (from 2 to 8) of monosyllabic non-words, with a German-like phonetic quality. For each string, subjects were required to repeat as many items as possible. We measured accuracy of repetition in both tests. On the basis of the mentioned literature, we hypothesized that subjects who develop exceptional skills in L2 pronunciation take advantage of enhanced rehearsal abilities, and/or improve their phonological store’s capacity as an effects of the linguistic training.

Subjects with low L2 proficiency (N = 23) showed a digit span of 6.04 ± 1.18 (mean ± SD) numbers during forward repetition, 4.15 ± 0.93 (mean ± SD) numbers during backward repetition, and a word span of 5.1 ± 0.61 (mean ± SD) words during non-word repetition. Subjects with high proficiency (N = 23) were able to repeat 6.7 ± 1.26 (mean ± SD) numbers forward, 4.89 ± 1 (mean ± SD) backward, and 5.55 ± 0.82 (mean ± SD) nonwords. Consistent with our expectations, we observed a correlation between L2 proficiency and span capacity during forward (one-tailed linear regression, $r = 0.253$, $P = 0.036^*$, N = 69) repetition. Backward (one-tailed linear regression, $r = 0.079$, $P = 0.518$, N = 69) and non-word repetition (one-tailed linear regression, $r = 0.193$, $P = 0.112$, N = 69) failed to be significant.

A correlation of proficiency and forward repetition is consistent with our expectations and coherent with the findings previously reported. The development of exceptional L2 pronunciation skills, even when it starts after puberty, is facilitated and/or favours WM capacities. Papagno and Vallar (1995) studied polyglots’ cognitive abilities, and observed an enhanced auditory digit span and higher performances during non-word repetition. Palladino and Ferrari (2008) investigated cognitive skills in children who exhibited difficulties learning a foreign language in the absence of any cognitive impairment. The researchers examined phonological encoding abilities and documented a poor word span in these children. Consistent with these findings, our results confirm a close relationship between phonological WM and the capacity to acquire foreign languages. Our preliminary investigation furthers previous evidence by showing that even in adulthood individuals who possess high WM capacities are likely to encounter less difficulties when learning L2. Training WM since childhood could be an indirect and effective way to develop high performances in other domains as well. Analogously, speaking multiple languages presumably affects WM span. Future longitudinal research should study the expansion of WM capacities over time, and systematically address the effect of linguistic competences on it. On the basis
of our preliminary study it is not possible to rule out the existence of a correlation between proficiency and backward/non-word repetition abilities. A bigger corpus of data is needed to further investigate this topic.

5. Intelligence and L2 phonetic abilities

Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought.

Neisser, 1998

The idea that intelligent behaviour reaches its highest expression in human beings has been claimed by philosophers of all times. Intelligence and intelligent behaviour are individually distinctive traits that greatly vary across individuals. Intellectual performance is context and situation-dependent, and shows a great variability over time even for the same subject. While the concept of intelligence possesses a very intuitive meaning, its scientific definition is challenging. Theorizations on the nature of intelligence have attempted to clarify this topic. The need for objective measures of intelligence first emerged at the end of the 19th century, as psychologist Alfred Binet (1857–1911) was asked to devise a system to assess cognitive abilities and intellectual skills in children. This instrument was needed to identify children with mental retardation who would need to be assigned to special school classes.

Different perspectives exist with respect to the nature of intelligence and the ways it should be addressed. Some researchers suggest that a unified general ability is at the core of it. Other researchers believe that a wide variety of skills and aptitudes should be considered essential attributes of intelligence and that sets of completely independent intelligences can be identified.

Psychologist Charles Spearman (1863–1945) proposed a theory on intelligence that has become very influential. While examining the results of different intelligence tests that investigated cognitive skills, Spearman observed that the scores on these tests positively correlated. People who rated high scores on one cognitive test were likely to perform well on other tests, and vice versa for those individuals with poor performances.
On the basis of these studies, Spearman suggested that a general factor, termed \(g\), influences intelligent performance and that various cognitive tasks picture different aspects of general intelligence. Spearman proposed that intelligence is a general cognitive ability that can be experimentally measured, and which can be conceptualized as the ability to think, analyze situations and solve problems (Spearman 1904). He proposed that specific skills (termed factors \(s\)), along with \(g\), are at work during the execution of specific tasks (Spearman 1927).

Thurstone (1938) was one of the firsts to propose the idea that intelligence is not a general ability, but that it consists of a number of primary mental abilities such as verbal comprehension, word fluency, inductive reasoning, associative memory, perceptual correctness, spatial visualization, and numerical ability.

5.1. Gardner’s theory of Multiple Intelligences

Gardner criticized the psychometric approach to intelligence initiated by Spearman, pointing out that such a traditional approach can address only certain human abilities, while it leaves out a wide variety of important skills. People are naturally predisposed to develop some skills instead of others, and they encounter less difficulty when they learn certain types of information as opposed to others (e.g. verbal versus visual information).

Gardner’s conceptualization (1983) is known as the theory of multiple intelligences. He identified core abilities that are essential to coping with everyday life situations and described seven types of intelligence: verbal-linguistic, logical-mathematical, visual-spatial, interpersonal, body-related or kinaesthetic, intrapersonal, musical. In 1999, he added a further type to this list, the naturalistic intelligence. He proposed that these different domains are independent and result from a complementary but independent evolutionary course.

Bodily-kinaesthetic intelligence concerns the ability to dedicate oneself to physical activity such as dance, craft making, and sports. The learning process works according to the learning by doing motto. People with enhanced bodily-kinaesthetic intelligence enjoy physical activity and building things. Extroversion and sensitivity to other people’s needs and feelings and a cooperative and respectful attitude are characteristics of interpersonal intelligence. Successful communication and empathetic skills are among those abilities that usually make individuals popular. These people typically learn best by working with others and often enjoy discussions and debates. Intrapersonal intelligence fosters introspection
and the development of the skills of self-awareness and reflection. Verbal-linguistic intelligence together with logical-mathematical intelligence have always been more at the centre of the attention and concerns of scholars. Verbal intelligence relies on the ability to use language for an effective communication. People with verbal-linguistic intelligence master syntax and the structure of language very well and have fewer difficulties learning new words and foreign languages. They typically enjoy explaining or teaching. Logical-mathematical intelligence relies on logic reasoning and abstraction, such as is necessary for mathematics and computer programming. Spatial intelligence deals with the visualization and manipulation of spatial information. It is often associated with good visual memory and artistic talent. Individuals with musical talent have a good sense for music, tones and rhythm. They are likely to enjoy playing instruments. Naturalistic intelligence deals with nature and nurturing. Greater sensitivity to nature and naturalistic beauties are among the characteristics of this type of intelligence, along with the interest for animal species and their environment, the passion and the care for growing and nurturing living beings.

Gardner’s theory suggests that the adoption of a unified educational method for all students is not the most suitable approach. Teachers should take advantage of the natural predispositions of the students and devote greater attention to those domains in which each student is weak. This theory has received some criticism both from the psychology and educational communities. Some critics argue that this approach is based on intuitions rather than experimental data, and that beyond those intelligences it is possible to identify traits of personality. Despite some scepticism, the perspective sketched by this theory has gained popularity over the last years. An important aspect of this approach is that it postulates that linguistic/verbal intelligence might constitute separate sets of skills that might develop independently, and which might be dissociated from general intelligence. We considered this perspective very interesting and decided to test it in our study. We used the Raven’s Advanced Progressive Matrices (Raven 1938) to measure the general intelligence of our volunteers, and the Mehrfachwahl-Wortschatz-Intelligenztest (VMLT, Lehrl 1999) to assess verbal intelligence.
5.2. Testing general and verbal intelligences

We used the Raven’s Progressive Matrices to assess nonverbal intelligence. Cultural factors such as social environment, status, schooling and occupation obviously influence the development of intellectual skills to a great extent. For instance, Gay and Cole (1967) showed that rice farmers in Liberia develop a good sense for estimation of rice quantities. Young adults in Botswana, who are continuously exposed to story telling from their early childhood, achieve good recall abilities for stories (Dube 1982).

In an analogous way, enhanced intellectual performances are likely to be observed among people with a high occupational status and a high educational level. Kohn and Schooler (1973) evaluated through questionnaires a large sample of men with varying occupational profiles (managers, farmers, etc.). They observed that more demanding jobs on a cognitive level correlated with higher levels of intellectual flexibility. The question whether the cognitive efforts imposed by those jobs foster the development of intellectual skills or whether, on the contrary, pre-existing abilities made people able to engage in such occupations remains unanswered. Probably, a reciprocal influence of the two factors explains this process. Nevertheless, there is no doubt that enrolling in scholastic education trains intellectual skills such as reasoning, problem solving and attention, and that those skills develop to a different extent in individuals.

The Raven’s Progressive Matrices make up a culture-free test devised to measure general intelligence irrespective of nationality, education, age, and sex. This test is widely used and has become a standard measure for reasoning and abstract thinking. It was originally developed by Raven (1938) to measure the g factor (Spearman 1904). The Raven’s Progressive Matrices are constituted of matrices made up of geometric patterns with a level of progressively increasing difficulty. For each of these matrices, a geometric part of the series is missing. To complete the missing part of the pattern, the candidate needs to perform a multiple-choice process (i.e. he needs to identify the correct pattern among multiple alternatives). This task can be performed only if the participant is able to carefully observe the alternatives and make sense of the perceptual complexity of the matrix to complete.

The Advanced Progressive Matrices is a version developed for adults and adolescents with skills above average who were likely to continue with further education to graduate level. It contains 48 matrices, divided into two sets of 12, and 36 items. As the Standard Progressive Matrices, the level of difficulty progressively increases from the first to the last item.
We used the Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B) to test verbal intelligence in our subjects. This German test is constituted of 37 strings of words. Each list contains 5 words, 4 of which are nonsense words. The task of the subject consists in identifying the correct word (one for each string). The level of difficulty of this choice increases from the first to the last string. Subjects who possess a rich vocabulary encounter less difficulties in this task.

We tested 60 participants on both the non-verbal IQ (Raven) and verbal IQ (MWT-B) and looked for correlations between these IQ measures and the subjects’ scores on the linguistic measures: performance (pure pronunciation proficiency, M. Jilka) and the “talent” scores. Here we measured three independent “talent” indicators. 1. the pronunciation talent rating by M Jilka (see Chapter 2); 2. the MLAT (Modern Language Aptitude Test) with subtests (phonetic coding, grammatical sensitivity and vocabulary learning); and 3. an immediate imitation task of an unknown language (imitating Hindi words which were rated by 5 native speakers of Hindi).

The results showed no significant correlations of either the verbal or the non-verbal IQ with our above mentioned measures 1 (the “pronunciation talent score”) and 3 (the Hindi imitation capacity), but it showed significant correlations with some of the MLAT scores and subscores. Thus, non-verbal IQ measured by the Raven Advanced Matrices was significantly correlated to the overall MLAT score (one tailed linear regression, $r = 0.44, P = 0.000^{**}, N = 66$) and all three subscores of the MLAT MLAT3 – phonetic coding ability ($r = 0.36, P = 0.000^{**}, N = 66$); MLAT4 – grammatical sensitivity ($r = 0.37, P = 0.000^{**}, N = 66$) and MLAT5 Vocabulary learning ($r = 0.28, P = 0.003^{**}, N = 66$). Verbal IQ was significantly correlated to the overall MLAT total score ($r = 0.26, P = 0.005^{**}, N = 66$) and its subscale phonetic coding ability, MLAT3 ($r = 0.37, P = 0.000^{**}, N = 66$). These are highly interesting results, because they confirmed our hypothesis that there should be no significant correlation between the aptitude or ability for foreign language pronunciation and IQ measures. However, there was a significant correlation between both intelligence measures and the scores on the Modern Language Aptitude Test (MLAT). This result, or effect, can be partly explained by the fact that the MLAT, like many other language aptitude test batteries also measure some form of intelligence (for more information on that see Chapter by Hu and Reiterer pp. 97 and Chapter by Nardo and Reiterer, pp. 213).
References

Amelang, M., Schwarz, G., Wegemund, A.

Arbuckle, J.L. AMOS users guide version 4.0. Chicago: Small Waters.

Atkins, W.B., and Baddeley, A.D.


Baddeley, A.

Baddeley, A.

Baddeley, A.D., and Hitch, G.J.

Barkley, R.

Batson, C.D.

Bench, C.J., Frith, C.D., Grasby, P.M., Friston, K.J., Paulesu, E., Frackowiak, R.S.J. Dolan, R.J.

Bialystok E.
2006 Effect of bilingualism and computer video game experience on the Simon task. Canadian journal of experimental psychology, 60, 68–79.

Bialystok, E., and Martin, M.M.

Bialystok, E., Craik, F.I.M., Klein, R., and Viswanathan, M.

Bialystok, E.
Bialystok, E., and Codd, J.

Brown, J.

Burgess, P.W., Gilbert, S.J., Dumontheil, I.

Catellanos, F.X., Sonuga-Barke, E.J.S., Milham, M.P., Tannock, R.

Coles, M.G.H., Gratton, G., Bashore, T.R., Eriksen, C.W., Donchin, E.

Cowan, N.

Craft, J.L., and Simon, J.R.

Davis, M.H.

de Jong, R., Wierda, M., Mulder, G., Mulder, L.J.

Dube, E.F.

Eimer, M.


Fitts, P.M., and Seeger, C.M.

Francis, W.

Fuster, J.M.

Gardner, H.

Gathercole, S.E., Pickering, S.J., Knight, C., Stegmann, Z.

Gathercole, S.E., and Adams, A.M.

Gathercole, S.E., and Baddeley, A.D.

Gollan, T.H., and Kroll, J.F.

Gratton, G., Coles, M.G., Sirevaag, E.J., Eriksen, C.W., and Donchin, E.

Guiora, A.Z., Acton, W.R., Erard, R., & Strickland, F.W.

Guiora, A., Paluszny, M., Beit-Hallahmi, B., Catford, J., Cooley, R. Dull C.

Guiora, A., Beit-Hallahmi, B., Brannon, R., Dull, C., Scovel, T.

Guiora, A., Brannon, R., Dull, C.

Guiora, A., Lane, H., Bosworth, L.
Gay, J., and Cole, M.  

Håkansson, J., and Montgomery, H.  

Hogan, R.  

Holmes, G.  

Hulme, C., Roodenrys, S., Brown, G., Mercer, R.  

Kohn, M., and Schooler, C.  

Kondo, H., Morishita, M., Osaka, N., Osaka, M., Fukuyama, H., Shibasaki, H.  

Kondo, H., Osaka, N., Osaka, M.  

Kris, E.  

Kroll, J.F., and Dijkstra, A.  

Leibetseder, M., Laireiter, A.R., Koeller, T.  

Leibetseder, M., Laireiter, A.R., Köller, T., Leibetseder, M., Laireiter, A.R., Riepler, A., Köller, T.  

Lehrl, S.  

Lu, C.H., and Proctor, R.W.  
Martin, M.M., and Bialystok, E.

Martinussen, R., and Tannock, R.

Martinussen, R., Hayden, J., Hogg-Johnson, S., Tannock, R.


Mehrabian, A., and Epstein, N.

Merzenich, M.M., and Jenkins, W.M.

Miller, G.A.
1956 The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Milner, B.

Murray, D.J.

Peal, E., and Lambert, W.E.

Palladino, P., Ferrari, M.

Papagno, C., and Vallar, G.

Papagno, C., Valentine, T., Baddeley, A.D.
Paulesu, E., Frith, C.D., Frackowiak, R.S.J.


Raven, J.C.

Ridderinkhof, K.R.

Rosenthal, R.

Rosenzweig, M.R., Bennett, E.L.


Salovey, P., Bedell, B., Detweiler, J.B., Mayer, J.D.

Schneider, W., Eschman, A., Zuccotto, A

Schumann, J.H., Holroyd, J., Campbell, R.N., Ward, F.A.

Service, E.

Shapiro, J., and Rucker, L.

Simon, J.R., and Berbaum, K.
Simon, J.R.

Simon, J.R., and Rudell, A.P.

Smith, E., Jonides, J., and Koeppe, R.A.
1996 Dissociating verbal and spatial working memory using PET. Cerebral Cortex, 6, 11–20.

Spearman, C.
1927 The abilities of man: Their nature and measurement. London. Macmillan.

Spearman, C.

Stotland, E., Kenneth, E., Mathews, J., Sherman, S.E., Hansson, R.O., Richardson, B.

Stuss, D.T., and Alexander, M.P.

Sundman, Marketta.

Swanson, H.L.

Swanson, H.L., Ashbaker, M.H., Lee, C.

Taylor, L., Guiora, A., Catford, J., Lane, L.
1970 Psychological variables and ability to pronounce a second language. Language and Speech, 14, 146–57.

Thurstone, L.L.
1938 Primary mental abilities, Chicago IL, University of Chicago Press.

Vallar, G., DiBetta, A.M., Silveri, M.C.

Warrington, E.J., Logue, V., Pratt, R.T.C.

Wechsler, D.
1939 The measurement of adult intelligence. Baltimore: Williams & Wilkins, 229.
Wiegand, K., and Wascher, E.
Personality and pronunciation talent in second language acquisition

Xiaochen Hu and Susanne Maria Reiterer

You, the language learner are the most important factor in the language learning process

– Rubin and Thompson

1. Introduction

One central claim in second language acquisition (SLA) research is that there is a specific talent for learning foreign languages, which exhibits considerable variation between learners (Dörnyei & Skehan 2003). The knowledge of language aptitude or talent is very important for language teachers, as the wise ancient Chinese man and educator Confucius said, a good teacher should be able to teach students in accordance with their aptitude ("Yin Cai Shi Jiao"). However, it is not clear, whether this specific talent for SLA is merely language specific in nature, or it is also influenced by some cognitive psychological constructs, such as intelligence, working memory, mental flexibility, etc. The impact of these psychological factors on pronunciation talent has been discussed in Chapter 3 (by Rota & Reiterer). Besides that, personality has caught the attention of psycholinguistic researchers, because it is perceived to be a stable, enduring, important aspect of the individual, which might have a major influence on the SLA behaviour. Early research has addressed the impact of personality factors such as extraversion, risk taking, lack of inhibition and self-esteem (for an overview, see MacIntyre & Noels 1994).

Language, after all, belongs to a person’s whole social being: it is part of one’s identity, and is used to convey this identity to other people. The learning of a foreign language involves far more than simply learning skills; it involves an adoption of new social and cultural behaviours and ways of being (Williams 1994). Moreover, if the “typical personality” of one culture is more introverted than that of the second culture, what
effect might this have on the self concept and persona of individuals speaking the two languages and participating in the two cultures? Do they perceive themselves as different? Do others? If yes, in what way?

As described above, one might have some intuitive ideas about the relationship between personality and SLA; but the scientific investigation of this relationship needs a more precise formulation of question. Thanks to the development of both personality and linguistic research, the impact of personality on SLA behaviour can be discussed from different aspects. For instance, extraverts are assumed to be successful sociable learners, who should generate more output, maximize contact and emphasize using the language. Or we can assume, individuals of higher openness are supposed to be more open to the foreign social and cultural behaviours or ways of being and will therefore be more motivated and successful in SLA behaviour. To test these hypotheses, a myriad of studies have used correlation approaches despite their shortcomings – a correlation approach cannot reveal any causal relationship, to test the strength of the link between a certain personality variable and a certain linguistic variable. Only a few studies have manipulated the former variable experimentally to examine the causal relationship between personality and language aspects (e.g. Guiora 1967; Guiora et al. 1972).

It is not easy to carry out studies in the research field of personality and SLA. This is partly because it is hard to find an appropriate level of analysis for both of the personality and the language variables (Furnham 1990). Theoreticians in the field of personality research wanted to explain linguistic behaviour at a global level and did not wish to analyze linguistic subsystems in detail. Psycholinguists and sociolinguists, on the other hand, were confused by the multiplicity of theories in the field of personality research. In addition, the complexity of selecting the appropriate measurement approach and instruments for personality and speech has been a deterrent for both linguists and psychologists, and so did the fact that various combinations of the selected measures often produced mixed results, making the interpretation of the findings difficult (Dörnyei 2005).

Therefore, it is important to clarify the subcomponents/concepts in both personality psychology and language aptitude. This effort should be made from both disciplines. For example, extraversion has received most of the attention, but the emerging picture about the relationship between SLA and extraversion was rather negative. Dewaele and Furnham (1999) explained that this negative image could be due to not having distinguished properly between written and oral language criteria. They
argued that the correlations were mostly significant in oral language, but not in written language. The current chapter tries to give an overview from another point of view: classifying the studies by different research perspectives of personality. Readers will find out that the research in the field of personality and SLA has been only focused on a few perspectives of personality (e.g. trait approach). Then the current study design and the first-stage behavioural results of the ongoing study will be presented.

2. How to define Personality

Personality (with its original meaning “mask” from the Latin word *persona*) has been filled with different elucidations throughout the history of psychological research. Gordon Allport, a prominent early figure in personality psychology, popularized the concept “personality” and provided the definition as ‘a dynamic organisation, inside the person, of psychophysical systems that create the person’s characteristic patterns of behaviour, thoughts and feelings’ (Allport 1961). The field of personality psychology has been long beset by a chaotic plethora of personality constructs, like a “Tower of Babel”, that sometimes differ in label while measuring nearly the same thing, and sometimes have the same label while measuring very different things (Funder 2001). In other words, it is like a typical scenario of what Ludwig Wittgenstein would have called “language confusion” (Sprachverwirrung). Therefore, the precise understanding of personality should be connected to its theoretical context and the relevant research approaches (how theoretical constructs are manifested). Interestingly, the complexity of the “Tower of Babel” even caused the difficulties of classification of the research approaches. However, most theorists of personality agreed on at least the following five perspectives: psychoanalytic, learning/behaviourist, dispositional, cognitive, and biological perspectives (Asendorpf 2007; Carver & Scheier 2004; Funder 2001; Larsen & Buss 2007; Maltby et al. 2007).

**Psychoanalytic perspective** originated in the writings of Sigmund Freud. The thematic focus of this perspective, which gave rise to the term psychodynamic, is the idea that personality is a dynamic set of processes, always in motion. Freud developed a topographical model of mind, including three regions of mind: conscious, preconscious, and unconscious; and a structural model of personality, which is composed of ego, superego, and id.
Learning, or behaviourist perspective of personality has sparked something of a philosophical argument, whether behaviour is “controlled” or not. Is behaviour controlled from within the person, or is it controlled by events and processes outside the person? Early learning theory had the creed that behaviour is merely a function of external stimuli and environmentally imposed reinforcement contingencies; these early models did not care about what happened within the “black box” – or the person. The representing theories are from classical conditioning to instrumental conditioning. Recent theorists have backed away from this assumption and paid increasing attention to the self-system, e.g. Rotter (1990) with his “Locus of Control Expectancies” and Bandura (1999) with his “Social Cognitive Theory of Personality”.

Dispositional perspective of personality contains two major themes: one is that people display consistency or continuity in their actions, thoughts, and feelings; the other is that people differ from each other in many ways. Each person’s personality consists of a pattern of dispositional qualities, and the composition of the pattern differs from one person to another. The “Tower of Babel” phenomenon is embodied here, obviously. The same label can have different meanings. The most important issue within the dispositional perspective is to differentiate a personality trait and a personality type (Figure 1). In typology, types are usually regarded as categories that are distinct and discontinuous. In contrast, discussions of traits usually assume that people differ along continuous variables or dimensions.

Trait approach has dominated personality psychology for a long period and still has important influence on both personality research and applied areas (e.g. linguistics). One of the key tools for developing a personality inventory from the trait perspective is factor analysis. Two important models, the Cattell’s sixteen factors model (Cattell et al. 1977) and the modern Five-Factor Model (e.g. Costa & McCrae 1985, 1992), were derived based on the lexical approach. It is the factor analysis of whole spectra of words for describing personality in the natural human language. However, the Theory of Eysenck, the three dimensional construct (extraversion, neuroticism, psychoticism), was developed based on the typology of Hippocrates and Galen and related observations by Jung and Wundt (Eysenck 1967). Another important issue from the trait perspective is the “person-situation” debate. This issue will be addressed in the following paragraph when we discuss anxiety.

Cognitive perspective regards personality as an information processing system, which is based on three different sources, including architecture,
parameters of processing, and knowledge (Asendorpf 2007). The architecture of this system is supposed to be universal for human beings. The parameters of processing can be found in cognitive, emotional and motive areas: in the cognitive area, such as the processing speed, threshold and capacity of the cognitive components (perceiving, remembering, interpreting, anticipating, and believing); in the emotional area, such as different temperament (affect, arousal level, and attention). Knowledge can be differentiated into explicit and implicit, or declarative and procedural, knowledge.

**Biological perspective** embraces the key notion that the basis of all psychological phenomena lies in the neural-biological system. Just a few years ago it became possible to regard research in neuroanatomy and physiology as relevant to personality (Canli 2004; Gray et al. 2005; Wright et al. 2007), with the help of newly developed brain imaging methods. Several personality theories have been tested within the biological perspective so far. Other possibilities of testing include animal studies and the behavioural genetics approach (e.g. twin studies).

Most of the existing studies on personality and SLA aptitude used the theoretical framework of the dispositional (trait and type) perspective. Only a few studies made use of the other perspectives. This phenomenon
probably has its cause in the limited investigation tools available on the shelf. As we can see, the dispositional approach has been producing countless questionnaires for assessing different personality traits and types, which can easily be implemented in other research settings. Dörnyei (2003, 2005) claimed that individual differences research is inextricably linked to psychometrics, with the issue of questionnaire design being at the forefront. Nevertheless, this opinion seems strict. We are currently situated in a new research era of brain imaging, which provides the chance to directly map brain anatomy and activities onto psychological phenomena. We believe that this rapid development will have abundant attributions in manner of thoughts and theory constructions of behavioural and cognition research (Coltheart & Friston 2008). Future research on the relationship between personality and language aptitude will definitely benefit from these recent developments. Therefore, we incorporated the biological viewpoint into our own research about the interplay between personality and pronunciation aptitude, as later sections of this chapter will show.

3. Psychoanalytic perspective of personality and SLA

The seminal work of establishing a link between the psychoanalytic perspective of personality and SLA behaviour (especially pronunciation aptitude) is the work of Guiora (1979), in which the concept of permeability of the language ego boundary was developed from certain psychoanalytic ego constructs. The permeability of language ego boundary does not equal to a “weak ego”, but suggests the ability to move back and forth between languages and the “personalities” that seem to come with them. This is related to the common observation that one “feels like a different person” when speaking a second language and often acts very differently as well. The early flexibility of ego boundaries is reflected in the ease of assimilating native-like pronunciation by young children. Here, pronunciation is viewed as the most important contribution of language ego to self-representation.

According to Guiora, the permeability of ego boundaries was supposed to be the root for both: the pronunciation aspect of the SLA behaviour and the personality’s empathetic capacity. Here, empathetic capacity is defined as individuals’ sensitivity to the feelings of others (Guiora 1967). Studies have been conducted using the so-called Micromental Expression (MME) paradigm, which is based on the proposition
that a subject’s ability to identify changes in facial expression at various speeds taps his sensitivity to the affective states of another, or his empathetic capacity. The initial two studies (Guiora 1967; Taylor et al. 1969) with MME paradigm seemed to directly support their hypothesis. However, the reliability of MME has been questioned afterwards, which is thought to be merely a noisy measure of perceptual ability or the ability of an individual to tolerate anxiety and anticipation, not necessarily a measure of empathetic capacity (Taylor et al. 1969).

Therefore, Guiora and colleagues (1972) turned to the well-known “alcohol study”. The purpose of the “alcohol study” was to examine the effects of a small amount of alcohol on the ability of college students to pronounce words and phrases in a foreign and totally unknown language (Thai). The result was that the group of students with alcohol did pronounce better than the control group (with placebo). This indicated that the experimentally induced lowering of inhibitions or enhancing of the permeability of ego boundaries will lead to a corresponding facilitation of the pronunciation flexibility. Nevertheless, Brown (1973) suggested that the effect of alcohol might just have relaxed the muscles. Additionally, Schumann and colleagues (1978) carried out an interesting study investigating the pronunciation ability of subjects under three conditions: baseline, hypnosis, and post-hypnosis. Thai was again used as the target language. The result, which seemed to support Guiora’s theory, indicated that deeply hypnotized subjects performed significantly better than the less hypnotized one.

Another interesting but still quite a philosophical aspect of the psychoanalytic perspective is the work of Granger (2004). She tried to use psychoanalytic framework of “identity” and “conflict” to explain the silence in the SLA process. Silence in the SLA process usually reflects an inability to express oneself because of an assumed lack of competence in the foreign/second language. However, according to Granger, “Something significant is indeed happening, even when learners are silent”. Second language learners need a period during which they reinterpret the world and their place within it through the structures and functions of the new language. This suggests that there is much more to silence than the absence of speech; silence in some second language learners is symptomatic of “the loss, ambivalence, and conflict that accompany a translation between two languages, a physical suspension between two selves”. Therefore, the experience of silence is the interrelated process of identity construction and SLA.
4. Learning perspective of personality and SLA

Although the learning perspective has been dealing with the “learning process”, the literature on SLA and personality from this perspective is fairly sparse. Two aspects have still caught the attention of researchers, that is, locus of control expectancy (Rotter 1990) and self-efficacy (Bandura 1999). Nevertheless, they were often described as the motivation factor by linguists. For example, self-confidence, a concept similar to self-efficacy, is understood as a component of motivation (Clément et al. 1977). From the psychological point of view, however, these are rather psychological constructs that could influence motivation, not the motivation per se.

Locus of control expectancy indicates the internal or external attribution of the success or failure. This aspect has been often addressed in the literature on SLA, while empirical studies were rarely done. Williams and Burden (1999) has conducted a qualitative research on school children's attribution of success and failure of learning a second language. They found that the “motivation enhancing attribution” is important for learning. Cautiously, the qualitative approach they used, which is based on the grounded theory (Strauss & Corbin 1990), is known to be efficient in theory construction, whereas more quantitative evidence is needed for verifying theories.

Self-efficacy is the central concept of the social cognitive theory of personality (Bandura 1999). It refers to people's beliefs in their capabilities to perform in ways that give them some control over events that affect their lives. Some researchers have conceptualized a generalized sense of self-efficacy. Meanwhile, self-efficacy is often understood as being domain-specific or even task-specific; that is, one can have more or less different self-beliefs in different domains or particular situations of functioning. Accordingly, language learning specific self-efficacy has been already constructed in the past studies, and was found to be a good predictor for the SLA behaviour. For instance, Wong (2005) found a significant correlation between the language specific self-efficacy and the self-rating of importance of learning strategies ($r = 0.72$) in pre-service teachers of English in Malaysia. So far, few studies in the field have been done from the learning perspective. We suggest that more studies should be carried out by relating SLA behaviour or its subcomponents to a more general personality construct (e.g. general self-efficacy).
5. Dispositional perspective of personality and SLA

5.1. Personality types

The “Mayer-Briggs Type Indicator” (MBTI; Myers and Briggs 1976), is currently the most often used personality type inventory in the SLA field (Dörnyei 2005). It originates from Carl Gustav Jung’s theory of personality type and measures personality preference on four scales: Extraversion – Introversion (E – I), Sensing – Intuition (S – N), Thinking – Feeling (T – F), and Judgment – Perception (J – P).

Some researchers have linked personality type to learning style (Bailey et al. 2000; Lawrence 1997). Therefore, the personality types are supposed to be able to predict the SLA achievement (Bailey et al. 2000). However, empirical studies using MBTI have produced mostly weak or mixed results. For instance, Ehrman and Oxford (1995) investigated a large sample of 855 Foreign Service Institute students and obtained several significant but weak correlations between the personality types and reading and speaking proficiency measure (around .20). Similarly, Carrell and colleagues (1996) have only found some trend of correlation between the MBTI scores and a grammar measure in 76 students of English. Although the MBTI is a standard tool for assessing personality types, however, the researchers of SLA field did not find it an appropriate indicator in language related issues. The reason for introducing MBTI into SLA research is that it can be linked to cognitive style, as well as learning style (Bailey et al. 2000; Dörnyei 2005). The negative results in the past studies indicated that, more specific theoretical link (not just stating that personality type is related to cognitive style) need to be proposed.

5.2. Personality traits: Extraversion – Introversion

Extraversion – Introversion as personality traits have been investigated intensively in the SLA field and researchers have found contradictory results. Deweale and Furnham (1999) concluded that the differentiation between oral and written forms of materials for assessing linguistic ability can solve this problem: linguistic variables extracted from oral form tended to find relationships with the degree of extraversion of the speakers.

Most early studies were based on Eysenck’s personality theory (EPT), which was measured by either Eysenck Personality Inventory (EPI: Eysenck & Eysenck 1964), or Eysenck Personality Questionnaire (EPQ: Eysenck et al. 1975). This predominance might partly be explained by the
test’s simplicity of administration. As one of the early researchers stated honestly, “The EPI was chosen because it is easy to administer (it requires only ten minutes) and is easy to score” (Busch 1982). Few studies were based on other scales, such as Freiburger Personality Inventory (FPI: Fahrenberg & Selg 1970; e.g. in Vogel & Vogel 1986), Early School Personality Questionnaire (ESPQ: Coan & Cattell 1966; e.g. in Strong 1983) and California Children Q-Set (CCQ: cf. van Lieshout & Haselager 1992; e.g. in Verhoeven & Vermeer 2002).

Dewaele and Furnham (1999) examined the relationship between extraversion (using EPI) and second language in oral speech situations extensively and revealed that extraverts had higher fluency (not accuracy). This effect can be found in both formal and informal situations. It seems to be more significant if situation complexity is increased. The Vogel & Vogel’s FPI study (1986) also revealed a strong relationship between extraversion and backchannel behaviour in a film/retelling task. Interestingly, the construction of the trait extraversion and emotionality in FPI was based on EPT.

The Strong’s study (1983) using ESPQ, however, failed to find a significant relationship between extraversion and fluency. Noteworthy is that ESPQ is constructed by factor analysis for assessing children’s pre-disposition. The trait extraversion was a second-order factor which composed of three first-order factors: Cyclothymia–Schizothymia, Surgency–Desurgency and Parmia–Threctia (Cattell 1959). Therefore, the label extraversion here is quite different from that of EPT. This case tells us, how important it is for researchers, to understand the exact meaning of the same label in personality research.

Another early EPI study by Busch (1982) should be mentioned here. Busch investigated a group of Japanese college students in their English proficiency both through a standardized English test and an oral interview. Only the pronunciation subcomponent of oral interviews was negatively correlated with extraversion ($r = -0.38$). This negative result seemed paradoxical here, but we have to point out that in this study the inter-rater coefficient for pronunciation was low ($r = 0.54$). This methodological issue has not been emphasized in the early studies. We strongly suggest that researchers ought to be cautious when interpreting the results of studies using single-rater measurements for linguistic abilities, which could raise questions of reliability and validity.

Recently, a few studies have started to use personality questionnaires which are based on the Five-Factor Model (extraversion, neuroticism, openness, conscientiousness, and agreeableness) to investigate the relation-
ship between personality and SLA behaviour. Verhoeven & Vermeer (2002) tested a group of teenage language learners on their communication competence (with three subcomponents) and their personality traits using CCQ. The results showed that extraversion was highly correlated with the strategic competence ($r = .51$); conscientiousness was slightly correlated with organizational competence ($r = .28$); and, more importantly, openness correlated with all the three subcomponents ($r = .51$ for organizational competence; $r = .29$ for pragmatic competence; $r = .48$ for strategic competence). In this study, the influence of extraversion was again detected. Interestingly, the other traits, such as openness, were also found to relate to the communication competence. As mentioned previously, the Five-Factor Model is derived by factor analysis, which has received agreement by most researchers in the field. The trend of using personality instrument based on the Five-Factor Model is promising, since it allows researchers to investigate effects of all the five most important traits at the same time.

5.3. Anxiety

Anxiety has been regarded as one of the best predictors of success in learning a second language (Gardner 1985). The vast majority of theoretical findings converge with language teachers’ and learners’ experience that language anxiety influences language behaviour negatively (Dörnyei 2005). However, the definition of anxiety has been rather inconsistent in the body of SLA research. According to MacIntyre and Gardner (1991), the volumes of research on anxiety and SLA behaviour can be divided into three kinds: trait anxiety, state anxiety and situation specific language anxiety. Accordingly, in the research body of the trait perspective of personality, the “person-situation” debate has also been one of the central themes (Funder 2001). Therefore, research on the relationship between trait and SLA behaviour will definitely be beneficial for both – the SLA field and the personality field. Trait aspect considers anxiety as a general personality trait of a person that is relevant across several different situations. State aspect is interested in “here-and-now” experience of anxiety as an emotional state. Early psychologists have already differentiated trait and state concepts (Levitt 1980) and have developed amounts of scales for measuring trait and state anxiety (Spielberger 1983; Taylor 1953; Zuckerman 1960). However, the research results obtained by using these standardized tests of anxiety trait or state appeared to be rather disappointing. Their inability to capture the essence of foreign language
anxiety or to satisfactorily demonstrate a role for anxiety in the language learning process seems to be leading research toward the situation specific approach.

The situation specific approach examines the specific forms of anxiety that occur consistently over time within a given situation. MacIntyre (1999) defines situation specific language anxiety as involving the “worry and negative reaction arousal when learning or using a second language”. In the past decades, several questionnaires have been developed, including French Class Anxiety Scale (Gardner & Smythe 1975), English Use Anxiety (Clément et al. 1977), English Test Anxiety (Clément et al. 1980), Foreign Language Classroom Anxiety Scale (FLCAS: Horwitz et al. 1986), Second Language Speaking Anxiety Scale (SLSAS: Woodrow 2006), etc. These questionnaires are successful in manifesting the foreign language related anxiety. Especially the FLCAS was found to have captured a relatively independent, uniquely second language related factor, which is different from general trait anxiety, or other situation specific anxiety (e.g. test anxiety, communication anxiety; Horwitz 2001).

Several conclusions can be drawn from the body of language anxiety literature (for details, see MacIntyre & Gardner 1991): language anxiety negatively effects performance in the second language; listening and speaking are the most anxiety-provoking aspects in SLA situations; language anxiety is more relevant to language learning among adults than among children; the causality between language anxiety and SLA achievement is not clear, whether anxiety is a cause or an effect of bad achievement.

Situation specific language anxiety does explain some SLA phenomena to some degree. However, other investigators found that language anxiety is not a stable “trait” among experienced language learners (Dewaele 2002). That means that some other factors could influence language anxiety, such as SLA context (Woodrow 2006), interpersonality relationship (Dörnyei & Kormos 2000), other personality factors, such as perfectionism (Gregersen & Horwitz 2002), and emotional intelligence (Deweale et al. 2008). This suggests that although the construct of situation specific language anxiety was welcomed by most researchers, it was still not the most accurate measure of language related anxiety. We propose that the next step ought to examine the interplay between language anxiety (or trait anxiety) and other personality factors.

Eysenck (1979) once reconceptualized anxiety in terms of cognitive interference. He suggested that anxious persons divide their attention between task-related and self-related cognition, making cognitive performance less efficient. Furthermore, anxious persons might however be aware
of this interference and attempt to compensate by increased effort. This explains why “the potential effects of language anxiety on cognitive processing in second language may be quite subtle” (MacIntyre & Gardner 1994). The question here is what causes the effort for compensation? Does it relate to motivation in language learning, or even to some other personality trait? We propose that the behavioural withdrawal-approach system (BIS/BAS), a biological personality concept proposed by Gray (1972, 1981, 1987b, 1990), might serve as the cause for the individual difference in making an effort, given the same degree of (language) anxiety. More about this aspect can be found in part 7 of this chapter.

6. Cognitive perspective of personality and SLA

Personality can be considered as an information processing pattern (Asendorpf 2007). Cognitive approaches to personality focus on differences in how people process information. There are at least three levels of cognition of interest to personality psychologists: perception, interpretation, beliefs and desires. The cognitive perspective has a broad framework, however, SLA studies seemed to have only captured a few points, such as cognitive styles, which are supposed to be related to one of the cognitive theory of personality – the field independence-dependence (FI/D) theory (Dörnyei 2005). FI/D theory is typically measured by the Embedded Figures Test (EFT). It reveals personality through perception (Larsen & Buss 2007). Field-independent people are more likely to analyze information into its component parts and to distinguish the essential from the inessential. Field dependents, in contrast, are more likely to deal with information structures as wholes, or “gestalts”. In relation to SLA, the former was supposed to analyze linguistic material, and perhaps learn systematically; the latter was thought to engage in communicative language use and to “talk to learn” (Dörnyei & Skehan 2003).

The empirical results showed there was low (around 0.3) or no correlation between FI/D theory and SLA achievement. Furthermore, the correlations are lowered when intelligence scores are partialed out (Chapelle & Green 1992). This could be because most researchers have used EFT as a measure of the FI/D, while EFT actually only measures the cognitive reconstructing ability, which is highly overlapping with measures of fluid intelligence, for instance, the Raven’s Matrices Test (McKenna 1983).

So far, the researches on FI/D theory and SLA remain ambiguous. Hopefully, future studies of better design will change this situation. For
example, as Chapelle and Green (1992) have argued, other components of the FI/D theory, such as the frame of reference (whether people rely on internal or external referents) and interpersonal competences (capacity to work effectively with other people), might still have separate research promise.

7. Biological/neuroscientific perspective of personality and current research

Human beings are biological creatures; therefore, biological indices can serve as the evidence of the inheritance and stableness (dispositions) of personality. Investigations of the biological basis of personality have led to the development of several influential models of personality (Canli 2004), such as those by Eysenck (1967, 1990), Gray (1982), Cloninger (Cloninger et al. 1993) and Zuckerman (1991). To the knowledge of the authors, none of the SLA-personality studies have adopted biological viewpoints. The current section focuses on introducing Eysenck’s biological model of extraversion (arousal) and Gray’s BIS/BAS model. Then neural imaging studies of personality are discussed.

7.1. Eysenck’s Biological Model of Extraversion

Eysenck (1967) suggested that extraverts and introverts are mainly differentiated by their basal arousal level. Extraverts have a lower arousal level than introverts do. And it is posited that the effects of arousal on cognitive performance conform to an inverted-U function (Yerkes & Dodson 1908), in which moderate levels of arousal have a beneficial effect on performance, whereas low or high levels have a detrimental effect, especially in complex cognitive tasks (Humphreys & Revelle 1984). The consequence is that introverts are expected to perform better than extraverts in low-arousal tasks but to be outperformed by extraverts in high arousal tasks. Recently, Lieberman and Rosenthal (2001) proposed a more contemporary version of the arousal account: the concept of arousal was linked to catecholamine activity, in particular, to dopamine and norepinephrine. Norepinephrine is more related to wakefulness and anxiety, and dopamine to positive affect and novelty detection.
7.2. Gray’s BIS/BAS Model

Gray (1981, 1982) has proposed another biological model of personality, which holds that two general motivational systems underlie behaviour and affect: a behaviour inhibition system (BIS) and a behaviour activation system (BAS). The aversive motivational system BIS comprises the septo-hippocampal system, its monoaminergic afferents from the brainstem, and its neocortical projection to the frontal lobe. Gray has argued that this physiological mechanism controls the experience of anxiety in response to anxiety-relevant cues (Gray 1972, 1977, 1978, 1982, 1987b, 1990). The core concept of BIS is anxiety. It is sensitive to the signal of punishment, nonreward, and novelty. It inhibits behaviour that may lead to negative or painful outcomes. The neural basis of appetitive motivational system BAS is so far less specific than that of BIS, though catecholaminergic, especially dopaminergic pathways are believed to play a central role (Stellar & Stellar 1985). This system is said to be sensitive to the signal of reward, nonpunishment, and escape from punishment. Activity in this system causes the person to begin (or to increase) movement toward goals. Obviously, the BIS/BAS dimensions are different from the widely accepted neuroticism/extraversion dimensions (for overview, see Carver et al. 2000), because the BIS/BAS dimensions reflect individual differences in two systems for action control (approach, withdrawal), which might be related to cognitive aspects of control (Gray & Braver 2002).

Since BAS and BIS represent distinct structures in the nervous system (being separable both pharmacologically and by brain lesion), their sensitivities are presumed to be orthogonal (Gray 1987a, 1987b; Quay 1993). We suggest that the orthogonal relationship of BIS/BAS personality might be also presented in the SLA behaviour. As mentioned before, language specific anxiety (related to BIS) plays an important role in SLA behaviour, while the effort that was made to compensate for the anxiety induced cognitive interference might be different across individuals (Eysenck 1979). Some researchers have the idea that one function of the prefrontal cortex may be to exert compensatory control over behaviour when additional demands are imposed (Braver et al. 1997; Bunge et al. 2000). Therefore, we suggest that, given the same degree of (language) anxiety, the individual differences in making effort (compensation) during SLA pronunciation tasks might be reflected by BAS.
7.3. Brain imaging methods to personality

By means of brain imaging techniques (e.g. functional magnetic resonance imaging, or fMRI), Canli and colleagues (2001) initiated the investigation on the moderating influence of personality on brain reactivity to emotional stimuli. They reported that in a sample of healthy women, extraversion was correlated with brain reactivity to positive stimuli in the regions of frontal cortex and left middle gyrus; whereas neuroticism was correlated with brain reactivity to negative stimuli at the same region. In a further fMRI study, Canli and colleagues found a positive correlation between extraversion scores and the reactivity to happy facial expressions in amygdala (Canli et al. 2002). More recently, studies using fMRI or other brain imaging techniques have also found significant correlations between brain activity (resting state or working state) and specific personality traits such as extraversion-introversion (Deckersbach et al. 2006; Eisenberger et al. 2005; Haas et al. 2006; Mobbs et al. 2005; O’Gorman et al. 2006; Vaidja et al. 2007), Neuroticism (Deckersbach et al. 2006; Eisenberger et al. 2005), psychoticism (O’Gorman et al. 2006), novelty seeking (Bermpohl et al. 2008), cheerfulness (Rapp et al. 2008). These results provide important evidence for the biological basis of extraversion or other personality traits and indicate that there are systematic individual differences in the patterns of brain activation in response to cognitive or affective stimuli, and sometimes in the brain resting state.

More importantly, individual differences according to BIS/BAS personality theory were also found to be able to modulate the brain activity. BIS/BAS dimensions can be measured with BIS–BAS questionnaire by Carver and White (1994). The following studies are all based on this scale.

Electroencephalogram (EEG) studies have shown that individual differences in BIS and BAS predict differences in prefrontal hemispheric asymmetries in resting brain activity (Coan & Allen 2003; Hewig et al. 2006; Harmon-Jones & Allen 1997; Sutton & Davidson 1997).

Using fMRI technique, Gray and Braver (2002) tested whether BIS/BAS dimensions could predict differences in working memory-related activation in the caudal anterior cingulate cortex (ACC) after preexposure to emotional films of positive or negative valence. Working memory (WM) load is known to modulate neural activity in the caudal ACC, a brain region critical for the cognitive control of behaviour. The results showed that the self-reported BAS had a trend to correlate with WM performance ($r = .27$), but it significantly correlated with WM-related acti-
vation in the caudal ACC ($r = -.84$). This suggests that neuroimaging is a more powerful tool to detect subtle relationships between variables than traditional behavioural methods. This might be explained by the fact that observed behaviour constitutes the sum of all neural activation patterns, whereas brain imaging can focus on a single brain structure, and may therefore pick up relationships that are subtle to be noticed at the behavioural level of analysis (Canli & Amin 2002).

Furthermore, Gray and colleagues (2005) replicated the above-mentioned study in a larger independent sample, by controlling fluid intelligence level and analyzing brain regions out of the ACC that also support cognitive control and working memory. The results indicated that higher BAS was negatively related to brain activity in the dorsal ACC, the lateral prefrontal cortex (LPFC), and parietal areas. The results suggested that higher BAS is related to less mental effort and indicated a greater neural-processing efficiency. Similarly, in the SLA domain, Reiterer and colleagues (2005) also found that subjects with higher L2 proficiency had lower brain activity measured by EEG, which indicated a higher neural-processing efficiency. Additionally, in our current research about the neural underpinnings of pronunciation ability we also found similarities of this principle: higher efficiency in phonetically talented individuals (for details, see Chapter 5 by Reiterer). In the current project, we expect to reveal the contribution of personality BAS to the neural-processing efficiency in second language pronunciation related brain areas, in order to provide a psychobiological explanation for the SLA pronunciation aptitude.

8. Short summary

To summarize, personality seems to play an important role in SLA behaviour. However, the relationship between personality and SLA is still not clear. As most linguists might have noticed, differentiation in forms of language materials (oral or written), differentiation between overall proficiency and specific aptitude, as well as differentiation in subcomponents of linguistic variables (phonetic, semantic, grammatical or pragmatic), are important for obtaining a clearer picture. In this review, we emphasized another aspect: one should make clear which personality one is talking about (see Part 2). Past research in this field has only focused on a few perspectives of personality, particularly, on the dispositional or trait approach. This phenomenon might be due to the easiness of admin-
isting a personality questionnaire. We can conclude from the trait perspective that extraversion (mostly based on Eysenck’s personality theory; see Part 5.2) and situation specific language anxiety (see Part 5.3) served as the important variables in SLA behaviour (especially in speaking situations); while the effects were relatively subtle.

Besides the fruitful studies from the trait approaches, studies from other perspectives have also yielded some interesting insights. From the psychoanalytic perspective, ego boundary or empathetic capacity (see Part 3) might be an important variable for the SLA aptitude, especially the pronunciation aptitude. From the learning perspective, self-efficacy might also play an important role (see Part 4). Unfortunately, studies from other viewpoints of personality are rather limited.

9. Current research design

It is a commonplace observation that people differ extremely in their phonetic-articulatory aspect of second language aptitude. The current study aimed to find out the influence of personality on the second language pronunciation talent by integrating several research approaches (dispositional, cognitive, and biological perspectives). From the dispositional point of view, personality is perceived to be a stable, enduring, important aspect of the individual, and it can be measured by self-report questionnaires on the behavioural level. From the cognitive perspective, personality is regarded as an information processing system. The parameters, which can regulate information processing efficiency, will determine the performance/behaviour. The cognitive viewpoint can be easily incorporated within the biological approach. We would consider the nervous system as a “container” of the information processing system, in which the information processing efficiency can be reflected by the physiological activity measured by brain imaging techniques. The parameters, which are supposed to regulate the information processing efficiency, will be thus related to the individual differences of neural-processing efficiency of the resting state or during a particular task. For example, the study of Gray and colleagues (2005) has shown the regulational effect of BAS personality on the brain reactivity to the high cognitive loaded working memory task.

The challenge of the current study lies in determining the parameters of the information process. As the cognitive model of personality indicates, the parameters can be found in cognitive, affective, and motive
areas. A careful selection of task-related parameters is very important. Based on the current review, our ongoing research has following assumptions:

1) Some personality traits (e.g. extraversion) are supposed to be correlated with pronunciation aptitude.

2) There might be an interplay between personality traits (e.g. interplay between anxiety and BAS as suggested in Part 7.2) or between traits and other cognitive parameters.

3) The relationships are presented either on the behavioural level, or on the brain activity level.

To verify the above-mentioned assumptions, three procedures have to be applied: finding out appropriate indices for pronunciation aptitude (see Chapter 2 by Jilka), carrying out pronunciation specific brain imaging experiments, and administering an exhaustive assessment of personality. To date, a group of 62 native German-speaking students (31 men and 31 women; all right handed; age = 26.0 ± 4.5; verbal IQ = 126.2 ± 11.2; non-verbal IQ = 130.1 ± 12.4) have already completed all the three procedures of the tasks (second language: L2 English, age of onset = 10.5 ± 1.2). The assessment of personality was accomplished by internet-based tools and by classical “on the spot” testing before or after subjects performed an fMRI experiment. The instruments for personality assessment are listed below (also see Table 1):

I. NEO-FFI

NEO-FFI (Costa & McCrae 1985, 1989, 1992) is a widely-used self-report measure of personality based on the Five-Factor Model. The current study used the German version, which is translated from the original American NEO-FFI scale by Borkenau and Ostendorf (1993). FFI refers to Five Factor Inventory. The five factors included neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. NEO refers to the first three factors. The test (questionnaire) was administered using the internet. Subjects needed to rate 60 statements on a 5-point scale from “strongly agree” to “strongly disagree”. The NEO-FFI was chosen because it is a widely accepted questionnaire of traits currently. Additionally, its two fundamental dimensions, neuroticism and extraversion, are comparable to the definition of Eysenck’s Personality Theory (EPT), which allows us to compare our results to the former traits studies on SLA behaviour.
<table>
<thead>
<tr>
<th>Test Name</th>
<th>Initiation</th>
<th>Administration</th>
<th>Variables</th>
<th>Short Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEO-Five-Factor Inventory</td>
<td>NEO-FFI</td>
<td>Internet Testing</td>
<td>Neuroticism</td>
<td>Tendency to experience distress</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Extraversion</td>
<td>Sociability, dominance and warmth</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Openness to experience</td>
<td>Intelligent, imaginative, perceptive, creativity, differentiated emotions, aesthetic sensitivity, need for variety, and unconventional values, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Agreeableness</td>
<td>Altruism, nurturance, caring, and emotional support, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>Thorough, neat, well organized, diligent, and achievement-oriented, etc.</td>
</tr>
<tr>
<td>Behavioral Inhibition System/ Behavioural Activation System</td>
<td>BIS/BAS</td>
<td>Internet Testing</td>
<td>BIS</td>
<td>Reactions to the anticipation of punishment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BAS – Drive</td>
<td>Persistence of pursuit of desired goals</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BAS – Fun Seeking</td>
<td>Desire for new rewards, willingness to approach a potentially rewarding event on the spur of the moment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BAS – Reward Responsiveness</td>
<td>Positive responses to the occurrence or anticipation of reward</td>
</tr>
<tr>
<td>E-Scale</td>
<td>E-Scale</td>
<td>Internet Testing</td>
<td>Readiness for empathy</td>
<td>Fictitiously imagine the behaviour and experience of another</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Social Concern</td>
<td>The extent to which a person tends to experience empathic behaviour in real life situations</td>
</tr>
<tr>
<td>Positive and Negative Affect Schedule</td>
<td>PANAS</td>
<td>On the spot of fMRI experiment</td>
<td>Positive affects</td>
<td>Interested, exited, strong, enthusiastic, proud, alert, inspired, determined, attentive, and active</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Negative affects</td>
<td>Distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery, and afraid</td>
</tr>
<tr>
<td>State-Trait Anxiety Inventory</td>
<td>STAI</td>
<td>On the spot of fMRI experiment</td>
<td>State anxiety</td>
<td>Transitory emotional state, subjective, consciously perceived feeling of tension and apprehension, heightened autonomic nervous system activity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Trait anxiety</td>
<td>General tendency to respond with anxiety to perceived threats in the environment</td>
</tr>
</tbody>
</table>
Interpretation: Neuroticism represents individual differences in the tendency to experience distress and the cognitive and behavioural style that follow from this tendency. The definition of Extraversion bears less consensus in the history. It refers mostly to social aspects of the individual. McCrae & Costa (1989) suggested that extraversion is also a mixture of dominance and warmth, which is close to Eysenck’s construct (Eysenck & Eysenck 1975). The dimension of openness to experience is rather broad. It can be intelligent, imaginative, perceptive, or can represent creativity, differentiated emotions, aesthetic sensitivity, need for variety, and unconventional values. Agreeableness represents the most humane aspect of humanity – characteristics such as altruism, nurturance, caring, and emotional support. Conscientiousness reflects the individual differences in whether they are thorough, neat, well-organized, diligent, and achievement-oriented.

II. BIS/BAS

BIS/BAS (Carver & White, 1994) is designed to capture two action control systems (approach, withdrawal) based on Gray’s biological theory of personality (1972, 1981). Current study used the German adaptation of the original scale (Strobel et al., 2001). BIS refers to behavioural inhibition system. BAS refers to behavioural approach system. This test was also administered through the internet. Subjects needed to finish rating 24 statements on a 4-point scale. The BIS/BAS scale was chosen based on assumption 2.

Interpretation: The BIS scale, or the so-called punishment sensitivity scale, includes all items referencing reactions to the anticipation of punishment. In contrast to the unidimensional character of BIS scale, the BAS contains three subscales: the Drive scale is made of items pertaining to the persistent pursuit of desired goals; the Fun Seeking scale has items reflecting both a desire for new rewards and a willingness to approach a potentially rewarding event on the spur of the moment; and the Reward Responsiveness scale has items that focus on positive responses to the occurrence or anticipation of reward. The German adaptation has less consensus on the differentiation of three subscales on the BAS dimension (Strobel et al. 2001).
III. Additional scales

i. E-Scale
E-Scale (Leibetseder & Laireser 1994; Leibetseder et al. 2001) is a German questionnaire which is designed to assess empathy. Factor analysis showed that the scale contains two major dimensions: readiness for empathy (Einfühlungsbereitschaft) and social concern (Betroffenheit). Readiness for empathy refers to fictitiously imagining the behaviour and experience of another. Social concern refers to the extent to which a person tends to experience empathetic behaviour in real life situations. The test was again administered through the internet. Subjects had to rate 25 items on a 5-point scale. The questionnaire of empathy is chosen, based on the studies from the psychoanalytic perspectives of personality, which suggest a high degree relationship between empathetic capacity and pronunciation aptitude.

ii. PANAS
Positive and Negative Affect Schedule (PANAS; Watson et al. 1988) is a measure of the tendencies to experience positive and negative affects. Carver and White (1994) found that the PANAS negative affectivity scale was related to the BIS scale, but not to the BAS scale (and subscales); the latter, not the former, were related to the positive affectivity scale of the PANAS. Therefore, this test serves as a natural control of the reliability of the internet testing of BIS/BAS. This questionnaire was administered in a subgroup of subjects. Subjects had to evaluate their overall affect during the experiments.

iii. STAI
State-Trait Anxiety Inventory (Spielberger 1983) is often used for assessing the trait and state aspects of anxiety. According to the “person-situation” debate within the research of influence of anxiety on SLA behaviour (see part 5.3), most researchers agreed that the situation specific language anxiety is better in predicting the SLA behaviour other than trait or state anxiety. Therefore, we do not expect to find a relationship between these two scores and pronunciation aptitude. However, since BIS is highly related to the anxiety trait (Carver & White, 1994), this scale was also administered in a subgroup of the subjects on the spot for controlling the reliability of the online testing.
10. Behavioural results

10.1. Test reliability

The first-stage analysis of the behavioural data included the correlation analysis between pronunciation talent assessed by Jilka and several personality scores. All the correlation analyses were two-tailed. The talent score scaled from 1 to 6, with smaller numbers indicating higher pronunciation talent. Although the talent score suffered from the suspicion of the validity of single-rater assessment, we considered this score to be highly reliable, since it was highly correlated with the median scores of another five native speaker raters ($r = .85$). As mentioned before, the intercorrelation between BIS/BAS and PANAS or STAI can be used as the test reliability for BIS/BAS. There was significant correlation between BIS and STAI trait ($r = .65$), and between BIS and STAI state ($r = .52$). There was also a significant correlation between BIS and the negative affectivity scale of the PANAS ($r = .34$) but not with the positive scale ($r = -.07$). This indicated that the BIS scores were highly reliable. The BAS scale, however, was significantly correlated with both the PANAS positive subscale ($r = .33$) and the negative subscale ($r = .27$). The relationship to the positive subscale was consistent to the assumption. For the negative subscale, post hoc analysis on the relationship between the negative subscale and the three subscales of BAS (drive, fun seeking, and reward responsiveness) did not reveal any significant correlations among them. Therefore, the BAS scale can be also considered reliable. Thus, although the BIS/BAS questionnaire was administered through internet testing, we can still confirm the reliability of the test. This might be also related to the test reliability of NEO-FFI, since two questionnaires had the same testing condition.

10.2. Correlation between personality traits and pronunciation talent

The pronunciation talent score was found not related to verbal IQ ($r = -.115$) or nonverbal IQ ($r = -.01$). This indicated that intelligence did not serve as a determining factor for the pronunciation aptitude in current group of subjects. The pronunciation talent score was, however, significantly correlated with the PANAS positive subscale ($r = -.34$) and its several sub-items (excited, $r = -.42$; proud, $r = -.28$; determined, $r = -.27$). This revealed that the subjects with a greater degree of pronunciation talent experienced more positive affects such as being excited, proud.
and determined during the phonetic-articulation tasks. This is quite reasonable, since people tend to be happier if they are asked to show their strengths. However, whether these positive affects turn to facilitate the pronunciation performance is not clear.

For the NEO-FFI scale, no significant correlations were found to the extraversion, openness to experience or neuroticism; whereas moderate correlations were found to conscientiousness ($r = .31$) and to agreeableness ($r = .25$). No relationship to extraversion might point out to the fact that pronunciation aptitude was a very specific (phonetic-articulatory) aspect of the oral language. It is different from other oral language proficiency measures (such as classroom speech rating), that it doesn’t need extraverts’ social capability, where they might have some pragmatical advantages over introverts on the speaking language proficiency measures. The correlation between SLA and agreeableness or conscientiousness indicates that more talented persons tend to be more agreeable and less conscientious. We still need to investigate these data further to make less explicit and final conclusions.

Interestingly, the readiness for empathy measured by E-Scale had also a significant relationship to pronunciation talent ($r = -.28$). This indicates that more talented persons tended to have more readiness for empathy or, in other words, were able to fictitiously imagine the behaviour and experience of another. This result was consistent with the afore-mentioned research on the pronunciation aptitude and ego boundary or empathetic capacity (see part 3, psychoanalytic perspectives). More specific investigations are needed to clarify this relationship we’ve observed, for example, to incorporate questionnaire scores of empathy into the analysis of brain imaging data, or to use another empathy questionnaire for retesting the relationship.

So far, the behavioural results have shown that some personality traits do have significant correlational relationships to pronunciation aptitude. Nevertheless, the correlations were not found in the assumed trait-extraversion, but in other personality factors (agreeableness, conscientiousness, and readiness to empathy). Therefore, assumption 1 of the research design has been partly confirmed. The function of the BAS dimension was not present in the current result on the behavioural level. But we cannot rule out the possibility that the BAS dimension to be present on the neural level, since the BAS is a more biologically based personality construct *per se*. This possibility is similar to what Gray and Braver (2002) have found: BAS has significant correlation to working memory (WM) related brain activity in caudal anterior cin-
gulated cortex (ACC), but has only a trend of correlation to the behavioural performance of WM. After all, more work still needs to be done in order to verify the other assumptions; further results are forthcoming.

**Conclusion:** As we have pooled the knowledge both from the review of former literature and the first-stage behavioural data, we can conclude that personality plays an important role in the second language acquisition (SLA) behaviour. This effect also seems to exist in a very specific phonetic-articulatory aspect of the second language – pronunciation talent. As we have reviewed different research approaches in the research body of SLA and personality in the past four decades, we are confident in using the trait approach for measuring of personality on the behavioural level; additionally, we need to incorporate cognitive and biological perspectives into the current study. This attempt will hopefully add to the current or even open another area of SLA – personality research.

**References**

Allport, G.W.  

Asendorpf, J.B.  

Bailey, P., Onwuegbuzie, A.J., & Daley, C.E.  

Bandura, A.  


Borkenau, P., & Ostendorf, F.  

Brown, H.D.

Bunge, S.A., Klingberg, T., Jacobsen, R.B., & Gabrieli, J.D.E.

Busch, D.

Canli, T.

Canli, T., & Amin, Z.

Canli, T., Sivers, H., Whitfield, S.L., Gotlib, I.H., & Gabrieli, J.D.

Canli, T., Zhao, Z., Desmond, J.E., Kang, E., Gross, J., & Gabrieli, J.D.

Carrell, P.L. & Prince, M.S., & Astika, G.G.

Carver, C.S., & Scheier, M.F.

Carver, C.S., Sutton, S.K., & Scheier, M.F.

Carver, C.S., & White, T.L.

Cattell, R.B.

Cattell, R.B., Eber, H. W., & Tatsuoka, M.M.

Chapelle, C. & Green, P.

Clément, R., Gardner, R.C., & Smythe, P.C.
1977 Motivational variables in second language acquisition: A study of

Clément, R., Gardner, R.C., & Smythe, P.C.  

Cloninger, C.R., Svrakic, D., & Przybeck, T.  

Coan, J.A. & Allen, J.J.B.  

Coan, R.W., & Cattell, R.B.  
1966 Early school personality questionnaire. Champaign, IL: Institute for Personality and Ability Testing.

Coltheart, M. & Friston, K.J.  
2008 Controversial debate on “functional neuroimaging has already told us a lot about cognition: Yes or no?” (Moderator: A. Villringer). 29th International Congress of Psychology, Berlin.

Costa, P.T., & McCrae, R.R.  
1985 The NEO Personality Inventory manual. Odessa, FL: Psychological Assessment Resources.

Costa, P.T., & McCrae, R.R.  

Costa, P.T., & McCrae, R.R.  
1992 Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory: Professional manual. Odessa, Florida: Psychological Assessment Resources.


Dewaele, J.M.  

Dewaele, J.M., & Furnham, A.  

Dewaele, J.M., Petrides, K.W., & Furnham, A.  
Dörnyei, Z.

Dörnyei, Z.

Dörnyei, Z., & Kormos, J.
2000 The role of individual and social variables in oral task performance. Language Teaching Research, 4, 275–300.

Dörnyei, Z., & Skehan, P.

Ehrman, M.E., & Oxford, R.L.

Eisenberger, N., Lieberman, M.D., & Satpute, A.B.

Eysenck, H.J.

Eysenck, M.W.

Eysenck, H.J.

Eysenck, H.J., & Eysenck, S.B.G.

Eysenck, H.J., & Eysenck, S.B.G.

Fahrenberg, J., & Selg, H.

Funder, D.C.

Furnham, A.
1990 Language and personality. In H. Giles and W.P. Robinson (Eds.), Handbook of language and social psychology. Chichester, UK: John Wiley & Sons. 73–95.

Gardner, R.C.
Gardner, R.C. & Smythe, P.C.  

Granger, C.A.  

Gray, J.A.  

Gray, J.A.  

Gray, J.A.  

Gray, J.A.  

Gray, J.A.  

Gray, J.A.  

Gray, J.A.  

Gray, J.A.  

Gray, J.A., & Braver, T.S.  
2002 Personality predicted working-memory-related activation in the caudal anterolateral cingulated cortex. *Cognitive, Affective, and Behavioral Neuroscience,* 2 (1), 64–75.

Gray, J.A., Burgess, G.C., Schaefer, A., Yarkoni, T., Larsen, R.J., & Braver, T.S.  

Gregersen, T., & Horwitz, E.  
Guiora, A.Z.

Guiora, A.Z.


Haas, B.W., Omura, K., Amin, Z., Constable, R.T., Canli, T.
2006 Functional connectivity with the anterior cingulate is associated with extraversion during the emotional Stroop task. *Social Neuroscience, 1* (1), 16–24.

Harmon-Jones, E., & Allen, J.J.B.

Hewig, J., Hagemann, D., Seifert, J., Naumann, E., & Bartussek, D.

Horwitz, E.K.

Horwitz, E.K., Horwitz, M.B., & Cope, J.

Humphreys, M.S., & Revelle, W.

Larsen, R.J., & Buss, D.M.

Lawrence, G.D.
1997 *Looking at type and learning styles.* Gainesville, FL: Center for Applications of Psychological Type.

Leibetseder, M., & Laireiter, A.R.

Leibetseder, M., Laireiter, A.R., Riepler, A., & Köller, T.
2001 E-Skala: Fragebogen zur Erfassung von Empathie – Beschreibung und psychometrische Eigenschaften. [E-Scale: Questionnaire for the Assess-
Levitt, E.E.

Lieberman, M.D., & Rosenthal, R.

MacIntyre, P.D.

MacIntyre, P.D., & Gardner, R.C.

MacIntyre, P.D., & Gardner, R.C.
1994  The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44 (2),* 283–305.

MacIntyre, P.D., & Noels, K.A.

Maltby, J., Day, L., & Macaskill, A.
2007  *Personality, individual differences and intelligence.* Estella, Bilbao, Spain: Pearson Education.

McCrae, R.R., & Costa, P.T.

McKenna, F.P.

Mobbs, D., Hagan, C.C., Azim, E., Menon, V., & Reiss, A.L.

Myers, I.B. & Briggs, K.


Quay, H.C.
Rapp, A.M., Wild, B., Erb, M., Rodden, F.A., Ruch, W., & Grodd, W.

Reiterer, S., Hemmelmann, C., Rappelsberger, P., & Berger, M.L.

Rotter, J.B.

Schumann, J.H., Holroyd, J., Campbell, R.N., & Ward, F.A.

Spielberger, C.D.

Stellar, J.R., & Stellar, E.

Strauss, A., & Corbin, J.

Strobel, A., Beauducel, A., Debener, S., & Brocke, B.

Strong, M.

Sutton, S.K., & Davidson, R.J.

Taylor, J.A.

Taylor, L.L., Guiora, A.Z., Catford, J.C., & Lane, H.L.

van Lieshout, C.F., & Haselager, G.J.
1992 Persoonlijkheidsfactoren in Q-sortpersons-beschrijvingen van kinde-

Verhoeven, L., & Vermeer, A.
2002 Communicative competence and personality dimensions in first and

Vogel, K., & Vogel, S.

Watson, D., Clark, L.A., & Tellegen, A.
1988 Development and validation of brief measures of positive and negative

Williams, M.
1994 Motivation in foreign and second language learning: An interactive

Williams, M., & Burden, R.
1999 Students’ developing conceptions of themselves as language learner.

Wong, M.S.L.
2005 Language learning strategies and language self-efficacy. Regional Lan-
guage Centre Journal, 36 (3), 245–269.

Woodrow, L.
2006 Anxiety and speaking as a second Language. Regional Language
Centre Journal, 37, 308–328.

Wright, C.I., Feczko, E., Dickerson, B., & Williams, D.
2007 Neuroanatomical correlates of personality in the elderly. NeuroImage,

Yerkes, R.M., & Dodson, J.D.
1908 The relation of strength of stimulus to the rapidity of habit formation.
Journal of Comparative Neurology and Psychology, 18, 459–482.

Zuckerman., M.
1960 The development of an affective adjective check list for the measure-

Zuckerman., M.
1991 Psychobiological of personality. Cambridge: Cambridge University
Press.
1. History of language assignments

Language-related functions were among the first to be ascribed a specific location in the human brain and have been a subject of intense research for well over a century. A “classical model” of language organization, based on data from aphasic patients with brain lesions, was popularized during the late 19th century. First in 1861, Broca reported a post-mortem study of a patient whose ability to articulate language was impaired. This patient had a lesion that encompassed the third frontal convolution in the left hemisphere. By deduction, the damaged area – later referred to as Broca’s area – was associated with the motor images of speech. Similarly, Wernicke (1874) reported a post-mortem study of a patient with impaired speech comprehension. Damage was found in the left posterior superior temporal cortex and this region – later referred to as Wernicke’s area – was associated with the auditory images of speech (see Figure Grodd/Wildgruber/Kumar 1 in the Colour figure section). Wernicke developed the model further to predict that if there were damage to the white matter tracts that connect Broca’s and Wernicke’s areas (the arcuate fasciculus), patients would have intact speech comprehension and production but a deficit repeating what was heard. This type of disconnection syndrome, referred to as ‘conduction aphasia’, was first reported by Lichtheim (1885). With respect to word reading Dejerine (1892) first distinguished two alexic syndromes: ‘alexia with agraphia’ and ‘alexia without agraphia’. Alexia with agraphia described patients who exhibited an acquired deficit in reading (alexia) and writing (agraphia) and this was associated with damage to the left angular gyrus. The left angular gyrus was therefore linked to memories of visual word forms (see Figure Grodd/Wildgruber/Kumar 2a in the Colour figure section). In contrast, alexia without agraphia (which is associated with lesions to the left occipital lobe and the splenium of the corpus callosum) was thought to arise from a disconnection of the left angular gyrus from the visual cortex (Damasio & Damasio 1983).
Connections between the angular gyrus and Wernicke’s area link visual word forms to the corresponding auditory word forms which can then be articulated as speech (see Geschwind 1965). Although some researchers still accept this classical basic scheme, a more detailed account of language organization is emerging, although it has not yet gained an overall approval (for review see Price 2000; Démonet et al. 2005). The major causes for these reconsiderations are refined neuroanatomical examinations and a vast amount of human brain mapping studies, which have boosted our knowledge of the underlying anatomical and cognitive substrates of language perception and production. Apart from the distinction between language production and comprehension, more specific linguistic aspects of language processing like semantics, phonology, syntax, and prosody have meanwhile gained major interest in research.

2. Neuroanatomy of language areas

2.1. Broca’s area

Broca’s area in the left hemisphere and its homologue in the right hemisphere are regions that usually refer to the pars triangularis (PTr) and pars opercularis (POp) of the inferior frontal gyrus. It is apically limited by the inferior frontal sulcus, dorsally by the precentral sulcus, and surrounds caudally the ramus horizontalis of the Sylvian fissure. The PTr lies immediately apical to the pars orbitalis and anterior to the POp. The POp is immediately posterior to the PTr and anterior to the precentral sulcus (see Figure Grodd/Wildgruber/Kumar 2b in the Colour figure section). PTr and POp defined by these structural landmarks can only be subdivided by their cytoarchitectonic appearance (Amunts et al. 1999) into the anterior and posterior cytoarchitectonic areas 45 and 44, respectively, by Brodmann’s classification of cortical areas 45 and 44, respectively, by Brodmann’s classification of cortical areas (BA, Brodmann 1909). The BA 45 is granular, containing a layer IV, whereas the BA 44 is dysgranular and distinguished from the more posterior agranular BA 6 in that it does not contain Betz cells in layer V. These differences in cytoarchitecture between BA 45 and 44 suggest a corresponding difference in connectivity between the two areas and the rest of the brain. BA 45 receives more afferent connections from prefrontal cortex, the superior temporal gyrus, and the superior temporal sulcus, compared to BA 44. In contrast, BA 44 receives more afferent connections from motor, somatosensory, and inferior parietal regions (Deacon 1992; Petrides & Pandya 2002). The dif-
ferences between BA 45 and 44 in cytoarchitecture and in connectivity also suggest that these areas perform different functions. Finally, the inferior frontal lobe and especially Broca’s area are characterized by great individual and interhemispheric variability in respect to its i) gyral and sulcal pattern, ii) volume of microscopically defined areas and iii) position with respect to the cortical profile. Therefore, it is recommended to use the anatomical probabilities maps of BA 44 and 45 for a detailed functional assignment (Amunts et al. 2004).

2.2. Wernicke’s area

Wernicke’s area, known as the sensory language area, is located directly adjacent to Heschl’s gyrus (BA 41 + 42) in the posterior part of the superior temporal gyrus, and, in contrast to Broca’s area, it is anatomically less well-defined (see Figure Grodd/Wildgruber/Kumar 2b in the Colour figure section). It embraces the planum temporale, parts of the medial temporal gyrus, and reaches into the inferior parietal lobe (parts of angular and supramarginal gyrus). Consensus exists that at least the posterior part of BA 22 of the middle temporal gyrus (MFG) is part of Wernicke’s area (Aboitiz and Garcia 1997), but eventually it also encircles parts of BA 42, 39, 40, and 37. Wernicke’s area, similarly to Broca’s area, reveals significant individual and hemispheric variability. Especially this variability and its striking right-left asymmetry could underlie language lateralization and account for the numerous reports on interhemispheric differences (Geschwind and Levitsky 1968).

2.3. Additional brain areas

In addition to Broca’s and Wernicke’s area, a number of other cortical and subcortical brain areas contribute substantially to the various language processes. Especially the primary motor cortex (M1), the premotor cortex, and the supplementary motor area (SMA) are involved in articulation and speech motor control. Besides these motor competent areas, subcortical regions in the basal ganglia and thalamus as well as the cerebellum also contribute to speech motor planning and execution of motor programs. By now, imaging studies investigating the cognitive-linguistic interface between speech perception and production have become a major focus of language research, and a number of additional areas have been implicated in semantic, syntactic, and grammatical processing (Ojemann 1994; Démonet et al. 2005; Noppeney 2008).
3. Functional brain imaging methods

Non-invasive functional imaging methods are a new source of data on language organization in the intact human brain. First reports utilizing positron emission tomography (PET) described the functional anatomy of auditory comprehension of language, the processing of single words, and role of the right hemisphere in figurative language aspects (Petersen et al. 1988; Démonet et al. 1992; Bottini et al. 1994). Additional time resolved brain mapping methods like electroencephalography (EEG) and magnetoencephalography (MEG) have emerged over the last decades and substantially contributed to our understanding of the temporal aspects of linguistic processing in speaking, understanding, reading and writing (Démonet et al. 2005). But with the advent of magnetic resonance imaging (MRI), new tools have appeared to localize human brain functions more precisely. The first one among those tools, i.e. functional magnetic resonance imaging (fMRI), is based on monitoring regional changes in blood oxygenation (see Figure Grodd/Wildgruber/Kumar 3 in the Colour figure section) resulting from neural activity. fMRI, often abbreviated as BOLD (blood oxygen level dependent) response (Ogawa et al., 1990, 1992), has nowadays gained worldwide acceptance to reveal brain activation associated with language, sensorimotor, visual, emotional, and cognitive processing. Although certain technical issues remain to be resolved, the capabilities of fMRI for localizing sensory, motor, and language areas are well established (Kim et al. 1993; Rao et al. 1993; Binder et al. 1994b; DeYoe et al. 1994). A recent addition to the MRI instrumentation is the development of Diffusion Weighted Imaging (DWI). This method uses three-dimensional measurements to determine the random motion of water molecules undergoing diffusion. A DWI derived technique is Diffusion-Tensor-Imaging (DTI), where at least 6 different directions are acquired to determine a Tensor for each voxel. The Eigenvectors and the corresponding Eigenvalues of the Tensor are consecutively used to apply tractography methods by which the three-dimensional course of white matter tracts can be exploited to study fiber pathways and connectivity between brain regions (see Figure Grodd/Wildgruber/Kumar 4 in the Colour figure section).
4. Imaging of language related areas

In this chapter we will give a short overview of some major fMRI and DTI findings in language production and perception mainly from experiments obtained at our institutions. It is beyond the scope of this introduction to present a complete and comprehensive survey of all significant imaging findings in language research. Language or, more specifically, language processing refers to the way human beings process speech or written language and understand it as language. Language processors activate linguistic representations during speaking, understanding, reading and writing. Brain imaging examinations on language processing can be separated experimentally into language production and language perception studies. And those studies mainly deal with language comprehension. Although it turns out that this distinction is usually not reflected in the obtained imaging pattern, we will follow this scheme to present some of the core findings.

4.1. Language production

During speaking, we select words in accordance with what we think our listener will understand. We have to generate a syntactic structure to concatenate them to meaningful sentences accompanied by an appropriate intonational contour and then to select the right phonemes out of the repertoire of a given language. Finally, these neuronal representations have to be transferred to concrete motor plans realized by movements of the mouth, jaw, tongue, palate, larynx and other articulators that are regulated on a millisecond basis to produce one sound every tenth of a second on average. This high accuracy of planning and production of sounds is probably the most complex human motor skill and involves the coordinated use of approximately 100 different muscles, which are distributed over three anatomically distinct structures: the respiratory, the laryngeal, and the supralaryngeal tract (Levelt 1995). These complex and varying neuronal patterns recruit all components of the motor system to various degrees and extend to numerous other cortical areas; and this can only partially be captured with fMRI and other mapping methods.
4.1.1. Syllable and word production

While the activation pattern in M1 follows the established somatotopic arrangement for lips, tongue and larynx, the premotor areas are thought to subserve speech motor planning. Simple overt syllable and word production regularly activates bilateral sensorimotor cortices corresponding to the somatotopic arrangement for lips, tongue and vocalisation musculature (Lotze et al. 2000) with a preponderance for syllable and word generation to the left hemisphere (see Figure Grodd/Wildgruber/Kumar 5a in the Colour figure section). Paradoxically, tongue movement evokes a stronger fMRI response in cortical areas, as well as in the cerebellum, than simple syllable and word production with the same pacing (see Figure Grodd/Wildgruber/Kumar 5b in the Colour figure section. The diminished activation with increasing sound complexity may indicate the recruitment of a larger neuronal network and a concomitant decrease of the local BOLD response, which consecutively fails to exceed the chosen threshold level (Riecker et al. 2000). The activation extent is also influenced by working memory load and here the frontal lobes play a critical role in the temporal organization and in serial aspects of speech processing. By comparing the ability to recite highly automated word strings (names of the months in normal against reverse order) Wildgruber et al. (1999) have shown that the increasing demand in working memory load yields an additional activation of the bilateral middle and inferior frontal gyri, the posterior parietal cortex and the left anterior cingulate gyrus (see Figure Grodd/Wildgruber/Kumar 5c in the Colour figure section).

4.1.2. Speech motor network

With respect to the underlying neuronal network, it is important to note that the cerebellum contributes to speech motor control with different compartments and varying intensity depending on the task and complexity (see Figure Grodd/Wildgruber/Kumar 6a in the Colour figure section). While simple horizontal tongue movements elicit bilateral activation in lobulus HVI of the anterior lobe, speaking and singing evoke activation contralateral to the dominant cerebral hemisphere in crus I (Ackermann et al. 1998). In order to delineate additional neuroanatomical correlates, Wildgruber et al. (2001) and Riecker et al. (2002) performed silent repetitions of the syllable “ta” at different rates and found that the spatial extent and magnitude of motor cortex activations are positive when correlated to production frequencies. In the basal ganglia, lower
repetition rates gave rise to higher magnitudes of activation within the left putamen. In contrast, cerebellar responses were rather restricted to fast performance and exhibited a shift in caudal direction during 5.5 Hz suggesting a differential impact of various cortical and subcortical areas on speech motor control (see Figure Grodd/Wildgruber/Kumar 6b in the Colour figure section). Calculation of rate/response functions revealed a negative linear relationship between repetition frequency and BOLD response within the left striatum, whereas both cerebellar hemispheres exhibited a step-wise increase of activation at ~3 Hz (see Figure Grodd/Wildgruber/Kumar 6c in the Colour figure section). These findings suggests that speech motor control may be organized into two separate networks: the "preparatory loop" for motor planning and preparation (SMA, dorso-lateral frontal cortex including Broca area, anterior insula, and superior cerebellum) and the "executive loop" for motor execution processes (sensorimotor cortex, basal ganglia, thalamus, and inferior cerebellum).

4.1.3. Speaking and singing

With respect to the acoustic output aside from spoken language, singing represents a second mode of auditory-vocal communication in humans. Here, fMRI revealed a different lateralization pattern with two complementary cerebral networks subserving singing and speaking (Wildgruber et al. 1996; Riecker et al. 2000). The reproduction of a non-lyrical tune elicits activation predominantly in the right motor cortex, the right anterior insula, and the left cerebellum, whereas the opposite response pattern emerges during speech (see Figure Grodd/Wildgruber/Kumar 7a in the Colour figure section). Interestingly, this different hemispheric lateralisation of speech and singing is more pronounced in the silent (covert) mode. The activation of the intrasylvian cortex is bound to overt task performance, in contrast to the hemodynamic responses within the motor cortex and cerebellum (see Figure Grodd/Wildgruber/Kumar 7b in the Colour figure section).

4.1.4. Reorganisation of speech production

The proper implementation of language functions depends on intact psychosocial development and is determined by the maturational stage of the central nervous system. Therefore brain injuries acquired early in life can cause reorganization of language and especially of speech motor
functions. This issue was addressed in a number of studies with children suffering from periventricular white matter lesions (PVL) which cause impairment of hand motor function and facial output connections (Staudt et al. 2001 and 2002). The patients had no obvious cognitive or language deficits and were examined as young adults (see Figure Grodd/Wildgruber/Kumar 8a in the Colour figure section). The cortical organization of language was assessed using two different activation tasks: To monitor language perception, a narrative story was presented. For the assessment of language production, the subjects were instructed to generate word chains silently. The results of the word generation demonstrated extensive changes in the cortical representation of language in all patients, although cortical language areas were not directly affected by the lesion (see Figure Grodd/Wildgruber/Kumar 8b in the Colour figure section). Right-hemispheric activation was found to occur homotopically to the left-hemispheric language areas and correlated with the severity of white matter damage and especially with facial pyramidal tract involvement. In addition the reorganization also embraced the cerebellum where the extent of laterality of the cerebellar activation correlated significantly with the laterality of the frontal lobe activation (Lidzba et al. 2008). It was suggested that the developing brain reacts to early focal lesions in the left hemisphere with a mirror-image organization of the entire cerebro-cerebellar network engaged in speech production.

4.2. Language perception

In order to address imaging studies on language perception, we present them based on the two main input routes: audition (spoken words and environmental sounds) and vision (written code, sign language, scene, and picture viewing).

4.2.1. Auditory input

The classical view concerning speech perception is the dominance of the left temporal cortex (i.e. Wernicke’s area). A number of studies (Binder et al. 2000, 2004; Hickok & Poeppel 2000) has shown that the anterior superior temporal gyrus (STG) and superior temporal sulcus (STS) surrounding the primary auditory cortex (i.e. medial part of Heschl’s gyrus) in both hemispheres are the main cortical substrates for the auditory representation of sounds and speech components. While pure tone stimu-
lation elicits bilateral activation often with a slight preponderance to the right side, auditory processing of speech components is linked to a varying degree of left-sided dominance (see Figure Grodd/Wildgruber/Kumar 9 in the Colour figure section). This left-sided accentuation may depend on a) the rate of speech changes over time (Belin et al. 1998), b) the number of phonological entities (Démonet et al. 1994) and c) whether input information is processed verbally or nonverbally (Thierry et al. 2003).

Besides this lateralization the neural correlates of auditory perception in the superior temporal cortex appear to be organized along a rostro-caudal gradient surrounding the primary auditory cortex (Alain et al. 2001; Maeder et al. 2001). This rostro-caudal gradient relates to a functional distinction of a “what and where” pathway comparable to that in the visual modality (Ungerleider et al. 1982; Haxby et al. 1991). Using fMRI Wessinger et al. (2001) described a “core” region involved in the perception of pure tones surrounded by “belt” regions, which are selectively activated by sounds with greater spectral complexity. Considering the rostral part of this system, Rauschecker and Tian (2000) identified small areas in the anterior lateral belt region responding to specific “meaningful” stimuli such as species specific vocalizations. In humans, Zatorre et al. (2004) described that the right anterior STS is sensitive to auditory object distinctiveness, for example, to identify the characteristic sound of a trumpet independent of the produced melody. A recent attempt to localise cortical areas specific to the human voice could show that specific neural fingerprints exist bilaterally in the early auditory areas (Heschl’s gyrus), in the planum temporale along the STS, and in the middle temporal gyrus (Formisano et al. 2008), which are insensitive to acoustic variations, but can retrieve what and to whom a person is listening.

In summary, the left STG seems to be functionally heterogeneous, as it is activated by a variety of conditions from phonological perception, access to lexical representations (Howard et al. 1992), and monitoring the speaker’s own voice, to word retrieval from semantic memory. Wise et al. (2001) distinguished two distinct subregions in the posterior left STG. These transient representations would be matched to the phonological form of words stored in lexical long-term memory. This left-lateralized component might be complemented by a homotopic region in the right hemisphere whose activity would depend on the frequency (Majerus et al. 2002). Wise et al. (1999) also described a small subregion, located dorsally to the left superior temporal sulcus, at the junction between the STG and the supramarginal gyrus, which is activated during speech
movements, and suggested that this region could contribute as an interface of auditory speech representations to their motor counterparts (see Figure Grodd/Wildgruber/Kumar 9b in the Colour figure section). Therefore, speech information coded by the auditory system is likely to be further transcoded into motor speech acts, when corresponding signals are sent dorsally from posterior temporal regions to the inferior parietal cortex (especially the supramarginal gyrus) and, ultimately, to the inferior dorsal premotor cortex via the arcuate fasciculus (Belin & Zatorre 2000). This transcoding process requires short-lived maintenance of speech representation via a system of phonological working memory. Addressing the substrates of phonological working memory, Paulesu et al. (1993) and Démonet et al. (1994) described specific activations located at the junction between the posterior STG and the inferior part of the supramarginal gyrus and proposed that this region harbours the substrates of the transient phonological store as defined in Baddeley’s model (Baddely 1986).

In language production, as well as in perception, significant linguistic and emotional information is conveyed by the prosodic features of speech. During the processing of emotional prosody multiple successive processing stages are involved as reported by (Wildgruber et al. 2006; Ethofer et al. 2006a; Wiethoff et al. 2008). Firstly, extraction of supra-segmental acoustic information is associated with the activation of predominantly right hemispheric primary and secondary acoustic regions (see Figure Grodd/Wildgruber/Kumar 10a in the Colour figure section). Secondly, representation of meaningful supra-segmental acoustic sequences is linked to higher-order acoustic areas housed within the right posterior STS (Wildgruber et al. 2005; Ethofer et al. 2006b). Thirdly, emotional judgement is linked to the bilateral inferior-frontal cortex (Wildgruber et al. 2004). Within this network, projections from primary to secondary acoustic representation within the middle STG seem to be predominantly stimulus-driven (bottom-up effect), whereas projections to the posterior STG and the inferior frontal cortex depend upon focussing of attention on evaluation (top-down effects). Considering implicit processing of emotional prosody, task-independent activation of the amygdalae and the medial orbito-frontal cortex has been observed (Sander et al. 2005; Ethofer et al. 2009; Kreifelts et al. 2009). Overall, these findings indicate that partially distinct networks subserve processing of phonetic and intonational information during speech perception (see Figure Grodd/Wildgruber/Kumar 10b in the Colour figure section).
4.2.2. Visual input

Reading is a highly specialized, evolutionarily developed human skill, by which graphic features associated with phonemes or syllables are integrated to access word meaning. After low-level perceptual analysis in the primary visual cortex, early processing of graphic stimuli elicits activation in the visual association cortex bilaterally (Cohen et al. 2000), especially in its ventral and medial parts (see Figure Grodd/Wildgruber/Kumar 11a in the Colour figure section). Petersen et al. (1988) have already stressed the importance of left medial extrastriate cortex (lingual gyrus) for reading of words and pseudowords. Silent reading of words compared with false font viewing activates portions of the left posterior inferior aspect of the medial fusiform gyrus, the posterior part of the left superior temporal gyrus, and the cerebellum (Price 2000; Turkeltaub et al. 2008).

Consequently, the posterior region of the left mid fusiform gyrus was denoted as the “visual word form area” (VWFA) by some authors (Cohen et al. 2002). This finding has not been validated in other studies where the so-called VWFA was also found to be activated by normal subjects during tasks that do not engage visual word form processing, such as naming colours or pictures, reading Braille, repeating auditorily presented words, and performing manual action responses to pictures of meaningless objects. Price and Devlin (2003) therefore concluded that neither neuropsychological nor neuroimaging data are consistent with any cortical region specialized for visual word form representations.

Reading and word retrieval consecutively activate the left posterior STS, parts of the left posterior inferior temporal cortex, and the angular gyrus (see Figure Grodd/Wildgruber/Kumar 11b in the Colour figure section). However, it is not clear at present whether this pattern of activation relates to lexicon access (Howard et al. 1992), phonological access, or both (Fiez and Petersen 1998). For overt reading, the right superior temporal activation in reading aloud is thought to be associated with the fact that subjects can perceive their own voice (Price et al 1996; Price 2000). In the classical model the involvement of the angular gyrus is assigned as a visual word form system (see Figure Grodd/Wildgruber/Kumar 2a in the Colour figure section). As proposed by Geschwind (1965), connections between the angular gyrus and Wernicke’s area link visual and corresponding auditory word form processing, which can be subsequently articulated. Today most researchers assume that the angular and the supramarginal gyrus are instead parts of a more distributed semantic system (Price 2000; Noppeney 2008; Noppeney et al. 2008). Additional evi-
dence comes from patients with lesions in left temporo-parietal regions extending beyond the Wernicke’s area and include the angular and supramarginal gyri which exhibit comprehension deficits concerning both auditory and visual input (Kertesz et al. 1982; Alexander et al. 1989; Hart & Gordon 1990).

Other perisylvian regions are activated during reading, especially when grapheme-to-phoneme conversions and general phonological manipulations are required. For instance, the supramarginal gyrus seems especially involved in studies dealing with pseudo-words or unfamiliar letter combinations (Price 1998). However, activation of the left posterior inferior frontal gyrus has also been clearly established in silent and even implicit reading (Price 2000; Price et al. 1996). Therefore, reading conditions are likely to involve phonological retrieval even when there is no speech output. Nevertheless, in the case of frequent, regular words, reading does not require precise phonological recoding, and phonological retrieval corresponds to direct and automatic access to their phonological form. In this condition, the neural network involved in reading (fusiform gyrus, the left STG, the supramarginal gyrus, and the left Broca’s area) might be modulated in such a way that activation in temporal and parietal regions associated with phonological encoding is reduced.

4.3. Language comprehension

The exploration of language comprehension is a core issue in cognitive neuroscience because it deals with the crucial function of how language is conveyed into meaning. Comprehension can be subdivided into three processing steps: initial structure building, semantic integration, and late syntactic integration. fMRI studies have revealed that semantic and syntactic processes are supported by separable temporo-frontal networks, with the syntactic processes involving the left STG, the left frontal operculum, and the basal ganglia in particular. With respect to the classical language model (see Figure Grodd/Wildgruber/Kumar 2 in the Colour figure section), where the posterior STS corresponds to ‘auditory word representations’, the left anterior insula and frontal operculum to ‘motor word representations’, and the left extrasylvian temporo-parietal regions to a ‘concept centre’, it can be resumed today that this model is outmoded, as there exists no more agreement, for example, which cortical areas make up the perceptive language system (Bogen and Bogen 1976) or what the specific linguistic role of Broca’s area is (Marie 1906; Mohr 1976; Reiterer et al. 2008).
One difficulty emerged from the inconsistency in the reading components. Critically, there are no consistent functional imaging data to indicate an anatomical region that corresponds to ‘visual word representations’ (Price 2000). Apart from the influence of word regularity within a given language, differences have been shown across languages. Paulesu et al. (1993) showed that reading English requires preponderant access to a lexicon of orthographic patterns that activates the left fusiform gyrus and the Broca’s area (see Figure Grodd/Wildgruber/Kumar 11b in the Colour figure section). In contrast, reading Italian, where letter-to-sound conversion is the predominant process, specifically activates the left STG. Japanese represents another important example of a language in which two different written codes coexist: Kanji (ideographic system) and Kana (syllabic system). Although some behavioural studies suggested a specific involvement of the right hemisphere for processing Kanji, neuroimaging studies have found selective left-sided activation (Nakamura et al. 2000). Sakurai et al. (2000) reported that activity was more pronounced in the lateral fusiform gyrus in Kanji, in contrast to the greater activation in the middle and inferior occipital gyri and the deep perisylvian temporo-parietal area in Kana, suggesting that Kanji and Kana are processed differently.

The other inconsistency exists with respect to assignments of inferior frontal lobe functions. One example is the variability of activation associated with the processing of two different kinds of grammatical structures, where the involvement of Broca’s area and the left premotor cortex seem to be more pronounced for hierarchical grammatical structures as compared to non-hierarchical structures (Bahlmann et al. 2008; Friederici et al. 2006). Secondly, spoken language constantly operates as a dual system, perceiving and producing utterances. These systems do not only alternate, but in many cases they partially or wholly operate in concert in a very short time window (Indefrey and Levelt 2004). Consequently, Hickok and Poeppel (2007) proposed a new dual stream model for speech processing. On the one hand, a ventral stream encompasses the superior and middle portions of the temporal lobe, which is involved in processing speech signals for comprehension (speech recognition). On the other hand, a dorsal stream comprising the posterior frontal lobe and the posterior aspect of the temporal lobe and parietal operculum, which is involved in translating acoustic speech signals into articulatory representations in the frontal lobe and is essential for speech development and normal speech production. The authors conclude that speech production tasks are associated to a greater extent with dorsal stream circuitry,
whereas speech comprehension tasks rely more on ventral stream circuitry (with shared neural tissue in the left STG). In addition, Hickok and Poeppel (2007) suggest – contrary to the common view that speech processing is mainly left-hemisphere dependent – that the ventral stream is bilaterally organized (although with important computational differences between the two hemispheres). Hence, the ventral stream itself comprises parallel processing streams. The dorsal stream, however, is strongly left-dominant, which could explain why production deficits are a prominent sequelae of dorsal temporal and frontal lesions and why left-hemisphere injury can substantially impair performance in speech perception tasks.

In summary, facing the increasing number of imaging studies related to language and the refinement in experimental approaches to assess language processing in the brain also clearly discloses that localisation with fMRI methods reaches its limits as soon as time-dependent processes in the second and sub second range are addressed. Due to the slow temporal behaviour of the underlying hemodynamic response function, fMRI is not able to disentangle the existing tight interrelationship between language perception, production and comprehension. It is therefore mandatory to apply complementary time resolved mapping methods like EEG and/or MEG to adequately address both time and space of language processes.

4.4. Connectivity of language areas

The human cortex is composed of a number of specialized areas, which are interconnected with each other. Therefore, a disconnection of fibre tracts can lead to dysfunction. This core finding – already proposed by Wernicke – constitute to a number of neurological disorders, which have already been reported in the 19th century and extensively reviewed by Geschwind in his seminal paper on “disconnection syndromes” (1965). Referring to earlier studies of Flechsig (1901), he also emphasized that such fibre connection pathways mainly appear between “parasensory” areas, the “association” areas in common neurological terms, as, in contrast, all primary receptive neocortical areas have connections only to adjacent areas. Nowadays the study of pathways of anatomical connectivity in vivo is achievable with DTI measurements and consecutive applied MRI tractography. The latter can be used to evaluate connectivity between voxels and to generate streamlines corresponding to estimated fiber trajectories (Basser et al. 2000; Conturo et al. 1999; Jones et al. 1999; Mori et al. 1999) or calculate maps of probability of connection.
from defined seed regions (Behrens et al. 2003a,b; Parker and Alexander 2003).

In the current state, a number of DTI studies on brain language pathways have demonstrated an unsurpassed ability to extract white fiber tracts (Catani et al. 2005; Catani et al. 2007; Catani & Mesulam 2008; Glasser and Rilling 2008; Powell et al. 2006). In reconstructing the arcuate fasciculus of the left hemisphere of right-handed healthy subjects using the two-region of interest approach, Catani et al. (2005) could show that beyond the classical arcuate pathway connecting Broca’s and Wernicke’s areas directly, there exists a previously undescribed, indirect pathway passing through inferior parietal cortex. The direct pathway runs medially and corresponds to the classical descriptions of the arcuate fasciculus. The indirect pathway runs laterally and is composed of an anterior segment connecting the inferior parietal cortex (Geschwind’s territory) with Broca’s area and a posterior segment connecting Geschwind’s with Wernicke’s area see Figure Grodd/Wildgruber/Kumar 12 in the Colour figure section). In another study Powell et al. (2006) have combined fMRI with DTI to assess hemispheric differences. Using a probabilistic tractography technique, they demonstrated consistent connections between Broca’s and Wernicke’s areas along the superior longitudinal fasciculus bilaterally, but more extensive fronto-temporal connectivity on the left than on the right. Both tract volumes and mean fractional anisotropy were significantly greater on the left. In another recent study, Catani et al. (2007) were able to show interhemispheric differences in the direct connection between Broca’s and Wernicke’s area, with extreme leftward lateralization in more than half of the subjects and bilateral symmetrical distribution in only 17.5% (Catani et al. 2007). Importantly, individuals with more symmetric patterns of connections are better overall at remembering words using semantic association. Moreover, the authors suggested that females are more likely to have a symmetrical pattern of connections.

Finally, it must be realized that DTI mapping onto single tracts is subject to a number of criticisms and has to await further methodological progress, especially with respect to creating group results, in view of our limited knowledge of human white matter anatomy – giant strides are needed to reach the same level of reliability and confidence that has been achieved in the field of fMRI.
References

Aboitiz, F. & García, V. R.

Ackermann, H., Riecker A., Mathiak K., Erb, M., Grodd, W. & Wildgruber, D.
1998 Rate dependent activation of a prefrontal-insular-cerebellar network during passive listening to train of click stimuli: an fMRI study. Neuroreport 12: 4087–4092

Ackermann, H., Wildgruber, D., Daum, I. & Grodd, W.

Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L.

Alexander, M.P., Hiltbrunner, B. & Fischer, R.S.


Baddeley, A.

Bahlo, J., Schubotz, R.I. & Friederici, A. D.
2008 Hierarchical artificial grammar processing engages Broca’s area NeuroImage 42: 525–534

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J. & Aldroubi, A.


Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C. & Samson, Y.
Belin, P. & Zatorre, R. J.
Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D.
Bogen, J. E. & Bogen, G. M.
Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P. & Frackowiak, R. S. J.
Broca, P.
Brodmann, K.
Broca, P.
Catani, M., Jones, D. K. & Ffytche, D. H.
Catani, M. & Mesulam, M.
Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A. & Michel, F.
Cohen, L., Lehericy, S., Chocon, F., Lemer, C., Rivaud, S., & Dehaene, S.  


Damasio, A. R. & Damasio, H.  

Deacon, T. W.  

Dejerine, J.  

Démonet, J-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J-L., Wise, R., Rascol, A. & Frackowiak, R.  

Démonet, J-F., Price, C., Wise, R. & Frackowiak, R. S.  

Démonet, J-F., Price, C., Wise, R. & Frackowiak, R. S.  

Démonet, J-F., Thierry G. & Cardebat, D.  

DeYoe, E. A., Bandettini, P., Neitz, J., Miller, D. & Winans, P.  

Ethofer, T., Erb, M., Anders, S., Wiethoff, S., Herbert, C., Saur, R., Grodd, W. & Wildgruber, D.  
2006a Effects of prosodic emotional intensity on activation of associative auditory cortex, NeuroReport 17, 249–253.

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W. & Wildgruber, D.  

Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P. & Wildgruber, D.  
Fiez, J. A. & Petersen, S. E.

Flechsig, P.
1901 Development (myelogenetic) localisation of the cerebral cortex in the human subject. Lancet, 2, 1027–1028

Formisano, E., De Martino, F., Bonte, M. & Goebel, R.

Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. & von Cramon, D. Y.

Geschwind, N.

Geschwind, N. & Levitsky, W.

Glasser, M. F. & Rilling, J.K.

Hart, J. & Gordon, B.


Hickok, G. & Poeppel, D.

Hickok, G. & Poeppel, D.

Howard, D., Patterson, K., Wise, R., Brown, W. D., Friston, K., Weiller, C. & Frackowiak, R.

Indefrey, P., Hellwig, F., Herzog, H., Seitz, R. J. & Hagoort, P.

Jones, D. K., Simmons, A., Williams, S.C. & Horsfield, M.A.
Kertesz, A., Sheppard, A. & MacKenzie, R.
Kim, S-G., Ashe, J., Georgopoulos, A. P., Merkle, H., Ellermann, J. M., Menon, R. S., Ogawa, S., Ugurbil, K.
Kreifelts, B., Szameitat, D., Alter, K., Szameitat, A., Sterr, A. & Wildgruber, D.
2009 Perception of different laughter types is linked to distinct neural correlates, submitted.
Levelt, W. J. M.
Lichtheim, L.
1885 On aphasia. Brain 7:433–484.
Lidzba, K., Wilke, M., Staudt, M., Krägeloh-Mann, I. & Grodd, W.
Lotze, M., Erb, M., Flor, H., Huelsmann, E., Godde, B. & Grodd, W.
2000 fMRI Evaluation of Somatotopic Representation in Human Primary Motor Cortex NeuroImage 11, 473–481
Lotze, M., Seggwies, G., Erb, M., Grodd, W. & Birbaumer, N.
2000 The representation of articulation in the primary sensorimotor cortex. NeuroReport 11:2985–2989
Majerus, S., Collette, F., Van der Linden, M., Peigneux, P., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A. & Salmon, E.
Marie, P.
Mohr, J. P.
Mori, S., Crain, B. J., Chacko, V.P. & van Zijl, P. C.
Nakamura, K., Honda, M., Okada, T., Hanakawa, T., Toma, K., Fukuyama, H., Konishi, J. & Shibasaki, H.

Noppeney, U.  

Ogawa, S., Lee, T. M., Nayak, A. S. & Glynn, P.  


Ojemann, G. A.  

Parker, G. J. & Alexander, D. C.  

Paulesu E, Frith, C. D. & Frackowiak, R. S. J.  

Petrides, M., & Pandya, D. N.  
2002 Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. European Journal of Neuroscience, 16(2), 291–310.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. & Raichle, M. E.  

Price, C. J., Wise, R. S. J. & Frackowiak, R. S. J.  

Price, C. J. & Devlin, J. T.  

Price, C. J., Wise, R. S. J. & Frackowiak, R. S. J.  

Price, C. J.  

Price, C. J.  


Rauschecker, J. P. & Tian, B.


Reiterer, S., Erb, M., Grodd, W., Wildgruber, D.

2008 Cerebral processing of timbre and loudness: fMRI evidence for a contribution of Broca’s Area to basic auditory discrimination. Brain Imaging and Behavior, 2: 1–10.


2000 Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional magnetic resonance imaging (fMRI) study. Brain and Language 75: 259–76.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G. & Grodd, W.


Riecker, A., Wildgruber, D., Dogil, G., Grodd, W. & Ackermann, H.


Riecker, A., Wildgruber, D., Mathiak, K., Grodd, W. & Ackermann, H.


Sakurai, Y., Momose, T., Iwata, M., Sudo, Y., Ohtomo, K., & Kanazawa, I.

2000 Different cortical activity in reading of Kanji words, Kana words and Kana nonwords. Cogn Brain Res 9: 111–115,

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M., Scherer, K.R. & Vuilleumier, P.


Staudt, M., Grodd, W., Niemann, G., Wildgruber, D., Erb, M. & Krägeloh-Mann, I.

2001 Early left periventricular brain lesions induce right hemispheric organization of speech. Neurology 57: 122–125

Staudt, M., Lidzba, K., Grodd, W., Wildgruber, D., Erb, M., Krägeloh-Mann, I.

Thierry, G., Giraud, A. L. & Price, C.  

Turkeltaub, P. E., Flowers, L., Lyon, L. G. & Eden, G. F.  

Ungerleider, L. G. & Mishkin, M.  

Wernicke, C.  
1874 Der aphasische Symptomenkomplex. Breslau: Cohn, Weigert.

Wessinger, C. M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J. & Rauschecker, J. P.  

Wildgruber, D., Ackermann, H., Klose, U., Kardatzki, B. & Grodd, W.  

Wildgruber, D., Kischka, U., Ackermann, H., Klose, U. & Grodd, W.  

Wildgruber, D., Ackermann, H. & Grodd, W.  


Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T. & Ackermann, H.  

Wildgruber, D., Ackermann, H., Kreifelts, B. & Ethofer, T.  

Wiesthoff, S., Wildgruber, D., Becker, H., Anders, S., Herbert, C., Grodd, W. & Ethofer, T.  

Wise, R. J., Greene, J., Buchel, C. & Scott, S. K.  
Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K. & Warburton, E. A.  
2001 Separate neural subsystems within ‘Wernicke’s area.’ Brain 124: 83–95,  
Zatorre, R. J., Bouffard, M. & Belin, P.  
2004 Sensitivity to auditory object features in human temporal neocortex.  
Brain and language talent: a synopsis

Susanne Maria Reiterer

1. Multilinguality matters

It might sound strange, but multilinguality, bilingualism, polyglottism or simply, second language learning, has become a major issue of scientific investigation in the field of brain research, specifically speaking, cognitive neuroscience. Albeit having more than hundred years of historical background, the big bang of neuroscientific investigations into the “bilingual brain”, as the most frequent cover term says, has happened only recently, within the last 15 years approximately.

Within the past few years, partly due to the rising availability of neuro-imaging facilities (functional magnetic resonance imaging, fMRI; Positron Emission Tomography, PET; Electro- and Magneto-encephalography, EEG+MEG) in the field of cognitive science research, imaging the “bilingual/multilingual” or “second language (L2) learning brain” has become increasingly popular, almost a mainstream research field in its own right.

Since Cognitive Neuroscience is composed of, or rather, influenced and fuelled by various disciplines, such as, medicine, neurology, biology, psychology, physics and engineering sciences, biosignal processing, information sciences, philosophy of mind, linguistics and language sciences, to name a few, the terminology used for research on the neuro-biological bases of second/foreign language acquisition, bilingualism and multilingualism issues became almost uniform.

All kinds of phenomena that deal in one way or the other with the process of learning more than one language or dialect at some point in time, to some degree, in some form, with either good or bad success or capacity, regardless of its final outcome, is subsumed under the umbrella term “bilingualism”. The classical, established distinctions used by psycholinguists or SLA (second language acquisition) researchers to discriminate, for instance, second language learning (the more formal classroom instruction setting) from second language acquisition (the more informal or natural surroundings setting), or on a larger scale, to discriminate bilinguals from second language learners, or, within bilingualism itself, to distin-
guish between the various forms of bilingualism (*even, uneven, consecutive, parallel, fluent bilinguals*, etc.), or within the language learners themselves (the *good* versus the *poor language learner*) have partly vanished and been replaced by a more uniform terminology.

The most frequently used terms in brain imaging studies dealing with multiple language learning, are the following collocations, which try to include information on
1. a time point (age) of onset of acquisition/learning (AOA/AOL) as well as
2. a level of fluency/proficiency/mastery of this second or further language:
   a) early and high proficiency bilinguals,
   b) late and high proficiency bilinguals;
   c) early and low proficiency and
   d) late and low proficiency bilinguals.

It has to be acknowledged that this frequent terminology is not consistently used and might lack the possibility for fine-grained distinctions, but it seems to have “emerged” under the constraints and for the benefits of interdisciplinary communication.

Depending on the field of provenience of the single researchers, the kinds of bilinguals or second language learners under scrutiny are described in more or less detail. When reading articles which span the time period between ~1990 and now, one can observe a trend towards providing always more and more details in the descriptions of the exact nature of the second language learners/bilinguals studied, including more factors of describing their “language learning history”. In the beginnings or first brain imaging studies of L2 learners, it seems that not even Age of Onset of learning was considered a crucial information to be provided, but increasingly, one can find L2-biographical information on e.g. exact nature of exposure to the languages, level of proficiency for the different subsystems of language (syntactic versus semantic or phonological processing), language learning environment and amounts or forms of training, motivation to learn, just to name a few.

Problems of misunderstanding, as in any interdisciplinary corner of science, do of course arise. For example, researchers of one field might not see why classical foreign language learners who have learned “some French” in school, are suddenly called “bilinguals” and there misunderstandings could arise in scientific writings and its audiences.
2. Brain and bilingualism: from individuality to universality and back

The reader might ask himself the question why dealing with the classical, almost historical questions in bilingual brain research should be relevant for such topics as proficiency level or even language talent. The implicit answer is threefold. On the one hand the quest into language talent, which is comparable to the search for the “language faculty” is so crucial to language learning as is the concept of DNA to all biological developmental issues and thus brings us directly to the brain. On the other hand, bilingualism research is strongly connected to individual differences in the learning process, which itself is contaminated by or composed of a myriad of factors, which will be elaborated later. Finally, individual differences of diverse polyglots were the root findings in the last century from where the offspring of today’s research took its itinerary. Thus, it will be necessary to describe some of the first findings about “polyglot” brains and the main issues treated since then, to finally be able to report on the newer and newest findings about manifestations of individual differences of language learners or users differing in their ability to learn and ultimately attain a foreign language.

Why did I call the title “from individuality to universality and back”? The first investigations of polyglot brains had been on single, individual (post mortem) brains. After a long era of case studies followed by generalizations for language recovery patterns after aphasia in multilingual patients and with the advent of group studies (which came into being much later in bilingualism brain research), generalizations into “universal” principles of normal (unimpaired) bilingual brain organization were requested and should explain the bilingual mind. Finding universal principles is of course a ubiquitous drive behind all scientific discovery, not only in this particular field of research. However, in this field it seems more difficult to establish universal principles for bilingual processing, because the bilinguals themselves are so highly variable and individual, so that their internal processes in terms of neuro-fingerprints and bio-signal-cascades will probably reflect this high individuality. This reason makes it difficult to deduct general principles of processing that are valid for all individuals in the same manner. In this way, controversial theories and findings with inconsistent results are already “preprogrammed”, as the history (past and recent) of this particular field of research testifies. Therefore, in very recent times, studies are more and more reporting on individual differences and brain organization (not only in bilinguals, but also other areas), that try to deduct principles as well, but where general-
izations are only valid for restricted subgroups of the population, namely those who share the individual features.

Research on the bilingual brain, in terms of medicine, neuropsychology or -physiology dates back to older times. It has already been investigated by neurologists in the past century (Scoresby-Jackson 1867; Pitres 1895; Minkowski 1927; Poetzl 1929). (For detailed reviews see Fabbro 2001a,b; Aglioti 1996). As already mentioned, the first cases, multilingual people who after a stroke or brain accident became aphasic, demonstrated various different recovery patterns, which were classified and generalized, but with only low to medium predictive power. Several recovery patterns have been described so far.

Amongst these are the following five most observed patterns:

1) **parallel recovery** (both languages are impaired and restored at the same rate) – **successive recovery** (a special case of 1) where one previously spoken language starts to recover only after complete recovery of the other);
2) **differential recovery** (both languages are restored differently, relative to their previous proficiency level);
3) **selective recovery** (one language recovers, the other is lost) – **antagonistic recovery** (a special case of 3) = as one language recovers, the second becomes impaired);
4) **blended recovery** (patients mix their languages inappropriately, pathological mixing and switching);
5) and finally **translation disorders**.

Translation disorders could occur only in one direction, e.g. the disorder being the inability to translate from L1 → L2, but the other way round (L2 → L1) would still be possible.

An indiosyncratic and rather amiable acquired distortion in multilingual aphasics is the phenomenon of **spontaneous translations** manifested as the compulsive need to translate everything which is being said by the patients themselves or by others around them (rudimentarily this phenomenon is even sometimes observable in normal people). Another form of translation disorder is **paradoxic translation** where patients can translate only into the language they cannot speak spontaneously (anymore) and not the reverse. One problem with all these observed patterns is that they do not show consistent localizations. For example, there can be big individual differences with the same lesion sites bringing forth different symptoms in different people.

One of the earliest rules predicting recovery after stroke was Ribot’s law (1881) – stating that recovery of the native language will be predict-
able in a multilingual patient with aphasia. A much later added, but essentially interesting law came from Kainz (Kainz 1960), and predicted the recovery of the least automatic language, arguing that the most automatic language would also be the most affected. Opposite to that, but historically the most famous and frequently reintroduced rule was the law of Pitres (Pitres 1895). It indicated that in acquired aphasia with a multilingual patient, recovery comes first and most completely in the language most used just before the injury, whether or not it is the patient’s mother tongue. Here, already at this time the importance of “fluency” level comes into play. In 1929, the Viennese neurologist, Otto Poetzl slightly reformulated this rule into “the best prediction can be made for that language of the aphasic which was the language of focus at the time when he had the accident, regardless of whether or not this was the patient’s most fluent language, not speaking at all about mother tongues” (Poetzl 1929:145). But O. Poetzl advanced another highly interesting idea, after having inspected the post mortem brain slices of various multilingual aphasics. He followed that lesions in the temporo-parietal areas of the left hemisphere might have special implications for the loss of the ability for learning new languages, and thus dubbed this area the special region for language talent, he wrote [in German] “[…] dass die Eigenleistung der hier zerstörten Region etwas zu tun hat mit der Mehrsprachigkeit bzw. mit der Anlage zu ihr, also mit dem Sprachtalent” (Poetzl 1929:153).

The fact that different recovery patterns were found and languages very often recovered separately, not in parallel, evoked a modularistic view of bilingual language representation. It gave rise to the imagination that bilinguals are different in the way their brains accommodate language, and that the various languages of a bilingual might also be “stored” or interconnected in different areas of the brain.

As early as 1867, the British neurologist, Robert Scoresby-Jackson (the nephew of the famous arctic explorer William Scoresby) surmised that Broca’s area was involved in or necessary for L1 processing, but the regions anterior to Broca for the foreign languages (L2 and L3 and so on). Over a hundred years later, this assumption was reformulated and visualized by modern brain imaging again (Kim et al. 1997). However, Pitres, already in 1895, was of the opinion that L1 and L2s use common areas. Furthermore, he assumed that after brain damage to a multilingual person, the languages are not lost, but only inhibited temporarily (effects of the lesion). This assumption was shared by Freud, Minkowski, Poetzl and Penfield.
Minkowski (1927), proposed a neurophysiological approach (instead of neuroanatomical approach) for L2 processing, which assumed essentially the same areas for L1 and L2, but with differentiated interactions acting variably upon these areas according to different linguistic constellations, as he called it. After an injury, some connections could be weakened (due to increased inhibition, raised activation threshold or unbalanced distribution of resources for the different languages).

Ever since, the most fervidly discussed or essential topic in this area has been the brain organisation of a bilingual/multilingual mind in localizationist terms. One could reformulate sloppily, “it is all about how the brain of a multilingual speaker is organized”, if it is at all different from that of a monolingual in neuromodulatory terms.

One of the crucial questions has been: Is there a common store for all languages or are there brain areas, larger networks, or even hemispheres which are specifically dedicated for each of the languages? In case of the latter possibility: Which are these activation patterns in detail and how can the different languages a speaker knows be traced and delineated in terms of brain function? Albeit the huge number of ancient (polyglot aphasia) and recent (brain imaging and neuropsychological) studies (for reviews see: Paradis 2004; Fabbro 2001a,b; Abutalebi et al. 2001; Perani & Abutalebi 2005; Halsband 2006; Wattendorf & Festmann 2008; de Boot 2008; Chee et al. 2005) the investigation of these questions has brought forth, there is still little consensus as to the exact nature of bilingual language representation, retaining it an issue of ongoing controversy, often even heated debate. The gross results of the generalizations of bilingual brain organization so far can be subsumed under 3 major viewpoints.

There is (1) the common storage viewpoint which claims a more or less precisely defined common network being responsible for the handling of all languages a speaker knows (e.g., Ojima et al. 2005; Illes et al. 1999; Klein et al. 2006; Hasegawa et al. 2002; Hernandez et al. 2001; Xue et al. 2004a).

The opposite view (2) could be called the modular view, which, in its extreme form would assign each language an own processing entity in the brain (most commonly this would be a different area or a different network responsible; e.g., Kim et al 1997; Dehaene et al. 1997; Rodriguez-Fornells et al. 2002; Kovelman et al. 2008; Perani et al. 1996).

There is also a moderate view, which could be termed the (3) partial overlap view. According to this view some areas show common activations for processing L1 and L2 and additionally other areas get activated by
the L2 (L3, L4) only (e.g., Vingerhoets et al. 2003; Marian et al. 2003; Chee et al. 2003; Lucas et al. 2004). This third view could be seen as having a second variant: the (3a) *core overlap/additional extensions* view. Under this viewpoint one could subsume all studies (e.g., Gandour et al. 2007; Hasegawa et al. 2002; Reiterer et al. 2005a,b; Briellmann et al. 2004; Chee et al. 2001; Perani et al. 2003; Yetkin et al. 1996; Yokoyama et al. 2006; Meschyan & Hernandez 2006; Abutalenbi et al. 2008a,b) which found a basic core overlap for L1 and L2 processing, with the extension of additional brain tissue (surrounding the core areas) being activated by the L2s, possibly as a function of fluency (see also caption on ‘*proficiency level*’).

One of the possible explanations why conclusions are difficult to draw and why the overall results are confusing in nature, could be the simple fact that bilinguals have very individual histories of language learning and therefore too many uncontrolled factors are in play. Confounding factors which can influence brain activation and cerebral processing of language(s) are on the one hand of a language-internal nature and regard primarily the design of the study and stimulus material. According to standard views on language processing sentence or text processing might affect brain organization differently than language processing on the word or syllable level, and yet other areas might be involved according to the computational language processes that are in question, e.g. production versus perception. (Production and perception usually have shared as well as distinct neural substrates and in some theories (e.g., the *motor theory of speech*, the two are indistinguishable)). Further differentiations regard the modularity of language (e.g., semantic, lexical, morphological, syntactic, phonological, articulatory, pragmatic levels). Most of the brain imaging studies conducted on bilinguals so far contain a variety of different stimuli and study designs, rendering comparability of results difficult (for review on bilingual studies according to different study paradigms see Gullberg & Indefrey 2006).

The second and even bigger source of confusion results from the intricate nature of the language acquisition process itself, which can be intra- and inter-individually different for each person, each social group, each language, each linguistic subsystem and each situation. This group of factors (language-external factors) is even more difficult to control for, because it would require rigorous collateral psychological testing and behavioural questioning, which in addition to elaborate brain imaging experiments is often highly time-consuming and faces practical limitations, given the amount of potentially influencing factors.
3. Factors that matter

As with every developmental process in nature, there are biological (nature) as well as environmental (nurture) factors in play which determine and shape the exact pathway of development and the outcome or success of the process, which, expressed in L2 terminology is called: ultimate attainment or level of proficiency reached. These factors influencing second language learning and acquisition will leave traces in the form of differential brain activation patterns in “different” bilinguals. Developing a simple dichotomy reflecting the above mentioned nature/nurture distinction for the sake of grouping these factors systematically (e.g., biological versus social), seems intuitive, but remains problematic. Grouping biological factors (nature part) influencing bilingual brain organization:

Here one would include: DNA, sex, hormones, handedness and age (as a consequence of brain maturation, changes in plasticity).

On the nurture side socio-cultural or linguistic factors would comprise: manner of acquisition/teaching method, amount and quality of input/training, exposure time, purpose of language use and linguistic environment, language attitudes of social group and individuals, exposure to or experience of bidialectalism and polyglottism (e.g., number of languages or registers spoken or heard, code switching habits), language type itself and linguistic subsystems (for L1 and L2). The problematic factors, however, which can neither be called purely “nature” nor “nurture”, are what we would term “psychological” factors: motivation, learning strategies/styles, domain general cognition, executive functioning, language control, intelligence/verbal intelligence, memory capacity (working memory), personality (e.g., anxiety, extraversion), empathy and language learning aptitude/talent/ability. It is questionable in how far these psychological variables are influenced by “nature” and are therefore rather predispositions, requiring renaming as “psycho-biological”, or, by “nurture” – i.e. environment and experience. This touches on a currently heated debate (monism vs. dualism) amongst neuroscientists, or scientists in general, ‘how hard wired and bio-chemical is the human mind/psyche?’ (e.g., Mohr 2003). However, the pragmatic division into “biological”, “psycho(bio)logical” and “socio-cultural-linguistic” should be sufficient (for the time being) to comprise and systematize most of the possible influencing variables.
3.1. Psychological factors influencing bilingual brain organisation

Starting out with the latter group, in the case of bilinguals, multilinguals and second language learners, little is known about the implications that motivation, individual learning styles or strategies, (verbal) intelligence and personality have on their language related brain organisation. However, some psychological aspects relating to bilingual language capacities have already been investigated with brain imaging techniques, e.g. executive functioning, cognitive control, working memory and to a much lesser degree, language ability/aptitude. With regard to executive functioning, Bialystok et al. (2005; 2007) investigated the effects of bilingualism on executive functioning and found that bilinguals develop higher levels of cognitive control, which again is reflected in their differential activation of left temporal, superior and inferior frontal areas and cingulate cortex. Executive functions, like language control are also necessary for the phenomenon of “language switching”, which, in some form, is always employed by bilinguals. The phenomenon itself has been approached by several brain imaging studies, pointing either to a sub-cortical involvement (Abutalebi et al. 2000, 2008a,b) or in most cases even to an increased involvement of non-language specific areas, like the prefrontal cortex (Rodriguez-Fornells et al. 2005; Price et al. 1999; Hernandez et al. 2000).1

General non-verbal IQ is said not to be correlated to second language learning ability (see more on this in chapter 3; pp. 67+). This has been demonstrated by an interesting rare neurological case (O’Connor & Hermelin 1991), investigating a polyglot linguist with high levels of second language proficiency for French, German and Spanish. The linguist was said to have attained a high language talent or multilingual expertise, despite the fact that he was diagnosed with a hydrocephalic brain injury and low IQ scores.

Investigating nonfluent bilinguals on a verbal working memory task Xue et al. (2004) found increased activations in a left dominated frontoparietal network for the L2. Another recent study by Chee et al. (2004), investigating phonological working memory in both, equal and unequal bilinguals, revealed more activation within the left insula and Broca’s area as well as less activation within the anterior cingulate for the equal bilinguals and less anterior cingulate activation for the unequal bilinguals’ group. They suggested that these patterns of differential activations show that more optimal engagement of phonological working memory in the group with higher language proficiency correlates also with better second

1. For a fully-fledged theory of language control see Green et al. 2006.
language attainment and discussed the terms “language attainment”, “higher language abilities”.

In a series of structural brain imaging studies using a phonetic learning paradigm of novel speech sounds, certain brain structures, primarily the left inferior parietal area, could be correlated with success (speed) of phonetic learning (Golestani et al. 2004, 2006). The authors advanced the hypothesis that brain anatomy itself could predict the success in learning foreign speech sounds. Exactly the same left parietal area has been found also by other researchers (Mechelli et al. 2004) investigating the influence of age of acquisition versus “proficiency level” in L2, to correlate with higher fluency levels (and lower age of acquisition) in bilinguals. In the beginnings of the 20th century this left inferior parietal area has already caught the attention of neurologists like Constantin von Economo and Otto Poetzl (Poetzl 1929; von Economo 1931) as being in some form connected to the phenomenon of foreign language learning talent. Another recent study (Amunts et al. 2004) explicitly investigated exceptional language talent in an interpreter who was known to have spoken over 50 languages fluently. By looking at the cytoarchitecture of his conserved brain slices they revealed significant differences between this man’s cell structure in Broca’s area as compared to normal reference brains. Thus, language talent, or high ability and a high number of languages spoken could influence brain structure. Section 4 will turn back in more detail to this last mentioned factor.

3.2. Linguistic factors

In a study by Proverbio (Proverbio et al. 2002; Evans et al. 2002) the factor polyglottism as an influencing factor has been brought into play. Proverbio and colleagues investigated how multiple languages are represented in the brain and recorded event-related potentials (ERPs) from right-handed polyglots and monolinguals during a task involving silent reading. The results showed that the multilinguals differed in their hemispheric activation patterns for semantic and syntactic processes. The fact that the bilingual speakers in this study were highly fluent and had acquired both languages in early infancy suggested that the brain activation patterns did not depend on the age of acquisition or the fluency level, as in the case of late, not-so-proficient L2 language learners, but on the functional organization of the bilinguals’ brain due to polyglotism based on brain plasticity. This led the researchers to the conclusion that multilinguality itself or the number of languages known could be one of the most powerful predictors of bilingual brain organisation.
Furthermore, language type (Klein et al. 1995; Chee et al. 1999; Ruschemeyer et al. 2005; Tham et al. 2005), grammatical complexity (Yokoyama et al. 2006), orthographical transparency (Meschyan & Hernandez 2006) as well as difference in linguistic subsystems (semantic vs. phonological: Marian et al. 2003; Pillai et al. 2003; procedural versus declarative: Ullman 2001; semantic vs. syntax: Hernandez & Li 2007; Wartenburger et al. 2003) all have been claimed to have their specific influences on the brains of speakers of more than one language. For instance, Wartenburger and colleagues (2003) investigated the effects of AOA and proficiency level on the neural correlates of grammatical and semantic processing in Italian-German bilinguals who learned the second language at different ages and had different proficiency levels. What they found was that the different subsystems of language (semantics vs. syntax) interacted differently with AOA and proficiency level on the neural substrate (measured by fMRI). While semantic processing largely interacted with or depended on proficiency level, AOA mainly affected the cortical representation of grammatical processes. This supports the view that grammar and semantics differently affect the brain organization in second language speakers.

3.3. Socio-cultural factors

A few decades ago manner of acquisition/teaching method (informal “acquiring” versus formal “learning”) was an intensely discussed issue which led to speculations about the differential engagement of the hemispheres (either more left or more right) for either the formal or the informal modes of learning (Vaid & Hall 1991). Purpose of language use, or, the impact of linguistic environment was investigated by Evans et al. (2002) with the resulting observation that lateralization in bilinguals is strongly affected by the specific language environment during development, reflected in more RH involvement for the later learned language in bilinguals brought up in areas where this language is not regularly heard.

Closely connected to language environment is the concept of exposure time, which was singled out as an important variable affecting the brain patterns in bilinguals, even in the case that both languages were acquired early and with a comparable level of proficiency (Perani et al. 2003). Not informal, but formal exposure time was the subject of a recent EEG study of our group (Reiterer et al. 2005a,b; see Fig. 1). The researchers measured the impact of amount of high-level linguistic university training (linguistic expertise) in a second language on the electrical brain synchronization patterns of university language students (L2 English; L1
German; age of onset 9 years for both groups) versus non-language students with much lower amounts (5 years difference) of formal English instruction and exposure. Results revealed characteristic differences in the alpha synchronization patterns for the groups during all language processing tasks, irrespective of the stimulus language (L1 or L2). The group with high amount of training displayed focal synchronization patterns only over specific electrode pairs of the left temporo-parietal areas in addition to a significant decrease of synchronization (against baseline) in bilateral prefrontal areas. The low training group showed widespread synchronization increases covering the entire left and partly also the right hemisphere. The results are discussed within an efficiency of processing paradigm (cortical efficiency; core overlap/partial extension view) differentiating between higher and lower levels of proficiency.

Figure 1. EEG alpha (8–12Hz) brain maps: Left column: Non language students (low proficiency group), right column: English language students (high proficiency group). Upper panel: L2 English, lower panel: L1 German. Bold lines represent synchronization increases between electrode places; dotted lines synchronization decreases (against a baseline task, e.g. looking into grey flickering screen).
Since the (generally attention-sensitive) alpha synchronization patterns\(^2\) did not discriminate along the languages, but the groups, the authors followed that either highly advanced and long-term language training at university (5–6 years) backpropagates on, reshapes or refines also the processing habits of the first language (resulting in generalized language processing strategies), or language ability, aptitude or pre-existing factors could be in play which differentiated the groups beforehand. In general, it is still an open question whether the observed brain differences in imaging studies result from genetic predispositions triggering enhanced language ability, or from structural reorganizations induced by language experience. For clarification more future research into this intricate issue is needed.

3.4. Biological factors

Evidence for the link between genetic predisposition (DNA) and first language learning is by now abundant (e.g., Bishop 1999, 2006; Fisher 2005; Fisher & Francks 2006; Dediu 2008; Dediu & Ladd 2007; Liegeois et al. 2003; Spinath et al. 2004; Vargha-Khadem et al. 2005) and it is becoming clear that there is strong evidence for a ‘language acquisition device’, an individually different language learning ability as being rudimentarily but strongly rooted in genes and successive brain development. However, genetic research connected to second language acquisition is still scarce, but a logical step towards further research, if one assumes no qualitative difference between L1 and L2. Some indirect evidence of the genetic influence on L2 comes from studies on bilingual down syndrome children, who display the same language difficulties for L1 and L2 (Bird et al. 2005), and from a twin study by Sakai et al. (2004) who found influences of grammatical L2 learning on the left dorsal inferior frontal gyrus, correlated in twins, concluding that a cortical mechanism underlying L2 acquisition also critically depends on shared genetic, not only environmental factors. Other biological factors, as mentioned earlier, are handedness and sex.

\(^2\) The EEG alpha band classically spans the frequency range from 8 to 12 Hertz and is thought to reflect primarily attention related processes. However recent research also showed that within this range one can observe other phenomena, like memory and language processing. (For more information see Reiterer et al. 2005b; Reiterer et al. 2009; Weiss et al. 2003)
Again, as with DNA, handedness playing a role in or being connected to linguistic and brain development in L1, is a widely assumed and repetitively discussed issue, however with less clear results (e.g., White et al. 1994; Foundas et al. 2002; Josse & Tzourio-Mazoyer 2004). The classical view holds that left handers would more frequently be right hemisphere dominant for language (Knecht et al. 2000), but strong claims about such ratios were also attenuated recently (Knecht et al. 2000a; Jansen et al. 2007). However, as with all studies in the field of neural representation of language, studies about the impacts of such factors on second languages and their representation are much less frequent. One of the rare studies (Andreou & Karapetsas 2004) investigated the electrophysiological responses of 30 left-handed highly proficient bilingual males in response to visually presented linguistic stimuli. The absence of statistically significant differences between left and right occipital and temporal lobes in all the components of the waveforms obtained, indicated a bilateral pattern of lateralization in left-handed males for both their native as well as their foreign language. This experiment would suggest a similar impact of handedness on L2 as it has on L1. In a purely behavioural experiment of the same group (Andreou et al. 2005), investigating handedness together with sex, L2 proficiency level and other factors (faculty choice), the results demonstrated that sex (women having performed better on semantic and syntactic tasks) and proficiency level were the most affecting factors on L2 performance, but handedness had a minor importance in explaining the results and only when combined with other factors.

Sex is another variable assumed to play a role in individual differences of language performance as well as its neuroscientific underpinnings (Hartshorne & Ullman 2006). Due to the paucity of studies for L2 in this respect, a few studies relevant for L1 will be presented here. Sex has been a further factor resulting in highly inconsistent results (Harrington & Farias 2008; Sommer et al. 2004). However, some studies did report influence of sex differences on brain organization, mostly for the phonological domain and in the form that women are less strictly left lateralized and show more bilateral diffuse language activations, whereas in males activation is more restricted and lateralized to the left hemisphere (Coney 2002; Shaywitz et al. 1995; Plante et al. 2006). Closely related to the question of sex differences, (Friederici et al. 2008) investigated sex differences together with the influence of hormones on infant brain responses to language. Using a phonological discrimination paradigm, they showed that the brain responses of few weeks old infants systematically varied as a function of biological sex and testosterone level. Females who generally
have lower testosterone levels demonstrated a clear phonological discrimination effect with a bilateral distribution. In male infants this effect systematically varied as a function of testosterone level. Males with high testosterone showed no discrimination effect, whereas males with low testosterone displayed a discrimination effect, which was clearly left-lateralized. This innovative investigation provides evidence for a strong influence of testosterone on language function and lateralization already present during the first weeks of life. Future studies would be needed to clarify whether such influences of gender or hormones are also present in brain functioning of second language processing.

Most of the attention, however, within brain imaging of second languages has been drawn towards the factor of age of acquisition (AOA). If coinciding with supposed critical or sensitive periods in brain development and loss of cortical plasticity, AOA is believed to play a major role in qualitatively distinguishing between L1 and L2 by affecting the brain processing efficiency of languages learned after that period in life in a way that is detrimental and inhibitory to native speaker like performance in late second language acquirers and learners. The vast amount of literature that has been written on that topic cannot be sufficiently reviewed here (for review see, for example, Birdsong 2006; Wattendorf & Festmann 2008). A glance at the brain imaging studies that have been conducted to that question so far gives the impression that there are at least as many studies in favour (e.g., Kim et al. 1997; Dehaene et al. 1997; Perani et al. 1996; Chee et al. 1999; Piske et al. 2002; Wartenburger et al. 2003; Mayberry & Lock 2003; Mechelli et al 2004; Zhang et al. 2005; Yokoyama et al. 2006; Hernandez & Li 2007) as against (e.g., Perani et al. 1998; Illes et al. 1999; Klein et al. 1999; Klein et al. 2002; Frenck-Mestre et al. 2005; Ojima et al. 2005; Friederici et al. 2002; Conboy & Mills 2006; Ofan & Zohary 2007) age of acquisition as being “the” ruling organizing principle in a bilingual’s brain.

Some studies also report more differentiated and partial results (e.g., Mechelli et al. 2004; Wartenburger et al. 2003; Hernandez et al. 2007; Yokoyama et al. 2006; Zhang et al. 2005; Kotz et al. 2008; Ojima et al. 2005), namely that age of onset is only observed to play a role for some part/sub-system of the language system, but not another (e.g., for syntax but not semantics, or vice versa, for phonology only, etc.) or that it plays a role only in low proficiency speakers, or that it strongly interacts with and depends on proficiency level.

The fact that AOA has stirred so many persisting speculations does not mean that it is the only influencing principle in bilingual brain organization. It was perhaps just amongst the first which were successfully ex-
plainable and amongst the most appealing because of ease of ‘operationalizability’ as an influencing variable in contrast to ‘proficiency’ as an influencing factor. However, proficiency level recently seems to have taken the lead and is amongst, if not at the time being, the preferred variable because of its explanatory power.

3.5. A critical note on the term proficiency level

Many studies looking at the brain organization of bi- or multilinguality, with or without explicitly investigating proficiency level, found in level of fluency explanatory power (e.g., Yetkin 1996; Perani et al. 1998, 2003; Chee et al. 2001; Caplan et al. 2003; Briellmann et al. 2004; Xue et al. 2004; Mechelli et al. 2004; Dodel et al. 2005; Tatsuno & Sakai 2005; Reiterer et al. 2005a,b; Ojima et al. 2005; Conboy & Mills 2006; Kotz et al. 2008; Majerus et al. 2008; Rossi et al. 2006).

A psychological concept, less well known within bilingualism research, which has already been coined as a term as early as four decades ago (Ertl 1969), called the ‘theory of cortical efficiency’ can explain many of the results obtained from studies on proficiency level as discriminating brain activations of bilinguals. This neuropsychological theory predicts that higher amounts and more distributed forms of brain activation go hand in hand with lower levels of performance in skills (not necessarily only but also applicable to language processing as well as second language learning, both being viewed as a skill; Haier et al. 1992; Just et al. 1996; Hasegawa et al. 2002).

However, a word of caution should be said about the term proficiency level itself (for further remarks on proficiency level in L2 please see also chapter by M. Jilka). Conceptually speaking, proficiency level is a misleading term. The explanatory power it consists of, could in part be explained by the fact that proficiency level is not a singular or ‘pure’, but complex factor, which functions as umbrella term and subsumes many of the other factors which have been mentioned above (like: training, aptitude, exposure time, etc.). Proficiency level can neither be regarded as biological, psychological, social nor linguistic factor because it is the sum of all. It is neutral because it is a measuring entity which expresses a degree of competence, a level of expertise. It is the measure and the outcome of the phenomenon of language learning itself, the acquisitional process measured at any given point in time.

Albeit its widespread usefulness, it is strictly speaking a fuzzy term for investigating the phenomenon of bilingualism in the brain, because it is a
sum factor reflecting the phenomenon itself, not a single variable (e.g. (AOL) that contributes to it. Thus, the comparison of “what contributes more to bilingual brain organization, age of onset or proficiency level” is a flawed one. It is like comparing “apple” (token) to “fruit” (type or category). The category “fruit” might always have more power or weight in explaining, for instance, food prevalence or consumption in a population, because it comprises a lot of single fruits. In other words, all these primary factors mentioned above (biological, psychological, social, etc.) are influencing the brain organization of the bilingual via their commonly generated product “proficiency level”.

4. Language Talent in the closer light of brain imaging

Having clarified the terminological problems with the term proficiency level as used in brain imaging studies of L2 learners, we can now turn to look in detail on the neuroimaging studies which have so far been carried out either by implicitly or explicitly looking at the question of some form of aptitude itself.

Very little is known about the manifestations of language talent in the brain. So far, there are only few imaging studies that we know of, which deal with some aspect of language talent as expressed by second language learning abilities. Only few neuro-imaging studies have explicitly investigated general foreign language talent, ability or competence. Therefore we adopt a broader view of “talent” and report here studies which use the terminologies: aptitude, ability, talent, expertise, attainment, being aware that these terms have slight differences in meaning.

Not directly a brain imaging study, but an interesting study to start out with which tried to investigate the influence of very early language experience on later language ability (Mayberry et al. 2002) provides insights into the intricate mechanisms that could bring about “talent” in learning languages, namely, the interdependence between early experience (nurture) and possible inborn qualities (nature). They showed that deaf and hearing individuals exposed to language in infancy perform comparably well in learning a new language later in life, whereas deaf individuals with little language experience in early life perform poorly later, regardless of whether the early language was signed or spoken and whether the later language was spoken or signed. These findings demonstrate that language-learning ability is also determined by the onset of language experience during early brain development, independent of the specific form of the experience.
This and genetic studies on language impairment (see section 3.4) testify to some degree the necessity for interdependence between nature and early nurture for paving the way to develop later “language talent”. Unfortunately there are no longitudinal studies about the brain development in linguistically gifted children to this date. Thus the remainder of this chapter will report on cross sectional studies performed on L2 adult or child speakers only.

Probably the first brain imaging study to really tap into individual differences in second language learning ability or “expertise” as the researchers would rather call it, was a morphometric magnetic resonance imaging study by Narly Golestani (Golestani et al. 2002) in which she investigated phonetic ability in differently “successful” L2 learners learning to identify difficult foreign phonetic contrasts. In two experiments (Golestani 2002, 2004) they found structural as well as functional neural differences correlating with the differences in learning speed and success of performance. Analysing the distribution of cortical grey and white matter structurally Golestani (2002) found higher amounts of white matter in the posterior parts of a parietal area correlating with both higher learning performance rates and speed of learning phonetic contrasts (Golestani 2002). They argued that pre-existing morphological differences, as in this case, degree of myelination, may predict language learning ability (the ease and rate of foreign language phonetic learning).

Within a similar study paradigm examining functional correlates of phonetic learning ability, they found more activity in the left angular gyrus in the more successful learners whereas in the less successful learners they found more activity within the frontal speech regions, which points to a less efficient processing strategy in the poorer learners (Golestani 2004). Generally speaking, they found similar areas to be involved in the phonetic identification process, regardless of whether a native or foreign language contrast was assessed. This means that the differences correlate more with differences in learning ability than with differences in language tasks. An implication for the theory of second language learning would be that only a small amount of phonetic training is sufficient to activate the same substrates underlying the perception of speech sounds to which one has been exposed since birth. This suggests that in principle the critical period of 6 months that is supposed to restrict later discrimination of non-native phonetic contrasts (Kuhl 1992, 2003), can be overcome by training at a later stage in adulthood, but the success and actual performance outcome is determined more by another factor: pre-existing individual differences in ability.
Similarly, in a more recent series of experiments, the same group (Golestani et al. 2007a,b) examined another group of phonetic learners using structural magnetic resonance imaging and diffusion tensor imaging. Voxel-based morphometry (VBM) indicated higher white matter (WM) density in left Heschl’s gyrus (HG) in faster learners compared to slower ones. Furthermore, they found that also white matter (WM) volume of left HG is larger in the faster learners compared with the other group. This finding was replicated in a reanalysis of the original groups tested in Golestani and others (2002). They also found that faster learners have a greater asymmetry (left > right) in parietal lobe volumes than slower learners and that the right insula and HG are more superiorly located in slower compared with faster learners. These results suggest that left auditory cortex WM anatomy, which likely reflects auditory processing efficiency, partly predicts individual differences in an aspect of language learning that relies on rapid temporal processing. They follow that a global displacement of components of a right hemispheric language network, possibly reflecting individual differences in the functional anatomy and lateralization of language processing, is predictive of speech sound learning.

In a further experiment (Golestani et al. 2007b), they tested productive phonetic ability in native French speakers who were scanned using anatomical MRI and asked them to pronounce a Persian consonant that does not exist in French and which can easily be distinguished from French speech sounds, a voiced uvular stop. The MR morphometric analysis revealed that individuals with more accurate pronunciation of the foreign sound have higher WM density in the left insula/prefrontal cortex and in the inferior parietal cortices bilaterally compared with poorer speakers or speech sound producers. These results suggest again a relationship between brain white matter anatomy and differences in pronunciation ability. Most recently, (Golestani 2008) found further traces of possible phonetic (perceptive) language aptitude in the gross morphology of the brain. They found that phoneticians have larger transverse gyrus volumes bilaterally than age and sex matched controls, and also that phoneticians are more likely than controls to have multiple or split left transverse gyri (also called Heschl’s gyri, in the left temporal cortex for primary acoustic processing).

In other words, they also find a gross morphology difference between groups, similar to what they found within healthy normal controls when comparing faster to slower phonetic learners (the faster ones were more likely than the slower to have multiple or split transverse gyri). These results suggest both gross morphological and volumetric differences in as-
pects of auditory cortex morphology between phoneticians and controls.

Another very recent brain mapping study in the phonetic speech learning domain by a Spanish researcher team (Diaz et al. 2008) provides important hints at preexisting individual differences in phonetic discrimination ability, i.e. phonetic “talent”. Event related potentials revealed individual differences between more or less “talented” or “able” phonetic perceivers not only in the foreign, but already in the native language sound system. These results point to a correlation between the native and the non-native phonetic abilities in so far as the native phonetic abilities may predict the successful learning of a new (foreign) phonetic system and this is presumed to stem from speech-specific rather than general acoustic mechanisms in phonetic mastery. This could represent another hint towards the existence of a specific phonetic ability, something we would call: phonetic language talent.

An interesting study (the only one in the domain of microstructure brain morphology, cytoarchitecture) about an outstanding general language talent was performed by the research team around Katrin Amunts (Amunts et al. 2004). They described the exceptional case of a German diplomat and interpreter in China, Emil Krebs, who was said to have known over 60 languages fluently in his life. Reportedly, he must have disposed of a strange personality, was very eager and quick to always learn new and more languages, and had creative but also strange habits of foreign language learning. For example, it is known that he learnt Armenian in only 9 weeks: 2 weeks for grammar, 3 for Old Armenian language and 4 weeks for the spoken language. After scrutinizing the cytoarchitecture (cell structure) of this multilingual’s brain slices, the researchers found out that the cell structures in area BA44 and 45 (Broca area in the left hemisphere) were significantly different than that of normal reference brains.

For more information on the capacities of outstanding language talents we want to refer to the research of Michael Erard (Erard 2005).

Turning back to functional imaging studies with a study about individual difference in the ability or speed of syntactic processing: Positron emission tomography (PET) (Waters et al. 2003) was used to determine the effect of working memory and speed of sentence processing on regional cerebral blood flow (rCBF) during syntactic processing in sentence comprehension. PET activity associated with making plausibility judgements about syntactically more complex sentences was compared to that associated with making judgements about synonymous syntactically simpler sentences. Two groups of subjects differing in working memory and
matched for speed of sentence processing both showed increases in rCBF in lateral posterioinferior frontal lobes bilaterally. These subjects were re-classified to form two groups of subjects who were matched for working memory but who differed in speed of sentence processing. Fast-performing subjects activated lateral postero-inferior frontal lobe bilaterally and slow-performing subjects showed activation of left superior temporal lobe. The results indicate that regional cerebral blood flow responses to syntactic comprehension tasks vary as a function of speed of sentence processing (a form of syntactic ability), but not as a function of working memory.

However, another fMRI study by Michael Chee and his group (Chee et al. 2004) found traces of language ability which were also correlated to working memory (for a detailed discussion of the relationship between language aptitude/ability and working memory see chapter 3). They investigated second language acquisition ability and its correlation with phonological working memory and found greater left insula and left Broca activation in high proficiency (equal) bilinguals and greater anterior cingulate activation in unequal (low proficiency or moderate) bilinguals. The activation of the left insula (see also Ackermann & Riecker 2004) was found to correlate most with high-level language attainment/ability and the better language acquisition skills themselves were found to go hand in hand with better working memory performance.

Individual differences in language ability in the semantic domain (learning new vocabulary items) have been tested by Breitenstein and colleagues (Breitenstein et al. 2005). They followed the question why some people pick up languages (in terms of new vocabulary) more easily than others. They used fMRI to elucidate which brain regions are modulated during the acquisition of a novel lexicon and which of these learning-related activity changes correlated with general semantic language knowledge. Fourteen healthy young subjects learned a novel vocabulary of 45 concrete nouns by an associative learning principle over the course of the fMRI experiment. The control condition was a no-learning task without any learning principle. Overall, increasing vocabulary proficiency was associated with modulations of activity within the left hippocampus and the left fusiform gyrus, regions involved in the binding and integration of multimodal stimuli, and with an increasing activation of the left inferior parietal cortex, the presumed neural store of phonological associations. None of these activity changes were observed during the control condition. Furthermore, subjects who showed less suppression of hippocampal activity over learning blocks scored higher on semantic
knowledge in their native language and learned the novel vocabulary more efficiently.

This possibly reflects a lexico-semantic language learning aptitude. The findings indicate that on the one hand the successful acquisition of a new lexicon depends on correlated amplitude changes between the left hippocampus and neocortical regions and on the other hand that learning-related hippocampus activity is a stable marker of individual differences in the ability to acquire and master vocabularies.

Turning back to the focus of our topic, namely aptitude in the phonetic domain, pronunciation aptitude:

The data from our own currently ongoing brain imaging research, also demonstrate that on a functional level, there are differences in brain activation between individuals that are more or less talented for L2 pronunciation. In this ongoing research (see Fig. 2, Reiterer 2 in Colour figure section) the languages to be pronounced “online” during the fMRI experiment were German as L1, English as L2 and Hindi as “L0” (a language with no previous experience at all).

Three German native speakers (mean age = 28 years) matched for age, level of education, and onset of L2 acquisition, but with different “talent” with respect to L2 pronunciation ability (as measured by our detailed elicitation techniques and speech production test battery, for detailed description please see chapters by M. Jilka) were our first cases studied.

The stimuli used were 12 pairs of German, English and Hindi words matched for number of syllables (N = 3), length (2s), semantic content, and read with neutral intonation by a male native speaker. The individuals had to carefully listen to each stimulus and repeat it by speaking into the inscanner microphone. Pronunciation of L1, L2 English, and L0 Hindi materials elicited activation patterns of a bilateral network mainly restricted to the superior temporal gyri, motor areas, insulae, basal ganglia and Left frontal areas. First level statistical analyses indicated that the extent of activation necessary to sustain the task was significantly greater for untalented speakers than for talented ones. The brain activation differences mainly correlated with the different pronunciation aptitude scores (see chapter M. Jilka for assessment of the scores) and not with the different languages (L1, L2 or L0). These results suggest that the primary factor of language aptitude (like proficiency and expertise) correlates with reduced effort in speech production, and enhanced cortical efficiency (see Fig. 1; also Reiterer et al. 2005; Reiterer et al. 2009).
In conclusion, our results of past and present research underline the significance of the concept of cortical efficiency, a concept which has been established in the domain of psychology for the learning of new skills (Haier et al. 1992; Ertl 1969; Just et al. 1996; Hasegawa et al. 2002), but is less well known in the area of second language acquisition research. However, since language learning and its gradual perfection in expertise and high level ability (talent) can be seen as a skill which is learned and improved (by the complex interplay of nature, early nurture and later training), it seems likely that the principle of efficient neural processing (from effortful to effortless) is an important lead hypothesis also in the cognitive neuroscience of individual differences of L2 processing.

The neuropsychology of talent (for more information and detailed reviews on that topic see also Dewaele 2002; Dörnyei & Skehan 2003) in general is a newly developing field and the few existing hypotheses so far are contradictory. On the one end of the continuum exists the notion that, despite having innate biological roots, which are manifested also in the brain, talent is not located in special areas or circuits of the brain (Waterhouse 1988). The classic hypothesis (Geschwind & Galaburda 1985a,b,c), on the other hand, states that an intricate interplay of genetic, hormonal and immunological influences triggers the development (delayed growth or accelerated growth) of special areas of the brain in such a way that lateralization phenomena arise. Delayed growth in one area might lead to increased growth somewhere else. Hyperdevelopment of the right hemisphere, according to this view, might enable exceptional musical talent. Generally, musical abilities, mainly musical training but to some degree also talent have already been and are still extensively investigated with brain imaging and neurophysiological techniques. There is now a growing body of literature (Andrade & Bhattacharya 2003; Anvari et al. 2002; Bentivoglio et al. 2003; Bhattacharya et al. 2001; Jäncke 2002; Koelsch 2003; Lopez 2003; Lotze 2003; Munte 2002; Pantev 2001; Schlaug 2001; Van Zuijen 2004; Zatorre 2003a,b) describing the intricate interplay between music abilities, possible other abilities (e.g., linguistic skills, timing, phonological working memory) or deficits and music training (for more information see chapter by D. Nardo). Surveys on language proficiency (Natsopoulos 2002) have reported that language proficiency is related to the degree of lateralization and handedness. Although some hypotheses about the region specific neural activation and recruitment patterns of exceptional talent are rudimentarily discussed, the brain processes involved in language talent have
yet to be characterized and understood within the fields of psychology and cognitive neuroscience (Kalbfleisch 2004), since the topic of language talent or language learning abilities has not yet been the subject of much systematic brain research.

It is, thus, at the time being a touch too early to definitely speak about special areas of language talent or involved networks or processing characteristics of special linguistic gifts. There are some reasons for this. On the one hand more large scale brain imaging studies comprising large subject pools, more longitudinal studies as well as replications hereof would still have to be performed before deducing more decisive conclusions. On the other hand language talent/aptitude itself has to be defined more closely and carefully in such studies as for defining which subcomponents of linguistic skills are in play for which a high expertise has been gained or aptitude is pre-existent. As already mentioned above, it is convenient to discriminate various subforms of “talent”: aptitude for sound discrimination (phonetic aptitude), aptitude for speech sound production (pronunciation aptitude), perhaps even special aptitude for prosody, as we investigated and found in our own present research, aptitude for syntactic relations (syntactic sensitivity, or “talent for grammar”), aptitude for vocabulary learning (lexical or semantic talent, etc.). There is ample evidence (see Dörnyei & Skehan 2003) that there are various subtypes of language learning aptitudes. Individuals can differ for specific subtypes, expressing high aptitude in one area and low aptitude in another and vice versa (partial language talent, e.g. the Joseph Conrad phenomenon), as well as high aptitude in most or even all subtypes (general or global language talent). Finally, on the other end of the ability continuum scale one can encounter also low aptitude in all or some subcomponents of language (general or partial low-talent for learning languages) as in generalized or specific developmental language impairments/syndromes.

Ultimately, defining language aptitude in its exact nature, is an endeavour, where future joint efforts of theoretical linguists (theoretical foundations of language and the language faculty), psycholinguists/psychologists (mental abilities in language) and neuroscientists (neural instantiations of language ability and its individual processing) would be required to complete the picture, enrich and evoke a common interdisciplinary theory.

Our own results indicate that talented, highly proficient learners develop efficient processing networks for language (Figure 1) and maintain/store those in small cortical areas for the efficient processing of languages that they have even never been exposed to before (see Figure Reiterer 2 in the
Colour figure section). The functional and/or anatomical connectivity among those areas allows them to confront any or most types of linguistic phenomena with efficacy and speed.

However, careful inspection of the individual L2 speakers is required in detail on a variable linguistic test battery, before being able to determine their strengths and weaknesses, their “talents” or the lack of it. These components should be investigated separately and repetitively (also by performing replication studies) before being able to deduct a theory of brain processing of foreign language learning talent or rather “talents”. However, since initially (within the last few years) a small but already respectable body of evidence has already been gathered for special traces of language talent (various subtypes), the future for this stream of research looks indeed very promising. Promising even for developing brain imaging-based tests of language talent for possible “pre-screening” purposes. These could be used for early interventions in language education, either as prevention of a language deficit or reinforcement of an already existing linguistic strength (see also chapter 11 by G Rota).

The advancement of brain imaging techniques (together with the whole field of biotechnology and biosignal analysis) could trigger a behavioural revolution for diagnosing, testing and describing specific language abilities, perhaps already at a young age. Once it is possible to detect early brain impairment and signs of delayed linguistic development in the brain, it will be equally possible to detect the opposite of the ability scale, namely linguistic aptitude.

References

Abutalebi, J.
Abutalebi, J., Cappa S, Perani D.
Abutalebi, J, Miozzo, A., Cappa, S.
Ackermann, H. & Riecker, A.
2004 ‘The contribution of the insula to motor aspects of speech production: a review and a hypothesis.’, *Brain Lang* 89 (2), 320–328.

Aglioti S, Beltramello A, Girardi F, Fabbro F.

Amunts, K., Schleicher, A. & Zilles, K.
2004 ‘Outstanding language competence and cytoarchitecture in Broca’s speech region.’, *Brain Lang* 89(2), 346–353.

Andrade P. E. & Bhattacharya J.

Andreou, G., Vlachos, F. & Andreou, E.

Andreou, G. & Karapetsas, A.

Anvari, S. H., Trainor, L. J., Woodside, J., Levy, B. A.

Bentivoglio, M.

Bhattacharya, J., Petsche, H. & Pereda, J.

Bialystok, E., Craik, F. I. M. & Freedman, M.

Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A. & Pantev, C.


Birdsong, D.

Bishop, D., Adams, C. & Norbury, C.

Bishop, D.


Caplan, D., Waters, G. & Alpert, N.
2003 ‘Effects of age and speed of processing on rCBF correlates of syntactic processing in sentence comprehension.’, *Hum Brain Mapp* 19 (2), 112–131.

Chee, M. W. L.


Chee, M. W., Soon, C. S. & Lee, H. L.

Chee, M. W., Hon, N., Lee, H. L. & Soon, C. S.

Chee, M. W., Tan, E. W. & Thiel, T.
1999 ‘Mandarin and English single word processing studied with functional magnetic resonance imaging.’, *J Neurosci* 19 (8), 3050–3056.

Chi, J. G., Dooling, E. C. & Gilles, F. H.

Conboy, B. T. & Mills, D. L.
2006 ‘Two languages, one developing brain: event-related potentials to words in bilingual toddlers.’, *Dev Sci* 9 (1), F1–12.

Coney, J.

Dediu, D.

Dediu, D. & Ladd, D. R.
2007 ‘Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin.’, *Proc Natl Acad Sci U S A* 104 (26), 10944–10949.

De Boet, K.

Dewaele, J. M.

Díaz, B., Baus, C., Escera, C., Costa, A., Sebastián-Gallés, N.

Dörnyei, Z. & Skehan, P.

Dodel, S., Golestani, N., Pallier, C., Elkouby, V., Bihan, D. L. & Poline, J.-B.

Erard, Michael.

Ertl, J. P. & Schafer, E. W.

Evans, J., Workman, L., Mayer, P. & Crowley, K.

Fabbro, F.

Fabbro, F.

Fisher, S. E. & Marcus, G. F.

Fisher, S. E.

Foundas, A. L., Leonard, C. M. & Hanna-Pladdy, B.
Frenck-Mestre, C., Anton, J. L., Roth, M., Vaid, J. & Viallet, F.  
2005 ‘Articulation in early and late bilinguals’ two languages: evidence from functional magnetic resonance imaging.’, *Neuroreport* 16 (7), 761–765.

Friederici, A. D., Pannekamp, A., Partsch, C., Ulmen, U., Oehler, K., Schmutzler, R. & Hesse, V.  
2008 ‘Sex hormone testosterone affects language organization in the infant brain.’, *Neuroreport* 19 (3), 283–286.

Friederici, A. D., Steinhauer, K. & Pfeifer, E.  
2002 ‘Brain signatures of artificial language processing: evidence challenging the critical period hypothesis.’, *Proc Natl Acad Sci USA* 99 (1), 529–534.

Geschwind, N. & Galaburda, A. M.  

Geschwind, N. & Galaburda, A. M.  
1985b Cerebral lateralization: Biological mechanisms, associations, and pathology: I. A hypothesis and a program for research. *Archives of Neurology*, 42, 564–578.

Geschwind, N. & Galaburda A. M.  
1985c Cerebral lateralization: Biological mechanisms, associations, and pathology: I. A hypothesis and a program for research. *Archives of Neurology*, 42, 634–654.

Gandour, J., Yunxia, T., Talavage, T., Wong, D., Dzemidzic, M., Yisheng, X., Xiaojian, L. & Lowe, M.  

Golestani, N., Price, C. & Scott, S.  

Golestani, N., Molko, N., Dehaene, S., LeBihan, D. & Pallier, C.  
2007 ‘Brain structure predicts the learning of foreign speech sounds.’, *Cereb Cortex* 17 (3), 575–582.

Golestani, N. & Pallier, C.  
2007a ‘Anatomical correlates of foreign speech sound production.’, *Cereb Cortex* 17 (4), 929–934.

Golestani, N., Alario, F.-X., Meriaux, S., Bihan, D. L., Dehaene, S. & Pallier, C.  
2006 ‘Syntax production in bilinguals.’, *Neuropsychologia* 44 (7), 1029–1040.

Golestani, N. & Zatorre, R. J.  

Golestani, N., Paus, T. & Zatorre, R. J.  
Brain and language talent: a synopsis


Kainz, F.

Kalbfleisch, M. L.


Klein, D., Zatorre, R. J., Chen, J.-K., Milner, B., Crane, J., Belin, P. & Bouffard, M.

Klein, D., Milner, B., Zatorre, R. J., Visca, R. & Olivier, A.
2002 ‘Cerebral organization in a right-handed trilingual patient with right-hemisphere speech: a positron emission tomography study.’, Neurocase 8 (5), 369–375.


Kotz, S., Holcomb, P. J. & Osterhout, L.

Kuhl, P. K., Tsao, F.-M. & Liu, H.-M.  

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B.  

Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M. & Vargha-Khadem, F.  

Lopez, L., Jurgens, R., Diekmann, V., Becker, W., Ried, S., Grozinger, B. & Erne, S.  

Lotze, M., Scheler, G., Tan, H. R., Braun, C. & Birbaumer, N.  

Lucas, T. H., McKhann, G. M. & Ojemann, G. A.  

Majerus, S., Belayachi, S., Smedt, B. D., Leclercq, A. L., Martinez, T., Schmidt, C., Weekes, B. & Maquet, P.  

Marian, V., Spivey, M. & Hirsch, J.  
2003 ‘Shared and separate systems in bilingual language processing: converging evidence from eyetracking and brain imaging.’, *Brain Lang* 86 (1), 70–82.

Mayberry, R. I. & Lock, E.  

Mayberry, R. I., Lock, E. & Kazmi, H.  


Meschyan, G. & Hernandez, A. E.  

Minkowski, M.  
Mohr, W. K.

Munte, T., Altenmueller, E. & Jäncke, L.

Natsopoulos, D., Koutselini, M., Kiosseoglou, G. & Kouandouris, F.

Ofan, R. H. & Zohary, E.

Ojima, S., Nakata, H. & Kakigi, R.

O’Connor, N. & Hermelin, B.

Pantev, C., Engelien, A., Candia, V. & Elbert, T.

Paradis, M.

Peper, J. S., Brouwer, R. M., Boomsma, D. I., Kahn, R. S. & Pol, H. E. H.

Perani, D. & Abutalebi, J.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F. & Fazio, F.
2003 ‘The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency.’, Hum Brain Mapp 19 (3), 170–182.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., Fazio, F. & Mehler, J.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F. & Mehler, J.


Ruschemeyer, S. A., Fiebach, C. J., Kempe, V. & Friederici, A. D.
2005 Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Hum Brain Mapp.*, 25 (2), 266–86.

Sakai, K. L., Miura, K., Narafu, N. & Muraishi, Y.
2004 ‘Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language.’, *Cereb Cortex* 14 (11), 1233–1239.

Schlaug, G.

Scoresby-Jackson, R. E.


Sommer, I. E. C., Aleman, A., Bouma, A. & Kahn, R. S.
2004 ‘Do women really have more bilateral language representation than men? A meta-analysis of functional imaging studies.’, *Brain* 127 (Pt 8), 1845–1852.

Spinath, F. M., Price, T. S., Dale, P. S. & Plomin, R.
2004 ‘The genetic and environmental origins of language disability and ability.’, *Child Dev* 75 (2), 445–454.

Tatsuno, Y. & Sakai, K. L.
2005 ‘Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands.’, *J Neurosci* 25 (7), 1637–1644.

Tham, W. W., Rickard Liow, S. J., Rajapakse, J. C., Choong Leong, T., Ng, S. E., Lim, W. E. & Ho, L. G.

Ullman, M. T.

Vaid, J. & Hall, D.

Van Zuijen, T. L., Sussman, E., Winkler, I., Naatanen, R. & Tervaniemi, M.


Zatorre, R.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M. & Tohkura, Y.
Foreign Accent Syndrome (FAS): An incidental “speech talent” following acquired brain damage?

Bettina Brendel and Hermann Ackermann

1. Early case histories

About a century ago, in 1907, the French neurologist Pierre Marie briefly commented upon a patient from the Paris area who had developed an Alsatian dialect after a stroke (cf. Whitaker 1982). The first more detailed case study of a similar phenomenon was published by Pick in 1919: A 26-year-old Czech soldier of the Austrian army – who had suffered from aphasia and right hemiplegia after a cerebrovascular event – spoke during the early period of recovery in his native language with a Polish accent. Pick noted a typical penultimate word stress pattern, clearly different from the prosodic features of the Czech language (initial stress), a shortening of long vowels, the lenisation of fricatives, and a nasalization of various sound categories. Prior to the onset of his disease, the patient had spent, however, several months of military duty in areas of a predominantly Polish-speaking population (“Österreichisch- und Russisch-Polen”). Thus, premorbid familiarity with the acquired foreign accent cannot be ruled out. As the “type specimen” of this constellation serves a case report published by Monrad-Krohn (1947). In 1941, a 30-year-old native Norwegian woman had been struck by a shrapnel on the left side of her head during an air raid over Oslo, giving rise to a traumatic brain injury with right-sided hemiplegia and “complete aphasia”. About two years later, “she walked into hospital without any noticeable limp and she spoke quite fluently but with such a decided foreign accent that I took her for German or French. She complained bitterly of constantly being taken for German in the shops, where, consequently, the assistants would sell her nothing” (page 410). Most noteworthy, the monolingual patient had had no preceding exposure to the German language. Whereas Monrad-Krohn (1947) referred to this constellation as “dysprosody”, Whitaker (1982) later coined the term “foreign accent syndrome” (FAS) – the most widely used label in the extant literature. Rarely, this acquired speech disorder has been called “pseudoaccent” (Lecours, Lhermitte & Bryans
2. Clinical characteristics and differential diagnosis

Whitaker (1982) proposed four mandatory clinical criteria of FAS: a) spoken language displays a foreign accent to the patient, her / his family members and friends as well as the examiner, b) the acquired accent clearly differs from the patient’s premorbid dialect and c) can be traced back to an acquired brain lesion, d) the patient is a monolingual speaker of his native language.

Some patients reported in the literature have fallen back to the accent of a language to which they had been exposed many years prior to acquired brain damage (“reverted accent”). (Edwards, Patel & Pople 2005; Reeves & Norton 2001; Roth, Fink, Cherney & Hall 1997; Seliger, Abrams & Horton 1992). Since, however, this constellation represents an “exclusionary factor in the identification of FAS” (Gurd & Coleman 2006, p. 425; see the fourth criterion in Whitaker 1982), these case reports will not be considered further. A FAS-like constellation sporadically has been observed in patients suffering from psychiatric conditions such as schizophrenia, dissociative disorders, or bipolar diseases (Reeves, Burke & Parker 2007; Reeves & Norton 2001; Tsuruga et al. 2008; Van Borsel, Janssens & Santens 2005; Verhoeven, Mariën, Engelborghs, D’Haenen & De Deyn 2005). Again, these constellations do not represent a FAS in its strict sense (but see Van Borsel et al. 2005), because of a “functional” rather than “organic” origin of the foreign accent. On the other hand, cases of “true” FAS may be mistaken as a psychiatric condition such as conversion disorders (Garst & Katz 2006).

Although a diagnosis of FAS – by definition – should be applied only to monolingual speakers (see above), it is noteworthy that the acquired foreign accent of a polyglot subject (native language = Spanish, second languages = French, English, Catalan) – sounding like French – affected exclusively her “mother tongue”, but not the second languages (Avila, González, Parcet & Belloch 2004; Gonzales-Alvarez et al. 2003).

FAS patients of the same linguistic background may display a (perceived) sound shift towards different language systems. For example, a change to Dutch, German, French, Spanish, Italian, Russian or Swedish accents has been reported in monolingual (American) English subjects. In many instances, however, there was little agreement among listeners.
about the origin of the "unlearned" accent. Thus, the accent of a native American English FAS patient – described in a paper by Laures-Gore and coworkers (2006) – has been categorized as Chinese, Dutch or Canadian, respectively, by different subjects. As another example, a native American English patient sounded either like a Slavic, Eastern European, French, or Nordic speaker (Blumstein, Alexander, Ryalls, Katz & Dworetzky 1987). Apart from a shift to another language, dialectal changes have also been observed after acquired brain damage. Thus, e.g., FAS patients with a background of American or Canadian English were reported to resemble speakers from England, Scotland, Australia or even from another part of Canada (Naidoo, Warriner, Oczkowski, Sévigny & Humphreys 2008). In consideration of this variability, Miller, Lowit & O’Sullivan (2006) suggested that the perceived foreign accent – which usually does not influence the intelligibility of a speaker – might depend upon the preceding experiences of a listener with other languages. In other words, the foreign accent might reside “in the ear of the listener” as “different listeners focus on different aspects of speech” (Miller et al. 2006, p. 403).

A recent survey found speech and language therapists in the United Kingdom to encounter on average just a single FAS patient within a time span of ten years (Coleman & Gurd 2006). Thus, this constellation is characterized by a rather low incidence, and – apart from the above-mentioned “historical reports” – only about 50 more or less detailed case studies have been published so far, at the exclusion of patients with a psychiatric background or subjects with previous exposure to the foreign accent emerging after acquired brain damage. Cerebrovascular disorders (35 cases) and traumatic brain injuries (9 cases) represent the most frequent etiological variants of this constellation (Table 1). More rarely, cerebral metastases of a breast tumour (Abel et al. 2009), multiple sclerosis (Bakker et al. 2004; Villaverde-González et al. 2003), or a progressive degenerative brain disease (Luzzi et al. 2008) have been documented. Most noteworthy, the foreign accent showed an abrupt onset even in neurodegenerative disorders. In four patients, an unambiguous medical diagnosis could not be established (Coelho & Robb 2001; Gurd, Coleman, Costello & Marshall 2001; Katz, Garst & Levitt 2008; Poulin, Macoir, Paquet, Fossard & Gagnon 2007).

As concerns lesion site, Blumstein & Kurowski (2006) assume “that the foreign accent syndrome emerges as a consequence of damage to the speech output motor system affecting the primary motor cortex and either cortico-cortical connections with it or its cortico-subcortical projections” (p. 352). Whereas this suggestion appears to hold for the major-
ity of the published cases, FAS has been found in some instances to be bound to lesions of the right cerebral hemisphere (Berthier, Ruiz, Massone, Starkstein & Leiguarda 1991; Dankovičová, Gurd, Marshall, MacMahon, Stuart-Smith & Coleman 2001; Gonzales-Alvarez et al. 2003; Laures-Gore et al. 2006; Miller et al. 2006), the corpus callosum (Hall, Anderson, Filley, Newcombe & Hughes 2003), or left-sided parietal cortex (Abel et al. 2009; Laures-Gore et al. 2006; Roth et al. 1997). Given a predominant association with damage to the anterior perisylvian cortex of the left hemisphere and the speech output motor system, a co-occurrence of FAS with other acquired disorders of speech production must be expected. Indeed, a recent review observed nearly two thirds of the FAS patients reported in the literature to have also suffered from a syndrome of non-fluent aphasia, and a quarter of the sample showed signs of apraxia of speech (Coleman & Gurd 2006). By contrast, dysarthria was rarely found associated with FAS, most probably, because of a more bilateral organization of the speech motor control network at the level of primary motor cortex and / or the cortico-bulbar tracts (see Ackermann & Ziegler 2009). Several case studies noted mutism or severe aphasic deficits (i.e. agrammatism) of spoken language to precede the emergence of an unlearned accent after acquired brain damage. As concerns the subsequent course of this constellation, FAS may represent a transient phenomenon, lasting for a few days or weeks only, or to persist over years.

So far, only anecdotal and informal data on the effects of speech therapy in FAS patients is available (e.g. Blumstein et al. 1987). For example, Garst and Katz (2006) observed a beneficial impact of “stress and timing work, using a pacing board to improve word- and sentence-level prosody”. Another subject found a mirror helpful which enhanced guidance of lip and jaw placement during speech production.

3. Phonetic features

The phonetic characteristics of the acoustic speech signal can be assigned to a segmental and a suprasegmental level, respectively. Segmental features describe the acoustic correlates of single speech sound categories, i.e. vowels and consonants. By contrast, suprasegmental or prosodic features refer to larger constituents such as syllables, words, and phrases (e.g. Sidtis & Van Lancker Sidtis 2003). The acoustic correlates of prosodic features comprise several durational aspects of speech utterances like “syllable, word, phrase, and breath group length, pausing, phrase final
lengthening, tempo, and rate” (Van Lancker Sidtis, Pachana, Cummings & Sidtis 2006, p. 135), loudness (amplitude / intensity), pitch (fundamental frequency) and voice quality.

3.1. Segmental level: consonant and vowel production

A series of studies noted voicing errors such as the devoicing of voiced stops and fricatives, including devoicing of final obstruents, or voicing of voiceless consonants as well as an exaggerated prevocing of voiced stops (Ardila, Rosselli & Ardila 1988; Blumstein et al. 1987; Coehlo et al. 2001; Gurd, Bessell, Bladon, & Bamford 1988; Ingram, McGormack, Kennedy 1992; Laures-Gore et al. 2006; Mariën & Verhoeven 2007; Miller et al. 2006; Van Borsel et al. 2005). Taken together, so far only a few investigations failed to document any voicing errors (Dankovičová et al. 2001; Katz et al. 2008; Kurowski, Blumstein & Alexander 1996; Moonis et al. 1993). As a further salient phonetic feature, fortitions, i.e. hyper-aspiration, exaggerated frication, stopping of fricatives and affrication of plosives, have been frequently reported (Christoph et al. 2004; Gurd et al. 1988, 2001; Ingram et al. 1992; Lippert-Gruener, Weinert, Greisbach & Wedekind 2005; Mariën & Verhoeven 2007; Scott, Clegg, Rudge & Burgess 2006;). However, lenisation processes or articulatory undershoot like hypo-aspiration, de-affrication, or reduced frication may emerge as well (Gurd et al. 1988; Miller et al. 2006; Pick 1919). Several case studies of native speakers of American English were able to document sound substitutions in terms of the production of alveolar flaps within a vowel-consonant-vowel (VCV) environment (e.g. “better”) as full stops (e.g. Bakker et al. 2004; Blumstein et al. 1987; Katz et al. 2008; Kurowski et al. 1996; Laures-Gore et al. 2006). Generally, manner errors appear to dominate over changes in place of articulation (see Katz et al. 2008 or Moen 2000).

Both shifts in vowel quality and length have been noted in FAS patients, and perceptual as well as formant analyses point at changes of vowel space in terms either of more lax (Carbary, Patterson & Snyder 2000; Kurowski et al. 1996; Lippert-Gruener et al. 2005; Whitaker 1982) or more tense productions (Blumstein et al. 1987; Graff-Radford et al. 1986; Ingram et al. 1992; Laures-Gore et al. 2006; Moonis et al. 1993). In some analogy to these alterations, a shortening of long / stressed syllables (Moen 1990; Pick 1919) as well as unstressed / short vowels could be noted (Dankovičová et al. 2001). Taken together, however, the prolongation or lengthening of (unstressed) vowels represents the more salient
durational change in FAS patients (e.g. Ardila et al. 1988; Carbary et al. 2000; Christoph et al. 2004; Graff-Radford et al. 1986; Ingram et al. 1992; Katz et al. 2008; Lippert-Gruener et al. 2005; Miller et al. 2006; Naidoo et al. 2008; Reeves et al. 2007; Ryalls & Whiteside 2006). In addition, diphthongizations (Ardila et al. 1988; Graff-Radford 1988; Lippert-Gruener et al. 2005, Naidoo et al. 2008) and monophthongizations of vowels (Dankovičová et al. 2001; Kurowski et al. 1996; Whittaker 1982) as well as reduced (Moen 1990, 2000) or more prominent lip rounding (Dankovičová et al. 2001; Miller et al. 2006) have been reported.

3.2. Suprasegmental (prosodic) features of spoken language

Depending on the kind of information transferred, the suprasegmental level encompasses linguistic (word stress, pitch accent, sentence mode, constituent boundaries), pragmatic (context-dependent aspects of meaning), affective (emotion, arousal), indexical (speaker-specific), or attitudinal aspects of spoken language (Van Lancker Sidtis et al. 2006). Some case studies of FAS explicitly mention affective / emotive prosody to be spared (Blumstein et al. 1987; Carbary et al. 2000; Christoph et al. 2004; Dankovičová et al. 2001; Hall et al. 2003; Katz et al. 2008; Miller et al. 2006). In contrast, a deviant pattern of linguistic prosody represents the core sign of FAS and has been observed in nearly all cases (see Blumstein & Kurowski 2006), and only three studies reported an unremarkable suprasegmental structure of verbal utterances (Moonis et al. 1993; Kurowski et al. 1996; Dankovičová et al. 2001). Patients are reported to produce abnormal “global” (sentence level) and “focal” (word / syllable level) intonation contours. At the sentence level, unusual intonation patterns such as a sharp phrase-final rise of the fundamental frequency (F₀) contour (Blumstein et al. 1987; Monrad-Krohn 1947; Moonis et al. 1993) or an exaggerated, abrupt pitch fall (Ingram et al. 1992) at this location could be observed, even pitch inversions have been noted (Takayama, Sugishita, Kido, Ogawa & Akiyama 1993). FAS speakers may also exhibit a restricted F₀ range (Graff-Radford et al. 1986; Mariën & Verhoeven 2007) or an inappropriate high variability of F₀ (Blumstein et al. 1987; Gurd et al. 2001; Laures-Gore et al. 2006; Moonis et al. 1996).

Nearly all case studies mention prolonged pauses between words and syllables which, presumably, contribute to perceived syllable segregation (Laures-Gore et al. 2006) or a staccato-like scanning speech production (Ingram et al. 1992, Hall et al. 2003; Moen 1990; Van Borsel et al. 2005). As a frequently observed prosodic abnormality, the shift of a stress-
timed – as in English or German – to a more syllable-timed rhythm and/or a tendency towards syllable isochrony have been noted (Blumstein et al. 1987; Coughlan, Lawson & O’Neill 2004; Gurd et al. 2001; Katz et al. 2008; Miller et al. 2006; Mariën & Verhoeven 2007; Varley, Whiteside, Hammill & Cooper 2006). These changes in syllabic structure are frequently caused by the insertion of an epenthetic vowel such as a schwa (Ardila et al. 1988; Blumstein et al. 1987; Ingram et al. 1992; Kurowski et al. 1996; Laures-Gore et al. 2006; Miller et al. 2006; Naidoo et al. 2008).

Ingram and coworkers (1990) performed a multiple regression analysis to determine the most salient phonetic-linguistic “features” associated with the perceived foreign accent in their patient. Four characteristics contributed to the impression of “foreignness”, i.e. terminal Fo changes, final obstruent devoicing, consonant strengthening (fortition), and shifts in vowel quality. In addition, Miller et al. (2006) emphasize altered speech rhythm in terms of the above-mentioned vowel epenthesis.

As a most important finding, the measured acoustic-phonetic features of a FAS patient’s verbal utterances often do not correspond to the perceived “foreign” accent or dialect. For example, Moonis et al. (1993) as well as Blumstein et al. (1987) report that the VOT values obtained from a French-sounding FAS subject, surprisingly, were found to be located within the range of English and not French stops. Furthermore, Dankovičová et al. (2001) investigated the phonetic features of a patient with Scottish-sounding accent and they were able to demonstrate that the perceived “Scottishness” mainly reflected segmental errors, only some of them representing typical features of that dialect. In other words, there is little evidence that the produced features are typical of the perceived particular sound system (Dankovičová et al. 2001; Gurd et al. 1988; Kurowski et al. 1996). This fuels the discussion whether FAS subjects produce a particular foreign accent or if the accent is rather “generic”, i.e. the accent is identified only as foreign but can not be assigned to any particular language (Blumstein et al. 1987; DiDio, Schulz & Gurd 2006; Miller et al. 2006).

3.3. Phonetic-linguistic meta-analyses

Any meta-analysis of the sound structure of the spoken language in FAS patients faces at least two limitations:

(a) The published case studies cover a very broad range of native languages, characterized by quite different sound structures each, including British, Irish, Scottish, American, and Australian English as well as
Spanish, Portuguese, French, Norwegian, Dutch, German, Japanese, and Mandarin Chinese.

(b) Both test materials and depth of analysis vary considerably across the extant investigations, extending from the “impressionistic” description of a patient’s verbal utterances to perceptual ratings conducted by independent listeners, to broad and narrow phonetic transcriptions, to detailed acoustic-instrumental assessments and, finally, to a combination of all these techniques (e.g. Avila et al., 2004; Blumstein et al., 1987; Di Dio et al., 2006; Miller et al., 2006; Varley et al., 2006).

Due to the variety of the native languages of the patients, the applied test material and analysis methods, the available data do not yet provide a coherent acoustic-phonetic description of FAS.

4. Psycho- and neurolinguistic models

Rather than a disorder per se, FAS has been assumed to reflect the compensation of or the adaptation to co-existing speech, language, or voice deficits. Against this suggestion argues, however, that FAS emerges early during the recovery process of neurological patients (Kurowski et al. 1996). Hence, this constellation appears to be a direct sequel of brain damage. It is still a matter of debate, however, whether FAS constitutes an independent constellation (Blumstein et al. 1987; Blumstein & Kurowski 2006; Coleman & Gurd 2006; Ingram et al. 1992) or rather a “primarily listener-bound” sequel of other syndromes of speech / language pathology (Van Borsel et al. 2005). Whereas Ardila and co-workers (1988) maintain FAS to reflect an epiphenomenon of aphasic deficits, other authors (Aronson 1990; Moen 2000; Naidoo et al. 2008; Mariën, Verhoeven, Engelborghs, Rooker, Pickut & De Deyn 2006; Varley et al. 2006) conclude that the observed error pattern in FAS patients resembles the segmental and suprasegmental deficits of apraxia of speech, considered to be a disorder of speech motor programming. As an example, Varley et al. (2006) assume this syndrome to reflect a disruption of automatized speech control processes in the sense that the “activation of phonetic gestalts for frequently encountered utterance units” (p. 366) is compromised. Consequently, these authors propose that FAS represents a subtype or a mild form of apraxia of speech. Speech output of apraxic patients is laborious, slow, often unintelligible, but most importantly, verbal utterances give rise to a “pathological”, not “foreign” character (Blumstein & Kurowski 2006). In contrast, spoken language of FAS pa-
tients is not perceived as disrupted and intelligibility is not or only mildly impaired. It has been suggested, therefore, that FAS does not violate phonetic rules of natural languages – in contrast to apraxia of speech or aphasia (Blumstein et al. 1987; Blumstein & Kurowski 2006, Miller et al. 2006). Hence, FAS might represent primarily a genuine disorder of linguistic prosody resulting from disturbed speech timing and the observed segmental errors are just a consequence of the prosodic abnormalities (Blumstein et al. 1987; Blumstein & Kurowski 2006; Katz et al. 2008).

Alternative models of the pathomechanism(s) of FAS assume an abnormally tense speech posture or a “misadjustment” of the “phonetic” or “vocal” settings of the vocal apparatus (Graff-Radford et al. 1986; Ingram et al. 1992; Naidoo et al. 2008). These suggestions trace back to a theory of speech production which assumes that the tension of the vocal tract, required for speech production, is specified by long-term muscular adjustments (cf. Ingram et al. 1992). The degree of pre-adjusted tension – varying across different languages – will affect, e.g., the amplitude of tongue and jaw movements which in turn influence the form and the size of the vowel space. For example, tense jaw movements will give rise to a more closed position / lower F1-values. By contrast, Moen (2006) explained FAS “in term of the framework of gestural phonology”. The theory of gestural or articulatory phonology – as proposed by Browman and Goldstein (1992) – considers gestures (movement patterns) to be the basic units for achieving particular articulatory targets. According to this model, the observed segmental and suprasegmental errors in FAS can be derived completely from two computational processes, i.e. an abnormal scaling and phasing of articulatory gestures. Within this context, articulatory undershoot and fortition processes reflect distorted scaling operations, whereas voicing errors (temporal misalignment of glottal and oral gestures) can be assigned to phasing problems.

5. Summary

FAS represents a rather rare speech disorder subsequent to acquired brain damage, in the vast majority of the reported cases a cerebrovascular event or traumatic brain injury. This constellation is characterized by the perceptual impression that a speaker produces his native language with a foreign accent. As a rule, intelligibility of spoken utterances remains uncompromised. Altogether, about 60 case studies have been published so far, including patients with a disorder which does not match
FAS in its strict sense. The available data do not yet provide a coherent picture, neither with respect to aetiology, lesion site, phonetic-linguistic profile, nor the underlying pathomechanisms. Among others, it remains to be established whether FAS is a “true disorder in its own” or just a phenomenon “in the ear of the listener” and whether various subtypes of this constellation must be separated (Garst & Katz, 2006).

References


Coelho, C. A., & Robb, M. P. 

Coleman, J. S., & Gurd, J. M. 

Coughlan, T, Lawson, S., & O’Neill, D. 

Critchley, M. 

Dankovičová, J., Gurd, J., Marshall, J., MacMahon, M., Stuart-Smith, J., Coleman, J., & Slater, D. 

DiDio, C., Schulz, J., & Gurd, J. M. 

Edwards, R. J., Patel, N. K., & Pople, I. K. 

Fridriksson, J., Ryalls, J., Rorden, C., Morgan, P. S., George, M. S., & Baylis, G. C. 

Garst, D., & Katz W. 

González-Alvarez, J., Parcet-Ibars, M. A., Avila, C., & Geffner-Sclarsky, D. 

Graff-Radford, N. R., Cooper, W. E., Colsher, P. L., & Damasio, A. R. 

Gurd, J. M., Bessell, N. J., Bladon, R. A. W., & Bamford, J. M. 

Gurd, J. M., & Coleman, J. S. 
Gurd, J. M., Coleman, J. S., Costello, A., & Marshall, J. C.  

Hall, D. A., Anderson, C. A., Filley, C. M., Newcombe, J., & Hughes, R. L.  

Hwang, C.-S., Lin, M.-H., & Lin S.-K.  

Ingram, J. C. L., McCormack, P. F., & Kennedy, M.  

Katz, W. F., Garst, D. M., & Levitt, J.  

Kurowski, K. M., Blumstein, S. E., & Alexander, M.  

Laures-Gore, J., Henson, J. C., Weismer, G., & Rambow, M.  

Lecours, A. R., Lhermitte, F., & Bryans, B.  

Lippert-Gruener, M., Weinert, U., Greisbach, T., & Wedekind, C.  

Luzzi, S., Viticchi, G., Piccirilli, M., Fabi, K., Pesallaccia, M., Bartolini, M., Provinciali, L., & Snowden J. S.  

Mariën, P., Verhoeven, J., Engelborghs, S., Rooker, S., Pickut, B. A., & De Deyn, P. P.  

Mariën, P., & Verhoeven, J.  

Miller, N., Lowit, A., & O’Sullivan, H.  

Moen, I.  
Moen, I.

Moen, I.


Monrad-Krohn, G. H.

Naidoo, R., Warriner, E. M., Oczkowski, W. J., Sévigny, A., & Humphreys, K. R.

Neilson, J. M., & McKeon, A. B.

Pick, A.

Poulin, S., Macoir, J., Paquet, N., Fossard, M., & Gagnon, L.

Reeves, R. R., Burke, R. S., & Parker, J. D.

Reeves, R. R., & Norton, J. W.

Roth, E. J., Fink, K., Cherney, L. R., & Hall, K. D.

Ryalls, J., & Whiteside, J.

Schiff, H. B., Alexander, M. P., Naeser, M. A., & Galaburda, A. M.

Scott, S. K., Clegg, F., Rudge, P., & Burgess, P.
Seliger, G. M, Abrams, G. M., & Horton, A.  

Sidtis, J. J., & Van Lancker Sidtis, D.  

Takayama, Y., Sugishita, M., Kido, T., Ogawa, M., & Akitiuchi, I.  

Tsuruga, K., Kobayashi, T., Hirai, N., & Kato, S.  

Van Borsel, J., Janssens, L, & Santens, P.  

Van Lancker Sidtis, D., Pachana, N., Cummings, J. L., & Sodtis, J.  

Varley R., Whiteside, S., Hammill, C., & Cooper, K.  

Verhoeven, J., Mariën, P., Engelborghs, S., D’Haenen, H., & De Deyn, P.  

Villaverde-González, R., Fernández-Villalba, E., Moreno-Escribano, A., Aliás-Linares, E., & García-Santos, J. M.  

Wendt, B., Bose, I., Sailer, M., Scheich, H., & Ackermann, H.  

Whitaker, H.  

Whitty, C. W. M.  
<table>
<thead>
<tr>
<th>Authors</th>
<th>Patient</th>
<th>Native language</th>
<th>Foreign accent</th>
<th>Changes of phonetic features</th>
<th>Lesion site</th>
<th>Aetiology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pick, 1919</td>
<td>26, m</td>
<td>Czech</td>
<td>Polish</td>
<td>+/+/+</td>
<td>left hemisphere</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Whitely, 1964</td>
<td>27, f</td>
<td>British</td>
<td>German</td>
<td>+/</td>
<td>/+</td>
<td>left pre-rolandic area just above the sylvian fissure</td>
</tr>
<tr>
<td>Whitaker, 1982</td>
<td>30, f</td>
<td>American</td>
<td>Spanish/German</td>
<td>+/+/+</td>
<td>left hemisphere</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Schiff et al., 1983 (cf. Blumstein et al., 1987)</td>
<td>60, m</td>
<td>Portuguese-American</td>
<td>Chinese</td>
<td>?/?/?</td>
<td>left premotor region/deep white matter (anterior and superior to the head of the left caudate nucleus)</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Graff-Radford et al., 1986</td>
<td>56, f</td>
<td>American</td>
<td>Nordic</td>
<td>–/+/+</td>
<td>left precentral gyrus/subjacent white matter</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Blumstein et al., 1987</td>
<td>62, f</td>
<td>American</td>
<td>Eastern European, Slavic, French, Nordic</td>
<td>+/+/+</td>
<td>left subcortical white matter of the pre-rolandic and post-rolandic gyr; enlarged internal capsule, decreased density in left lateral putamen</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Ardila et al., 1988</td>
<td>26, m</td>
<td>Spanish</td>
<td>English</td>
<td>+/+/+</td>
<td>left Broca's area</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Gurd et al., 1988</td>
<td>41, f</td>
<td>British</td>
<td>French</td>
<td>+/+/+</td>
<td>left basal ganglia/lentiform nucleus/lateral to the genu of left internal capsule</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Moen, 1990, 2006</td>
<td>61, f</td>
<td>Norwegian</td>
<td>English</td>
<td>+/+/+</td>
<td>left hemisphere</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Berthier et al., 1991</td>
<td>70, f</td>
<td>Spanish</td>
<td>“foreign”</td>
<td>+/+/+</td>
<td>left precentral gyrus, subjacent white matter</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td></td>
<td>58, m</td>
<td>Spanish</td>
<td>Slavic</td>
<td>+/+/+</td>
<td>right fronto-parietal white matter</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td>Authors</td>
<td>Patient</td>
<td>Native language</td>
<td>Foreign accent</td>
<td>Changes of phonetic features</td>
<td>Lesion site</td>
<td>Aetiology</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td></td>
<td>47, f</td>
<td>Spanish</td>
<td>Hungarian</td>
<td></td>
<td>left middle frontal gyrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34, f</td>
<td>Spanish</td>
<td>Slavic</td>
<td></td>
<td>left middle frontal gyrus</td>
<td></td>
</tr>
<tr>
<td>Ingram et al., 1992</td>
<td>56, f</td>
<td>Australian</td>
<td>Asian,</td>
<td></td>
<td>lentiform nucleus</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Swedish, German</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seliger et al., 1992</td>
<td>65, f</td>
<td>American</td>
<td>Irish</td>
<td></td>
<td>deep left hemisphere stroke; subcortical</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>infarction</td>
<td></td>
</tr>
<tr>
<td>Takayama et al., 1993</td>
<td>44, f</td>
<td>Japanese</td>
<td>Korean</td>
<td></td>
<td>posterior lateral aspect of the left</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>precentral gyrus</td>
<td></td>
</tr>
<tr>
<td>Kurowski et al., 1996</td>
<td>45, m</td>
<td>American</td>
<td>British,</td>
<td></td>
<td>left posterior supramarginal gyrus;</td>
<td>CVA cerebrovascular</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Scottish, Irish,Eastern</td>
<td></td>
<td>more extensive lesion in left subcortical</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>European</td>
<td></td>
<td>structures (lentico-lostriate territory incl.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>putamen, dorsolateral caudate, the lateral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>anterior limb of the internal capsule and anterior</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>periventricular white matter.</td>
<td></td>
</tr>
<tr>
<td>Roth et al., 1997</td>
<td>45, m</td>
<td>American</td>
<td>Dutch</td>
<td></td>
<td>left parietal</td>
<td></td>
</tr>
<tr>
<td>Dankovicová et al., 2001</td>
<td>43, f</td>
<td>British</td>
<td>Scottish</td>
<td></td>
<td>extensive lesion in the territory of the</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>right middle cerebral artery (incl. basal system</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>and right Sylvian fissure)</td>
<td></td>
</tr>
<tr>
<td>Hwang et al., 2001</td>
<td>40, f</td>
<td>Mandarin</td>
<td>American</td>
<td></td>
<td>MRT no lesions; hypoperfusion in the left</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>lateral temporal region with contralateral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>cerebellar diaschisis</td>
<td></td>
</tr>
<tr>
<td>Hall et al., 2003</td>
<td>53, f</td>
<td>American</td>
<td>French, French-</td>
<td></td>
<td>corpus callosum (slightly extending to the</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Canadian</td>
<td></td>
<td>left)</td>
<td></td>
</tr>
<tr>
<td>Coughlan et al., 2004</td>
<td>39, f</td>
<td>Irish</td>
<td>French</td>
<td></td>
<td>left internal capsule</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Patient</td>
<td>Native language</td>
<td>Foreign accent</td>
<td>Changes of phonetic features</td>
<td>Lesion site</td>
<td>Aetiology</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
<td>-----------------</td>
<td>----------------</td>
<td>------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Christoph et al., 2004</td>
<td>28, m</td>
<td>Brazilian Portugese</td>
<td>Generic, American</td>
<td>+ +</td>
<td>left pre-rolandic cortex extending to underlying white matter and inclusion of insula and frontal operculum</td>
<td></td>
</tr>
<tr>
<td>Edwards et al., 2005</td>
<td>58, m</td>
<td>English</td>
<td>Irish</td>
<td>/ /</td>
<td>left basal ganglia</td>
<td>CVA</td>
</tr>
<tr>
<td>Edwards et al., 2005</td>
<td>64, f</td>
<td>Scottish</td>
<td>Dutch, Swedish, Russian, German</td>
<td>/ /</td>
<td>left basal ganglia; internal capsule</td>
<td></td>
</tr>
<tr>
<td>Fridriksson et al., 2005</td>
<td>45, m</td>
<td>American</td>
<td>French, Greek, British</td>
<td>/ /</td>
<td>left basal ganglia</td>
<td>CVA</td>
</tr>
<tr>
<td>Laures-Gore et al., 2006</td>
<td>64, m</td>
<td>American</td>
<td>Chinese, Dutch, Canadian</td>
<td>+ +</td>
<td>right posterior temporal parietal cortex; old infarct in the left basal ganglia and left cerebellum</td>
<td>Cerebrovascular accident</td>
</tr>
<tr>
<td></td>
<td>67, f</td>
<td>American</td>
<td>Spanish, Jamaican</td>
<td>+ +</td>
<td>unknown</td>
<td></td>
</tr>
<tr>
<td>Miller et al., 2006</td>
<td>60, f</td>
<td>British</td>
<td>Italian, occasionally Eastern European</td>
<td>+ +</td>
<td>aneurysm in the anterior communicating artery (junction of A1 – A2); fronto-parietal craniotomy with displacement of right gyrus rectus; mild ventricular dilation with several small areas of low attenuation in the supratentorial white matter in both hemispheres; no focal damage</td>
<td></td>
</tr>
<tr>
<td>Ryalls &amp; Whiteside, 2006</td>
<td>57, f</td>
<td>American</td>
<td>British, Australian</td>
<td>/ /</td>
<td>left internal capsule</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Patient Gender</td>
<td>Native Language</td>
<td>Foreign Accent</td>
<td>Changes of Phonetic Features</td>
<td>Lesion Site</td>
<td>Aetiology</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------</td>
<td>----------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Scott et al., 2006</td>
<td>64, f</td>
<td>Scottish</td>
<td>German, Polish, South African</td>
<td>+ + +</td>
<td>left white matter underneath the pre-central sulcus</td>
<td></td>
</tr>
<tr>
<td>Varley et al., 2006</td>
<td>40, f</td>
<td>British</td>
<td>Swedish</td>
<td>+ + /</td>
<td>anterior portions of left middle cerebral artery; haemorrhagic changes within left putamen</td>
<td></td>
</tr>
<tr>
<td>Marien &amp; Verhoeven, 2007 / Marien et al., 2006</td>
<td>53, f</td>
<td>Dutch</td>
<td>French, German</td>
<td>+ + +</td>
<td>left fronto-parietal region; incl. IFG, precentral gyrus, anterior insula, postcentral gyrus, supramarginalis gyrus</td>
<td>CVA cerebrovascular accident</td>
</tr>
<tr>
<td></td>
<td>61, m</td>
<td>Dutch</td>
<td>distinctly foreign, North-African</td>
<td>+ – +</td>
<td>left basal ganglia, incl. posterior insula, medial temporal lobe and paraventricular white matter of left parietal lobe</td>
<td></td>
</tr>
<tr>
<td>Wendt et al., 2007</td>
<td>35, f</td>
<td>German</td>
<td>Russian</td>
<td>/ / +</td>
<td>area of left middle cerebral artery</td>
<td></td>
</tr>
<tr>
<td>Naidoo et al., 2008</td>
<td>50, f</td>
<td>Southern Ontario</td>
<td>Newfoundland</td>
<td>+ + +</td>
<td>left internal capsule, basal ganglia (lenticular nucleus), frontal corona radiata</td>
<td></td>
</tr>
<tr>
<td>Monrad-Krohn, 1947</td>
<td>30, f</td>
<td>Norwegian</td>
<td>German, French</td>
<td>/ / +</td>
<td>left frontal / fronto-parietal; old left cortical-subcortical infarction incl. anterior inferior parietal lobule, inferior motor cortex, and pars opercularis</td>
<td>TBI traumatic brain injury</td>
</tr>
<tr>
<td>Authors</td>
<td>Patient</td>
<td>Native language</td>
<td>Foreign accent</td>
<td>Changes of phonetic features</td>
<td>Lesion site</td>
<td>Aetiology</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
<td>-----------------</td>
<td>----------------</td>
<td>------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Moonis et al., 1993</td>
<td>59, m</td>
<td>American</td>
<td>French</td>
<td>+  +  –</td>
<td>left anterior dorsolateral inferior frontal gyrus and ipsilateral caudate nucleus; corticostriatal pathways</td>
<td>TBI</td>
</tr>
<tr>
<td>Carbary et al., 2000</td>
<td>51, m</td>
<td>English</td>
<td>“foreign sounding”</td>
<td>+  +  +</td>
<td>left posterior frontal lobe</td>
<td>TBI</td>
</tr>
<tr>
<td>González-Alvarez et al., 2003</td>
<td>51, f</td>
<td>Spanish</td>
<td>French, English, German</td>
<td>?  ?  ?</td>
<td>right caudate, putamen, internal capsule</td>
<td>TBI</td>
</tr>
<tr>
<td>Avila et al., 2004</td>
<td>51, f</td>
<td>Spanish</td>
<td>French</td>
<td>+  +  +</td>
<td>right temporal lobe and left inferior frontal corona radiate</td>
<td>TBI</td>
</tr>
<tr>
<td>Edwards et al., 2005</td>
<td>18, m</td>
<td>British</td>
<td>American</td>
<td>/  /  /</td>
<td>multiple small left hemisphere contusions</td>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td></td>
<td>53, f</td>
<td>English</td>
<td>Dutch</td>
<td>/  /  /</td>
<td>left motor cortex / subcortical contusions</td>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td></td>
<td>70, f</td>
<td>British</td>
<td>Welsh</td>
<td>/  /  /</td>
<td>left parietal, basal ganglia, internal capsule</td>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td>Lippert-Gruener et al., 2005</td>
<td>35, f</td>
<td>German</td>
<td>English</td>
<td>+  +  +</td>
<td>brain injury with traumatic left temporal haemorrhage</td>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td>Bakker et al., 2004</td>
<td>52, f</td>
<td>Canadian</td>
<td>Episodes of Dutch</td>
<td>+  +  +</td>
<td>deep white matter lesions in the corpus callosum, left parietal and left frontal lobe</td>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td>Abel et al., 2009</td>
<td>60, f</td>
<td>American</td>
<td>Swedish</td>
<td>/  /  /</td>
<td>left anterior parietal lobe</td>
<td>Carcinoma metastasis</td>
</tr>
<tr>
<td>Luzzi et al., 2008</td>
<td>64, f</td>
<td>Italien</td>
<td>Spanish</td>
<td>+  +  –</td>
<td>mild hypoperfusion of the left perisylvian speech area</td>
<td>Degenerative atrophy</td>
</tr>
<tr>
<td>Authors</td>
<td>Patient</td>
<td>Native language</td>
<td>Foreign accent</td>
<td>Changes of phonetic features</td>
<td>Lesion site</td>
<td>Aetiology</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Coelho et al., 2001</td>
<td>51, f</td>
<td>American</td>
<td>French / Italian</td>
<td>+</td>
<td>Unknown (MRT + CT normal)</td>
<td>cerebellar vermis; several small foci of hyperintensity in peripheral white matter of both frontal lobes, left inferior frontal corona radiate and left thalamus (were judged as non-specific); suspect of psychiatric (conversion) disorder.</td>
</tr>
<tr>
<td>Gurd et al., 2001</td>
<td>47, f</td>
<td>British</td>
<td>French</td>
<td>+</td>
<td></td>
<td>unknown / unclear</td>
</tr>
<tr>
<td>Poulin et al., 2007</td>
<td>74, m</td>
<td>Canadian</td>
<td>French</td>
<td>+ (but not specified)</td>
<td>bipolar disorder; diffuse hypometabolism in frontal, parietal and temporal lobes bilaterally; focal deficit in the anterior left temporal lobe which were related to asymmetric atrophy.</td>
<td>unknown / unclear</td>
</tr>
<tr>
<td>Katz et al, 2008</td>
<td>46, f</td>
<td>American</td>
<td>Swedish, Eastern European</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reeves &amp; Norton, 2001</td>
<td>65, m&lt;sup&gt;6&lt;/sup&gt;</td>
<td>American</td>
<td>British</td>
<td>+</td>
<td>paranoid delusions</td>
<td></td>
</tr>
<tr>
<td>Van Borsel et al., 2005</td>
<td>32, f</td>
<td>Dutch</td>
<td>Eastern European</td>
<td>+</td>
<td>psychogenic origin</td>
<td></td>
</tr>
<tr>
<td>Verhoeven et al, 2005</td>
<td>51, f</td>
<td>Dutch</td>
<td>French</td>
<td>+</td>
<td>conversion disorder</td>
<td>psychiatric / psychogenic</td>
</tr>
<tr>
<td>Reeves et al., 2007</td>
<td>30, m</td>
<td>American</td>
<td>Jamaican</td>
<td>+</td>
<td>schizophrenia</td>
<td>psychogenic</td>
</tr>
<tr>
<td></td>
<td>53, f</td>
<td>American</td>
<td>European</td>
<td>+</td>
<td>bipolar disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>66, m</td>
<td>American</td>
<td>British</td>
<td>+</td>
<td>schizophrenia</td>
<td></td>
</tr>
<tr>
<td>Tsuruga et al., 2008</td>
<td>44, f</td>
<td>Japanese</td>
<td>Chinese</td>
<td>-</td>
<td></td>
<td>dissociative (conversion) disorder</td>
</tr>
</tbody>
</table>

1–6 previously learned / exposed accent; ¹ worked in Polish-spoken areas for several months; ² during childhood extensive exposure to Irish English; ³ borned in Holland, moved to USA in the age of 5; ⁴ worked in Ireland for several years; ⁵ lived in Wales as a child; ⁶ caregiver (grandmother) in childhood spoke with a British accent; ⁺ observed symptom; – not existing symptom; / not specified; ? information is not available in secondary literature; a at the time of assessment (2 month p.o.), FAS recovered nearly completely; only residual prosodic deficits; b probably the same patient; c articles in Spanish / Japanese.
Musicality and phonetic language aptitude

Davide Nardo and Susanne Maria Reiterer

1. Introduction

1.1. A question of definitions

Several concepts are related to – and relevant for – the issue of musical talent. Musicality, musical ability, musical aptitude, musical intelligence and musical giftedness are just some examples. Unfortunately, it is very tricky to provide a single and simple definition of (musical) talent, even because such definition largely depends on both the theoretical and empirical context of a given author. However, most authors agree on two fundamental characteristics of talent: i) it is regarded as something special, or rather an exceptional capability in a given domain; ii) it is regarded as a potential, e.g. something capable of development (Jørgensen 2008).

Musicality is rather a loosely used term with many meanings (Jaffurs 2004). The term musicality refers to a sensitivity to, a knowledge of, or a talent for music. In psychology of music the term was first used by Révész (1953) to denote the ability to enjoy music aesthetically. However, Reimer (2003) prefers the term musical intelligence rather than musicality, in order to highlight its cognitive content instead of its similarities with talent, skill, or ability. According to him, there are many ways to be musically intelligent (i.e. in composing, performing, improvising, listening, etc.), and different individuals may show different levels of achievement in some of these abilities, but hardly in all of them. However, his definition of musicality goes beyond performance to include also aspects of music listening, which leads to an aesthetic experience.

As suggested by Jaffurs (2004), one way to define musicality in formal practice is to examine the standards for arts education, provided by the National Association for Music Education (MENC, 1994): i) singing, ii) performing on instruments, iii) improvising, iv) composing and arranging, v) reading and notating music, vi) understanding musical experience, vii) aesthetically evaluating, and viii) historical and cultural understanding. However, a survey with amateur musicians has shown that qualities like playing with expressiveness or feeling, timbre sensi-
tivity, repertoire readiness, imitating skills, and the ability to get along with other musicians were considered very important parts of an individual’s musicality (Green 2002). Thus, musicality is a multifaceted concept, more than just a skill, partially innate and yet strengthened by a nurturing environment.

On the other hand, talent can be defined as: i) a characteristic feature, aptitude, or disposition; ii) the natural endowments of a person; iii) a special, often creative or artistic aptitude; iv) general intelligence or mental power (ability). Aptitude can be characterized as: i) an inclination or tendency; ii) a natural ability (talent); iii) a capacity for learning; iv) a general suitability (aptness). A red thread draws a line throughout these concepts, conveying the meaning of something that is: i) somehow strongly connoted as innate (a gift, thus clearly separated from practice); ii) oriented towards something (a propensity or potential); iii) exceptional or extraordinary; and iv) closely related to a skill (ability, capacity). On these bases, we suggest a temporary definition of musical talent as a (predominantly) innate tendency to understand/appreciate, perform or create music outstandingly.

1.2. Major issues

When one is concerned with the measurement of musical talent, a series of fundamental questions arise: i) is talent an exclusively and exquisitely innate phenomenon, or does nurture play any role?; ii) is talent normally distributed in the population, or rather an “all-or-none” phenomenon?; iii) is talent a unitary trait, or rather a multi-dimensional one?; iv) if yes, how many sub-components make it up?; v) is talent similar to, or different from intelligence? In the present section, we will briefly introduce such issues, in order to make the reader aware of the complexity of musical talent and its measurement.

Although most authors agree that musicality is widely innate and hereditary, the “nature vs. nurture” controversy has been vivid also in this field, especially with respect to the amount of such innateness on the one hand, and the role contingent learning factors and the social environment play on the other hand, which should be neither neglected nor underestimated. Scientists devoted to talent measurement know how difficult it is to separate aptitude from achievement. This is another fundamental issue, particularly with respect to the music domain, where the ability (i.e. playing an instrument, solmizate, categorizing notes) can be so finely trained.
In the next sections, we will see how the different authors conceive musicality. However, we suggest that a helpful distinction could be drawn between *talent* (or *aptitude*) and *musicality* (or *ability*), where the former refers more to the innate component, and the latter to the resulting skill developed in interaction with the environment (through training, practice, etc.). This way, when we talk of *talent*, we are referring to an innate tendency to perform well in a given domain, the degree of which varies among different individuals, and which is rather independent of experience. Conversely, we could refer to an ability as the result of practice or learning, although it is highly probable that given the same amount of training, a talented person will outperform a non-talented one.

Is talent normally distributed in the population, or is it rather an “all-or-none” phenomenon? Although the popular belief mainly considers it in the second way (on the basis of which some individuals would be gifted, whilst others would not), the scientific literature reveals that music aptitude, like most of all human characteristics, is normally distributed in the population (Jørgensen 2008; Gordon 1989a). This implies that everybody has a certain degree of potential to achieve in music, with relatively fewer very high- and low-talented persons, and the majority with an average aptitude.

The concept of musicality is probably closer to the notion of *skill*, rather than to the traditional way psychologists see *intelligence*, unless we consider approaches like that of Howard Gardner (1983), who proposed the existence of several independent intelligences (see chapter by G. Rota). In section 3, we will review a series of experimental studies demonstrating a large independence between musicality and intelligence, when this latter is defined and measured as a general factor. Such evidence clearly demonstrates that musical aptitude cannot be considered an aspect (or a by-product) of intelligence, but rather an independent mental characteristic.

According to Gordon (1989a), there are two general points of view over music aptitude. The Gestaltists hold that music aptitude is a unitary trait of which overall intelligence is a substantial part. Conversely, the Atomists contend that music aptitude is multidimensional, consisting of various parts, none of which is significantly related to overall intelligence. The experimental evidence collected in almost one century of research on musicality testing converge at showing that there are different identifiable sub-components within musical talent, and that such sub-components are rather independent of intelligence, defined as a unitary trait.
1.3. What is measured

In section 1.1 we have provided some definitions of the various concepts related to musicality. Yet the question “what is measured?” (that is, what are the sub-components of musicality, or the fundamental abilities measured) can be risen. Nowadays there is good agreement on the importance of perception and cognition of musical patterns and structures. According to Shuter-Dyson (1999), there are five groups of fundamental abilities: tonal, rhythmic, kinesthetic, aesthetic and creative abilities, each of which could be subdivided into other sub-components, all of which are subject to improve with age and exposure (acculturation).

Tonal abilities comprise: i) pitch perception (i.e. the ability to discriminate different pitches); ii) sense of tonality (tonal reference, i.e. the development of a pitch system in which tone relations are specifically defined on the basis of their inferred relation to the tonic); and iii) harmony-polyphony (i.e. the ability to detect incorrectness or violations in chord sequences).

Rhythmic abilities are made up of different sub-components which correspond to different aspects of rhythm, like meter abstraction, perception of rhythmic structures, rhythmic anticipation (expectancy), practorhythmic factor (coordination of limbs in rhythmic movements), and tempo-tapping. Moreover, within rhythmic abilities a distinction between a figural and a metric perception has been suggested (Bamberger 1982), in which the former refers to the grouping of sounds into meaningful chunks, and the latter is focused on the steady pulse underlying the surface events of melody.

Kinesthetic abilities – the motor components – play a pivotal role in music performance like playing an instrument or singing. It has been observed that kinesthetic factors influence the ability to improvise (McPherson 1993/1994), and the bodily movements made by the performers contribute to expressivity of the performer (Davidson 1993). However, kinesthetic abilities also play a role in auditory perception. For instance, Mainwaring (1933) reported that kinesthetic cues were used by his subjects in order to recall tunes. On the other hand, Baily (1985) highlighted the need to study the way musical patterns may be represented cognitively as patterns of movements rather than as patterns of sound.

1. Rhythmic abilities are peculiar because on the one hand it has been shown that rhythm is processed relatively independently from pitch (Peretz & Coltheart 2003), whereas on the other hand there is evidence of reciprocal interactions between rhythm, melody and tonality (Gordon 1965; Sloboda 1985).
Aesthetic abilities are fundamentally related to expression, appreciation and emotion. An interesting research by Clarke (1993) with pianists required to imitate performances showed that the more the relationship between structure and expression was disrupted, the more inaccurate and unstable was the attempt at imitating. Gabrielsson (1982) has shown that a balanced combination of the structural, motional and emotional aspects adapted to the needs of a given individual and the actual musical content may be what is required for artistic performance. Swanick (1973) claimed that much cognitive activity is involved in aesthetic response to music, and that the intensity and quality of any emotional experience depends on this activity. Moreover, the ability to make predictions (expectations) as to what may follow is central to the process of understanding music, so that in general deviations arouse excitement.

Creative abilities – like aesthetic abilities – are rather a fuzzy concept. What is then creativity? Supposedly, a cognitive process resulting in the production of something which is both original and highly valuable (Sternberg 1996). Webster (1988) has claimed that a “collection of musical aptitudes” is necessary for a creative work in music: i) convergent skills (i.e. the above-mentioned abilities to recognize rhythmic and tonal patterns, musical syntax, etc.); ii) so-called divergent skills (i.e. musical extensiveness, flexibility, and originality); and iii) other abilities, like conceptual understanding, craftsmanship and aesthetic sensitivity. Musicality tests which attempt to measure musical creativity (Vaughan 1977; Webster 1983; Wang 1985) mainly exploit improvisation. Although musical creativity factors seem to be unrelated to other sub-components of musical aptitude (Swanner 1985), they show a certain degree of association with personality traits of imagination, curiosity and anxiety. Some studies (Kratus 1989, 1991) have shown a negative correlation between musical aptitude and the need to explore the musical pieces (a sub-component of the composition activity).

Correlations between general intelligence scores and musical ability tests are mostly found to be positive, but low (generally about 0.30; see Shuter-Dyson & Gabriel 1981).

In his work on cognitive abilities, Carroll (1993) reanalyzed hundreds of test data-sets and proposed a model which specifies what kinds of individual differences in cognitive abilities exist, and how they are related to one another. According to this model, there is a large number of distinct individual differences in cognitive ability, and the relationships among them can be derived by classifying them into three different strata: i) stratum I, specific abilities (among which the specific factors in
perceiving music and musical sounds); ii) stratum II, **broad abilities** (i.e. fluid and crystallized intelligence, general memory and learning, retrieval ability, cognitive speediness, etc.); iii) stratum III, **general intellectual ability** similar to Spearman’s “g” factor (1927).

By analyzing the relevant literature, Carroll identifies 31 factors of musical talent, which he further divides into four subgroups: i) **general sound discrimination factors**, comprising basic abilities to discriminate tones or patterns of tones with respect to their fundamental attributes of pitch, timbre, intensity, duration, and rhythm; ii) **sound-frequency discrimination factors**, similar to those of the previous group, but focused on discriminations of the frequency attributes of tones (i.e. detecting a changed note in a melody, detecting the number of notes in a chord, detecting a changed note in a chord, etc.); iii) **sound intensity and duration discrimination factors**, focusing on discriminations with respect to intensity (loudness), duration and rhythm (amplitude and temporal attributes of sounds and sequential patterns of sound), which depend on sensitivity to temporal and rhythmic aspects of tonal passages; iv) **musical sensitivity and judgment factors**, comprising judgments of the “musicality” of short musical passages (which sounds best?) based on phrasing, loudness, rhythm and harmony, in general independent of factors assessing simple auditory discriminations. This factor can be further divided into a **tonal imagery subfactor**, emphasizing melodic and harmonic aspects of music, and a **musical expression subfactor**, stressing those aspects arising from variations in phrasing, loudness and tempo.

Nearly all measures of musical aptitude depend to a great extent on tests of very elementary discriminations among tonal materials, with only little musical contexts. According to Carroll (1993), this could be due to the desire to minimize the effects of musical training (so that a test can be used to predict success in such training), but also to a failure to recognize the possibilities of preparing a test including an appropriate musical context. Expert musicians and music educators, so Carroll, tend to discredit simple tests of auditory discriminations (like Seashore’s) as possible tests of musical aptitude because they contain little or no musical meaning. However, it is difficult to develop tests with a desirable level of musical meaning that do not at the same time become tests influenced by musical training and experience.
2. **How to measure musicality**

Each author tends to have his own view of what exactly musical aptitude is, and what subcomponents it is made of. As a consequence, each musical aptitude test has been created from a different perspective, often criticizing the works of predecessors and attempting to combine, complete or improve them. This fact has also generated a series of theoretical proposals and empirical approaches that do not always coincide.

Several psychometric tests of musical talent and ability have been created in the last century, some approximate to tests used by musicians, others analyze music into its most elementary basic constituents. We cannot survey them all here (for a review, see Shuter-Dyson & Gabriel 1981), but we will examine the most popular ones, in order to highlight which issues emerge when attempting to measure musical aptitude, and what such issues tell us about the nature and the characteristics of musical talent. We will consider three major tests: Seashore’s test, especially for historical reasons; Wing’s test, for its cognitive implications; and Gordon’s tests, for its flexibility and popularity.

2.1. Seashore’s measures of musical talents

Seashore’s *Measures of Musical Talents* is the oldest standardized music test available, first published in 1919, and subsequently revised in 1939. The characteristic of this test is to focus on the very basic sensory capacities with a strong psychophysical approach, by presenting the subjects a series of pairs of tones and requiring to discriminate a certain physical characteristic between them.

The revised version of the test (1939) is made up of six subtests, each assessing a specific domain of musical aptitude: i) *pitch*; ii) *loudness*; iii) *rhythm*; iv) *time*; v) *timbre*; and vi) *tonal memory*. In the first five subtests, the subject is required to compare two items (notes or rhythmic patterns) and say whether they differ (and sometimes in which direction, i.e. higher/lower, stronger/weaker, longer/shorter). In the last subtest, the subject has to listen to a series of pairs of consecutive tones forming no melody, and to identify within each pair which tone in the second sequence differs in pitch from its corresponding tone in the first one. In each subtest only one factor is varied at a time, while others are kept constant and as simple as possible. This way, the test should be equally feasible for young and old, musicians and non-musicians, because it measures immediate sensory acts that do not improve with practice.
According to Seashore (1919), musical talent is a gift of nature, as it can be inherited but not acquired, and the measurement of musical ability chiefly regards inborn psychophysical and mental capacities as distinguished from skills acquired by training. In his view, musical talent is not unitary, rather there would be a hierarchy of related talents which work together, and such hierarchy would present different organizations in different individuals. Therefore, the main aim of the assessment of musical talent is to characterize the dominant traits, as well as determining both qualitatively and quantitatively the composition of each hierarchy of traits. Hence, the test permits a quantitative measure of the magnitude of each trait, and a description of the distribution of individual differences for each one as well. A collection of a large number of cases allows in turn the creation of a curve of distribution which can be referred to as a norm for the interpretation of individual records (expressed in percentile ranks).

Seashore proposes a classification of the essential traits of musical talent considering on the one hand the characteristics of sound which constitute music, and on the other hand the mental skills which are needed for the appreciation of musical sounds. As regards the characteristics of sound, he identifies three elements relevant for testing: pitch, time and intensity. Pitch is defined as the quality or the essence of sound, a basic element underlying more complex music phenomena such as timbre, consonance and harmony. On the contrary, rhythm is a combination of more fundamental elements (i.e. time and intensity). Thus, according to the author, this classification permits the arrangement of musical talents into the ability to appreciate and the ability to express respectively pitch, time and intensity.

As regards the mental skills, Seashore divides the capacity for the appreciation and expression of music into four fundamental abilities: i) sensory (i.e. hearing music); ii) motor (i.e. expressing music); iii) associational (i.e. understanding music); and iv) affective (i.e. feeling music and expressing feelings in music). He claims that, by combining these two classifications – the elements of musical sounds and the capacities of human individuals – the principal groups of musical talent would be obtained.

Seashore makes a clear distinction between what he calls cognitive and physiological thresholds. The cognitive threshold is a limit due to cognitive difficulties such as ignorance, misunderstanding, inattention, lack of application, confusion, disturbances, misleading thought, etc. Conversely, the physiological threshold is a limit determined by the character
of the physical structure of the inner ear. The author states that the cognitive threshold is no measure at all, but rather an indication of a lack of control over other conditions. Instead, a correct measurement should give the physiological threshold (or at least a proximate physiological threshold), given that the former is scarcely attainable. Anyway, according to him we cannot get a measure below the physiological threshold, and any error is due to the cognitive threshold. He also warns against what he calls the illusions, that is the influence of conscious or unconscious anticipation (expectation) on the judgmental process (i.e. illusion of pitch, intensity, timbre, etc.).

According to Seashore, absolute pitch is just an illusion cultivated by many musicians. In fact, one could identify a note sounded in isolation not by absolute pitch, but by memory of conditions of tuning, by difference in timbre, or by guess. However, the author believes that the sensitivity of the ear to pitch difference (as well as to other elements of sound) cannot be improved appreciably by practice. Practice can only improve the cognitive threshold by clearing up difficulties (such as information, observations, development of interest, isolation of the problems, application to the task, etc.) which could hinder an actual measure of discrimination. Thus, training in discrimination is not like the acquisition of a skill, because only the meaning of pitch can be refined through training.

Finally, Seashore asserts that the actual psychophysical capacity for pitch discrimination (and other elements of sound) does not improve with age and does not vary with sex. In fact, the records of younger children are just slightly inferior to those of the older, and this could be accounted for by the presence of conditions for observation which are overcome as experience grows with age. Furthermore, although scores of girls are normally superior to those of boys, this could be better explained by the relative lack of interest in music typically shown by pre-adolescent boys.

The historical importance and the scientific influence of Seashore’s test should not be underestimated, because it was a pioneering work and the first to be fully standardized. Nevertheless, its atomistic approach has been heavily criticized, and its weakness in predicting musical ability has been demonstrated, being even scarcely more efficient than general intelligence in predicting musical achievement (for a review see Wing 1970).

2. Absolute pitch (or perfect pitch) is the ability found in a minority of listeners to name an isolated musical tone presented without an objective reference tone or, conversely, to produce a tone identified by name only.
Let us briefly follow the major critiques Wing raises against Seashore’s approach.

Of course, by choosing pitch, intensity, rhythm, time, timbre and memory, Seashore has selected the most commonly accepted basic qualities of musical capacity. However, the elementary way he had tested them, has moved him from “music” (made up of patterns and relationships of tones) to mere sensory perception. To the musician, the pitch subtest is too simple, and measures too fine a degree of discrimination. The time and intensity subtests are probably the least satisfactory of the battery, because of their “distance” from actual music. In fact, they do not test for time and intensity as they are used in music. In an actual music performance, the correct length of a note is not based on a comparison with the note just played, but on the dynamic rhythmic progression of the melody. In the same way, intensity does not merely consist in noticing that one note is louder than another, but in getting an intensity change suited to the melodic line and the whole character of the piece. Memory and rhythm are probably the best of the Seashore’s subtests, because of their “closeness” to actual music. Nonetheless, it is questionable whether a test for memory on nonsense material is fully valid for musical (and therefore “meaningful”) material. Moreover, there is the possibility that a subject could score low in those subtests which do not gain his “attention”, because they are far away from actual music.

2.2. Wing’s standardised tests of musical intelligence

The Wing’s test has been devised and revised by the author between the late 1930s and 1970. In his work, the author intended to reconcile the pragmatic experience of musicians with the experimental experience of psychologists by creating a test of music aptitude independent from musical training, and capable of pointing out: i) the mental processes implicated in music fruition; ii) the distribution of musical aptitude in the population; iii) its development with age and learning; and iv) the influence of the environment and culture. In conceiving such measurement, the opposition between a nativist and an empiricist approach has been mediated by cognitivism, according to which music results from mental processes of organization and transformation of the physical stimuli. In order to find these processes, the author has devised a series of standardized subtests which employ structured musical material instead of elements of music (as in Seashore’s test).
By surveying the psychological literature of his time, Wing identifies a series of weaknesses in the previous tests of music aptitude (including Seashor’s): i) ignoring qualities a musician regards as desirable (i.e. appreciation), or emphasizing qualities a musician regards as of little importance (e.g. absolute pitch); ii) measuring only one aspect of music and treating it as a measure of a general musical capacity; iii) validating tests on very small groups; iv) lack of an adequate standardization of the test scores; v) application to a narrow age range, or difficulty in re-testing procedures; vi) lack of any attempt to correlate with teachers’ ranking; vii) neglecting the effect of musical training on the test scores. The author intended to conceive a test which would have compensated for all the above-mentioned weaknesses.

According to Wing, the ultimate version of its test satisfy the criteria a scientific psychological test of music aptitude should: i) being acceptable to musicians; ii) not being influenced by training; iii) allowing the assessment of a wide range of different capabilities; iv) providing information on several relevant aspects of musical talent; v) being statistically reliable; vi) providing a standardized score; vii) requiring short times for the administration; viii) correlating well with scores provided by music teachers; ix) being of practical use in musical education; x) being easy to administer even with younger children. Wing has standardized the scores for the English population of different age cohorts, calculating on this basis a musical age and a musical quotient. His test is split into two sub-tests, the first one measuring perceptive aspects (ability subtests), and the second one measuring more cognitive components (appreciation sub-tests).

Two terms which are central in Wing’s view of musical talent are musical ability and musical appreciation. Although strictly speaking the first refers to the ability to play an instrument, in a wider sense it includes the speed in learning to play, the ability to perform an aural test, and the ability to carry out musical activities such as composing. On the other hand, musical appreciation (which is distinguished from musical ability both by musicians and psychologists), is the power to recognize and evaluate artistic merit in music, and involves the deliberate aesthetic judgments of music as it actually exists in compositions, rather than ability to solve problems connected with the elementary materials of which music is composed. However, the author claims that music ability and appreciation should be connected in some way. Wing regards his measurement of appreciation as a revolutionary innovation, since previous tests did not deal with the essential aesthetic element involved in music, but were al-
most exclusively concerned with the simpler perceptual processes or with the knowledge of musical technicalities.

For Wing, nature is far more important than nurture, and he offers the following support for his stand: i) 11 year olds who score in the lowest quartile on his test take music lessons as often as those in the highest quartile; ii) test scores may continue to climb for some time after music lessons are over; iii) Wing scores of children, tested again after 5 years, correlate about 0.9 whether or not the subjects have had music lessons in the meantime; iv) high testing children do as well on unfamiliar music test items as on the more familiar; v) having two musical parents is associated with higher test scores than having only one such parent.

In comparison with previous tests, Wing's test is characterized by the employment of original material, higher reliability and validity, and an easier and more uniform administration. The subtests are designed in such a way that they require no special knowledge of musical technicalities, and within each, difficulty is graded so that the easiest items are suitable for children, while the hardest ones may be used to test the capacities of professional musicians. The items are made up of quite short melodic extracts, usually consisting of eight bars only, with each subtest containing 20 items as a rule. The whole set takes about one hour to be administered.

In the last version (Wing 1970), there are seven subtests which involve the following tasks: i) chord analysis, i.e. detecting the number of notes played in a single chord (sometimes made up of just one note); ii) pitch change, i.e. detecting an alteration of a single note in a repeated chord; iii) memory, i.e. detecting an alteration of a note in a short melody; iv) rhythmic accent, i.e. choosing the better rhythmic accent in two performances; v) harmony, i.e. judging the more appropriate of two harmonizations; vi) intensity, i.e. judging the more appropriate mode of varying loudness (crescendo, decrescendo, etc.) in two performances of the same melody; vii) phrasing, i.e. judging the more appropriate phrasing (grouping of notes by pauses, legato and staccato playing, etc.) in two performances.

Wing claims that it is important to assess not only the individual's general capacity for musical appreciation, but also the particular type of musical appreciation in which one is weak or strong.

Wing (1941) performed a factor analysis on a large dataset collected with his test, in order to characterize the dimensions underlying the various subtests. Results have identified three factors. The first factor could be treated as measuring the same group of mental processes that Wing dubs general musical ability. The second factor was found to divide the seven subtests into two classes, the first including those in which the
essential task of the listener is to judge the more appropriate musical arrangement (e.g. appreciation), the second comprising those in which the task is merely to perceive a change (e.g. ear acuity). According to Wing, this factor resembles the finding of similar bipolar factors in other tests of artistic and intellectual abilities, for example the opposition between the synthetic activity described as “intuition” (in which we implicitly comprehend the essential meaning or character of a whole), and the analytic activity which essentially consists in explicitly analyzing the whole into its component parts and the relations between them.

The third factor Wing extracted showed saturations with the two sub-tests of harmony and chord analysis. These are the only ones which essentially depend on listening to notes sounded simultaneously, whereas the others deal mainly with the melodic or rhythmic contour of the music played. Thus, according to Wing, this factor distinguishes those persons who have a better appreciation for harmony than for melody or rhythm.

Finally, Wing reports that rhythm seems to have a comparatively weak association with the general musical ability, this way anticipating what more recent research has firmly demonstrated (Peretz & Coltheart 2003). He claims that of all musical capacities, the ability to recognize rhythm is probably the most elementary, in fact, it develops early, is the most widely diffused, and may exist in almost complete independence of any deeper appreciation of higher developments of musical art.

2.3. Gordon’s measures of music audiation

As from the 1960s, Gordon has devised a series of musicality tests for various purposes, which have become very popular in the literature for several reasons: i) they can be used with subjects belonging to various age cohorts; ii) they can be used with subjects of different levels of expertise; iii) they measure different aspects of musicality, such as pitch ability, rhythmic ability, performance and expression preferences, ability to improvise, ability to score reading, etc; iv) they have been re-mastered following the digital era and recorded on CD.

Gordon stresses the distinction between aptitude and achievement. He defines music aptitude as “the potential to learn or achieve in music” (an inner possibility), whereas music achievement represents “what somebody has already learned in music” (an outer reality). According to the author, those students who show a high level of music achievement, must also have a high level of music aptitude, whereas vice-versa is not necessarily true, as revealed by their scoring on a music aptitude test. In fact, stu-
Students with low music aptitude who receive proper instruction may achieve more success than students with average music aptitude who receive improper instruction.

Another important distinction, is that between a developmental stage and a stabilized stage within music aptitude. According to Gordon, music aptitude is basically innate, but not inherited. Thus, heredity influences music aptitude, but it does not entirely determine it. In fact, although innate, it also depends on a rich music environment to come to fruition. Hence, music aptitude becomes a product of an innate potential plus some early environmental musical influences, remaining lacking in the case of an inappropriate environment. In Gordon’s view, music aptitude is therefore a product of both nature and nurture.

Gordon calls developmental music aptitude stage the period from birth to approximately age nine, a period in which according to him the environment would have a pronounced effect on music aptitude. During those years, a child’s music aptitude level would constantly fluctuate, and the potential may go up or down, according to the modulation of the environment. However, the effect of the environment would start to decrease shortly after birth and keep on diminishing with age, until about age nine music aptitude would stabilize and remain in what he calls stabilized music aptitude stage throughout adulthood.

The author claims that it is very important that children receive the highest quality of both informal music guidance and formal music instruction during the developmental music aptitude stage, because this would increase their immediate level of achievement, their overall level of music aptitude, and their life-time potential for music achievement. The younger the children, the better they may benefit from a high-quality music environment. On the contrary, inappropriate or lacking instructions or no exposure to music whatsoever would drastically reduce a child’s developmental music aptitude. However, although early environmental influences promote music aptitude, in Gordon’s view one’s music aptitude cannot reach a higher level than that with which one is born, and no one would be able to reach a level in music achievement higher than that at which his aptitude has stabilized.

Gordon considers unlikely the existence of completely separate aptitudes for composition, improvisation, instrument and vocal performance, rather, he suggests there would be different personality traits and psychomotor abilities, as well as separate sub-components of music aptitude, including preference and non-preference. According to him, aptitude would be unique but not unitary, and best represented by the inter-
Musicality and phonetic language aptitude

action of several human attributes. Nonetheless, it would have very little or even no relation to any other human trait, comprising race, religion, nationality or sex, and it would be also unrelated to the instrument one plays. It would be multidimensional, including a tonal aptitude (related to melody and harmony), a rhythm aptitude (related to tempo and meter), and expressive (related to phrasing, balance, style), improvisatory and creative aptitudes. Moreover, scores on both developmental and stabilized music aptitude tests would be normally distributed.

To better describe musical aptitude, Gordon coins the term *audiation*, defining it as the capacity to assimilate and comprehend in our mind music for which the sound is not physically present (delayed musical events). On the contrary, *aural perception* occurs when we hear sounds in the very moment they are being produced (immediate sound events). Obviously, we are able to audiate actual sounds only after we have aurally perceived them. Gordon claims that, although audiation would be fundamental to both aptitude and achievement, it would work differently in each. In fact, while audiation potential cannot be taught, “how to audiate” could, i.e. how to use one’s intrinsic audiation (aptitude) to maximize one’s own acquired music achievement (as influenced by the environment). From Gordon’s perspective, sound becomes music only through audiation, when we translate it in our mind and give it a meaning, although such meaning would differ on different occasions, as well as from one person to the other. According to him, we audiate when listening to, recalling, performing, interpreting, creating, improvising, reading, or writing music.

Gordon makes a comparison between music and speech, claiming that in the same way as speech communicates the meanings we have in mind, music performance communicates audiation (that is, the meaning of music). In fact, while listening to speech, we give meaning to what is said by connecting it with what we have heard on other occasions, and we create expectancies of what we will hear next, on the bases of our experience and understanding. Similarly, while listening to music, we give meaning to what we hear by connecting it with what we have heard on other occasions, and we create expectancies of what we will hear next, on the bases of our music achievement. Audiating is therefore the process of summarizing and generalizing from the specific music patterns we hear as a way to anticipate or predict what will follow. However, audiation is different from both imitation and memorization. We are able to store specific material in our memory without understanding it, but then we quickly forget it, and music makes no exception. Audiation leads to understanding, whereas imitation and memorization – when separated from audiation –
lead at best to emotional reaction. In the same way, without audiation a performer can neither improvise nor create.

Gordon identifies seven types of audiation, which serve as readiness for others, but are not hierarchically organized, and describes six stages of audiation, which conversely are hierarchical and cumulative, each of which establishing the basis for – and combining with – the next one.

Gordon has devised several tests, which can be divided into two groups with different features in accordance with the music aptitude stage they are designed for. The developmental music aptitude tests employ either tonal or rhythmic patterns, but not melodic patterns which combine both aspects. They use “same/different” or “same/not same” responses, and each question is identified by a simple picture of a familiar object. Developmental tests are audie (1989b), for children aged three to four, PMMA (Primary Measures of Music Audiation, 1979), for children in grades from kindergarten through three, and IMMA (Intermediate Measures of Music Audiation, 1982), an advanced version of PMMA designed for children aged six to eleven with a higher music aptitude. The patterns are always performed on electronic instruments, because according to the author, students in the developmental music aptitude stage are more interested in how music is constructed, rather than in its expressive components. Vice versa, the stabilized music aptitude tests (i.e. MAP and AMMA), employ music excerpts composed on purpose, performed with actual music instruments. They employ “same/different”, “like/different”, and “yes/no” responses, because options like high/low, up/down, short/long risk to transform a music aptitude into a music achievement test. Moreover, they also contain so-called “preference measures”, and each question is identified by a progressive number.

AMMA (Advanced Measures of Music Audiation, 1989a) is a test designed for high school students and college/university music and non-music majors. It is made up of 30 items, each of which consists of a short musical statement followed after four seconds by a short musical answer with the same number of notes. The subject is asked to decide whether they are the same or different. When the answer is different from the statement, the subject is asked to decide whether the difference is a tonal or rhythm change. Within a certain item there may be either one or more tonal changes, or one or more rhythmic changes, but not both, and the difference between the statement and the answer may occur at the beginning, in the middle, and/or at the end of the item. All items are programmed on a computer and performed on an electronic instrument. AMMA is also the test we chose in our current experiments.
MAP (Musical Aptitude Profile, 1965) is an eclectic battery designed to measure seven separate dimensions for students in grades four through twelve. It consists of three main sections: tonal imagery (comprising a melody and a harmony non-preference subtests); rhythm imagery (including a tempo and a meter non-preference subtests); and musical sensitivity (including three preference subtests: phrasing, balance, and style). The non-preference subtests work in a similar way to those of AMMA, conversely, in the preference subtests, the subject listens to two versions of a melody, and he is asked to decide which of the two “sounds better”. In phrasing the two versions are performed with different musical expression; in balance they begin in the same way, but end in a different way; and in style the two versions are performed at different tempos.

Gordon has also devised also other tests, to determine whether a student has the necessary readiness and ability to improvise (HIRR and RIRR), to help students to choose an appropriate musical instrument (ITPT), and to measure tonal, rhythm and notational audiation (ITML).

3. Musicality meets language talent

3.1. A common ground for music and language

Many authors claim that, beyond their respective differences, language and music share some common characteristics. Both of them are auditory phenomena that follow a time line (temporal aspect). Moreover, rhythm and melody in music can be compared to stress and intonation in language (Arleo 2000). Both of them are human universals consisting of perceptually discrete elements organized into hierarchically structured sequences, be it from the individual note to the larger constituent of a musical composition, or from phonemes to the discourse units (Sloboda 1985; Patel 2003). Both of them share a series of fundamental characteristics, such as the processing of sounds, the conveyance of messages, the learning by exposure, the sharing of intrinsic features like pitch, volume, prominence, stress, tone, rhythm, and pauses (Fonseca Mora 2000). It has also been suggested that, in the same way in which rhythmic structures in the prosody influence the meaning of segments in English, rhythmic structure or patterns of accent strength affect the relative importance with which musical events are interpreted (Palmer & Kelly 1992). Infant-directed speech is music-like in a number of aspects (e.g., regular rhythms, slow tempo, pitch contours expanded and repeated with altered
lexical or segmental content and varying tempo, extended vowels) and has a distinct suprasegmental structure (Trehub & Trainor 1993). Moreover, musical abilities probably play an important role in the acquisition and the processing of language. In fact, infants acquire much information about word and phrase boundaries (and possibly even about word meaning), through different types of prosodic (thus musical) cues of language, such as speech melody, metre, rhythm and timbre (Jusczyk 1999). Finally, tonal languages rely on the decoding of pitch relations between phonemes, and non-tonal languages also require an accurate analysis of speech prosody to decode structure and meaning of speech (Koelsch & Siebel 2005).

3.2. Music training and language acquisition

We have already seen that musical aptitude is different from musical ability, in that the latter is affected by training. Before treating the relationships between musical aptitude and language acquisition, it is worth to take a look to those studies considering the relationship between music training and language acquisition.

The existence of a positive influence of music training on language acquisition in children and adolescents has been consistently reported. One study showed that students who receive a musical training are more successful in discriminating and performing French pronunciation than those who do not (Harrison 1979). Another reported that Asian students of English distinguished between minimal pairs more effectively when the sounds were presented contextually in songs and chants rather than when they were presented in word lists (Karimer 1984). It was also found that listening to music in a second language class improved auditory discrimination relevant to learning proficiency (Pinel 1990; Tomatis 1991). Moreover, it was demonstrated that a group receiving music lessons performed significantly better in both oral grammar and reading comprehension of French (Lowe 1998). On the other hand, Deutsch (1991) has demonstrated that the perception of pitches is influenced by the mother language spoken by the listener.

In a study by Spychiger (1993), primary school children received extra music lessons in place of other school subjects over the course of three years, and results showed that these children performed better than their peers in language and reading skills. Douglas and Willatts (1994) carried out a study in which children with reading difficulties were given a music training, whilst a control group undertook exercises in non-musical ac-
tivities. Significantly, reading scores for the music group increased, whereas scores for the control group did not. Similarly, Costa-Giomi (1999) demonstrated that two years of piano instruction significantly improved verbal abilities of ten to eleven year olds compared with controls. Finally, it was found that a musical training at a young age caused a significant improvement in short-term verbal memory in adulthood (Chan et al. 1998). On the other hand, Stokes (2001) found no correlation between music training and L2 acquisition in adult learners. Musical ability can predict aspects of first-language (L1) verbal ability, such as reading ability in children (Atterbury 1985; Anvari et al. 2002). Jakobson et al. (2003) reported enhanced verbal memory performance in musicians.

Milovanov et al. (2007) have studied the phonemic processing skills of musicians and non-musicians with the dichotic listening task in children and adults with varying degrees of musical aptitude (as assessed by Seashore’s test). Subjects were given phonetically meaningful – but semantically irrelevant – consonant/vowel syllables pairs presented to both ears, always two different pairs at a time. Results showed superior left ear monitoring skills among the adults who practiced music regularly, indicating altered hemispheric functioning, whereas other musically talented subjects did not have the ability to control left ear functioning in an equal manner, that is, the performance of musical children and their non-musical controls in the left ear condition did not differ. Thus, regular music practice may have a modulating effect on the brain’s linguistic organization.

An improving effect of musical practice on pitch processing in speech has been recently also demonstrated with the ERPs technique (Schön et al. 2004; Besson et al. 2007), suggesting that a set of common processes may be responsible for pitch processing in both music and in speech.

A very recent study (Pastuszek-Lipinska 2008) has investigated whether music training and education influence the perception and pro-

3. The Dichotic Listening Task is a technique devised to investigate the functional hemispheric specialization. It exploits the physiology of the auditory ascending paths, so that two thirds of the fibers in the auditory nerves go to the contralateral hemisphere, whereas one third of the fibers remains ipsilateral. When two different stimuli (e.g. two different words, or two different pitches) are presented simultaneously at the two ears, the subject typically reports the stimulus for which the hemisphere is specialized, i.e. the left hemisphere (right ear) for verbal material, and the right hemisphere (left ear) for musical material.
duction of L2 sounds. Subjects were Polish native speakers either musicians or non-musicians, who were asked to reproduce as accurately as they could sentences in the following foreign languages: English, Belgian Dutch, French, Italian, Spanish and Japanese. Results revealed that musicians outperformed non-musicians, in that the former produced more sentences, encountered fewer difficulties with the task, and were rated as more fluent with respect to the latter, demonstrating that music education exerted a measurable impact on speech perception and production. The author concluded that since the influence of musical expertise extends to speech processing, music education should be considered an enabling factor in the successful acquisition of L2.

On which basis would a musical training be able to improve language processing? Some explanations have been given, mainly based on the role musical (specifically rhythmic) processing would play in the development of short-term verbal memory (Karimer 1984; Chan et al. 1998) and temporal processing ability, i.e. the fine discriminations between rapidly changing acoustic events (Jakobson et al. 2003), and on the basis of a common neural substrate between musical rhythm processing and language reading (Douglas & Willatts 1994).

Furthermore, music has been employed as an alternative treatment for language impairments. For example, Benson et al. (1994) have used a music-based therapy (Melodic Intonation Therapy) in aphasic patients with severe left-hemisphere brain damages. In this therapy, word sequences are incorporated into a song, and after some time the melody is removed until the patient can speak without singing. Such therapy exploits the intonation and singing abilities preserved in the right hemisphere, which memorizes the phrases through music. Music therapy has also been employed to help children with speech and language impairment (SLI), who may have sufficient speech sounds and vocabulary, but may stop expressing themselves fully through speech. Sutton (1995) observed an interesting parallel between SLI children's progress in music and their progress in language: as they began to build music into phrases and structures, they also began to express themselves with their voices and construct simple sentences.

3.3. Music aptitude and language acquisition

There is evidence of a relationship between music skills and the acquisition of the mother tongue (L1). By investigating the relationship between reading and auditory abilities in English native speakers, Ewers (1950)
found significant correlations between reading scores and musical skills such as pitch discrimination, loudness, musical rhythm and tonal memory. In another study (Wheeler & Wheeler 1954), the Seashore pitch subtest was found to correlate with an auditory discrimination test for English sounds, and with a test of reading skills, but not with general intelligence. Holmes (1954) also reported that various auditory abilities play an important role in L1 spelling ability both in high school students (i.e. tonal movement, pitch, tonal memory, intensity, rhythm and melodic taste) and college students (i.e. tonal memory and pitch), irrespective of intelligence. More recently, Douglas and Willatts (1994) found an association between rhythmic ability and reading, rhythmic ability and spelling, but not between pitch discrimination ability and reading.

On the other hand, there is vast evidence of a significant relationship between music skills and second language (L2) acquisition. Dexter and Omwake (1934) investigated the relationship between the ability to discriminate pitches (Seashore’s pitch subtest) and pronunciation ability in French, and found that pitch correlated significantly with accent ratings. Another study (Eterno 1961) reported that both musical aptitude and musical training were capable of predicting foreign language pronunciation success. A study by Pimsleur et al. (1962) found that pitch and timbre discrimination were consistently related to the auditory comprehension of French. Interestingly, Leutenegger et al. (1965) investigated both the effects of musical aptitude on language learning ability (in French and Spanish), and the effects of language learning on musical aptitude, also controlling for sex and intelligence. Although results on the whole did not show strong relationships between the Seashore subtests and foreign language achievement, in the female group the tonal memory scores significantly predicted the achievement scores in French.

By using the Seashore test and some pronunciation tests, Arellano & Draper (1972) found a strong correlation between timbre and intonation, timbre and phones, rhythm and intonation, rhythm and phones, and tonal memory and phones, demonstrating the existence of a relationship between perceptual musical skills and productive phonetic aspects. Moreover, Fish (1984) found a strong correlation between pitch discrimination and sound discrimination, as well as between sound discrimination and the playing of a musical instrument. However, no correlations between pitch discrimination ability and pronunciation ability of German (L2) phonemes, pronunciation ability of Germans phonemes and musical background, and sound discrimination ability and pronunciation ability of German phonemes were found. Therefore, music was re-
lated to language perception, but had little influence on language production.

Similarly, by using the Seashore test and some pronunciation production measures Brutten et al. (1985) found no significant correlations. By examining the association between musical and language aptitude, Stevenson (1999) found a correlation between rhythmic ability and the ability to reproduce words in a foreign language, as well as between the ability to sing back melodies and the ability to reproduce foreign language words. Tucker (2000) examined foreign language aptitude in English and Japanese native speakers by using the MLAT (and its Japanese translation) and self-assessments of foreign language proficiency, and Seashore’s test. Results showed significant correlations between the MLAT and self-assessments and the tonal memory subtest, as well as between the MLAT and the rhythm subtest. Moreover, Anvari et al. (2002) found significant correlations between musical aptitude and both phonological awareness and reading development.

Morgan (2003) has investigated the relationship between both receptive and productive aspects of music (Gordon’s IMMA and vocal notes reproduction), and those of French as L2 (vowel discrimination and accent production). Her results showed correlations between perception of rhythm and speech perception, perception of rhythm and accent production, music production and speech perception, and between music production and accent production, thus demonstrating crossed influences within the same experimental design. Gilleece (2006) has investigated the relationship between musical and foreign language aptitude in English native speakers, also controlling for the role played by general intelligence. She also assessed both receptive and productive aspects, the former by means of Bentley’s test for music, and a test based on MLAT plus a discrimination task of Chinese and Czech words for language; the latter by means of imitation of short rhythm patterns for music, and imitation of Korean words and Spanish sentences for language. Results showed a significant correlation between receptive musical and language aptitudes, as well as a significant correlation between productive skills in language and music, both irrespective of general intelligence.

Slevc and Myiake (2006) have investigated the relationship between musical aptitude (as assessed by Wing’s test) and L2 proficiency (i.e., receptive and productive phonology, syntax and lexical knowledge) in adult Japanese native speakers who were learning English as L2, while controlling for age of L2 immersion, patterns of language use and exposure, and phonological short-term memory. Their results showed that
Musical aptitude predicted both receptive and productive L2 phonetical ability, but not syntax and lexical knowledge, thus demonstrating that musical skills may facilitate the acquisition of L2 sounds.

Finally, Milovanov et al. (2008) examined the relationship between musical aptitude (as measured by Seashore’s test) and L2 pronunciation skills (word discrimination and repetition), comparing children with superior performance in foreign language production with children with less-advanced production skills. Sound processing accuracy was examined by means of Event-Related Potential (ERP) recordings and behavioral measures. Results showed that children with good linguistic skills had better musical skills than children with less accurate linguistic skills. Moreover, the ERP data showed that children with good linguistic skills have more pronounced sound-change evoked activation with the music stimuli than children with less accurate linguistic skills. The authors conclude that musical and linguistic skills could partly be based on shared neural mechanisms.

3.4. Empirical evidence from our study on musicality and phonetic ability

In our current research project we pursued the question whether musicality as a whole, as well as various more detailed musicality measures, correlate with linguistic abilities, especially talent for L2 pronunciation. For this aim we employed various measures for musicality:

1) Gordon’s AMMA (extensively reviewed in section 2.3.), composed of two subscales: i) a scale for rhythm discrimination ability; and ii) a scale for pitch discrimination ability; which results in a total score of musicality.

2) An additional introspective questionnaire eliciting self reported abilities in the domain of music: i) singing capacity (performance) and the liking for singing; ii) dancing ability (performance) and the liking for dancing; and iii) instrument playing (the number of instruments

---

4. Event-Related Potentials (ERPs) are a neurophysiologic technique consisting in the systematic averaging of many Electroencephalografic (EEG) samples following the presentation of a certain stimulus. By averaging many samples, the noise in the evoked signal is reduced, while the commonalities are enhanced. This results in a graph representing the electrical activity of a given neural pool as a function of time (expressed in milliseconds). Each electrical (positive or negative) peak appearing in the graph is then identified as a “component” associated with a specific stage of the cognitive processing of the stimulus.
played, the performance therein, and the degree of enjoyment connected to the playing of an instrument). The scale for the self scoring ranged from 1 point (minimum) to 5 points (maximum).

We correlated these scores revealing musicality to the second language related tests, which comprised:
1) an L2 pronunciation talent score (see chapter 2);
2) an L2 pronunciation performance/proficiency score (see chapter 2);
3) the MLAT (Modern Language Aptitude Test by Carroll & Sapon), short form, (described in more detail in chapter 2) with three subtests:
   3a) MLAT3 (phonetic coding ability);
   3b) MLAT4 (grammatical sensitivity);
   3c) MLAT5 (vocabulary learning);
   3d) MLAT total score subsuming the three subparts;
4) a subtest of the TOEFL battery for English grammar (see chapter 2)

Results and discussion

Results are reported for a cohort of 66 individuals (33 males, age range 20–40 years, males: mean age 26.49 years +/- 5.36; females: mean age 25.31 years +/- 4.47) taking part in the study. Correlation coefficient ($r$) was computed after Pearson, 2-tailed, with a level of probability of $P < .05 (*)$ and $P < .01 (**)$.

Gender differences were not detected, with the exception of the liking for dancing ($t$-test for independent samples yielded highly sig. difference, $P = .001**$, higher scores for females) and the performance or capacity to dance (with $P = .033*$, higher scores for females). However, as we will see later, dancing was not amongst the musical abilities to predict any of the language related skills.

First of all, a trivial result emerged. All musicality tests were highly correlated amongst each other and likewise, all linguistic measures were highly correlated amongst each other with correlation coefficients ranging from .3 to .9, $P < .01**$. This underlines the validity of our tests and measures taken.

Our aim was to see how the linguistic measures for the various L2 abilities (aptitude and performance) interact with the musicality measures.

Results are summarized in Fig.1 (see Figure Nardo & Reiterer 1 in the Colour figure section).
Most importantly, the two L2 aptitude measures which correlated in a significant way with all the employed musicality scores were: our own pronunciation talent score and MLAT4 score (grammatical sensitivity).

Within the musicality measures, the pronunciation talent score correlated most highly with both the liking for singing and the self reported singing capacity ($r = .4$, $P = .000**$, see figure 1) and secondly with Gordon’s rhythm score ($r = .24$, $P = .01*$).

The aptitude test for grammatical sensitivity (MLAT4) also yielded its strongest and highest correlation with the musicality measures of singing ($r = .31$, $P = .001**$), followed by the liking for playing an instrument ($r = .26$, $P = .005**$).

Further language ability measures which correlated with the scores measured by Gordon’s AMMA (pitch and rhythm) and the questions about singing (but not with playing or liking instruments) were: the actual pronunciation performance, MLAT3 (perceptive phonetic coding ability) as well as MLAT total score and the TOEFL grammar test. (For details see Fig. 1.)

The only test from the linguistic measures with almost no significant correlations (except for a correlation at the lower end for the rhythm perception – $r = .19$, $P = .034*$) was the MLAT5, the vocabulary learning subtest of the aptitude battery. In this test the task was to quickly learn new unknown L2 vocabulary, which is a skill that belongs to the lexico-semantic domain and draws on the capacity for associative memory.

In conclusion, we can say that the strongest correlations for musicality as measured by almost all our subtests (except for dancing ability) were found in productive phonetic talent (as measured by our pronunciation talent score) as well as the aptitude for grammatical sensitivity.

Among the musicality measures, the score which correlated in a significant way with all the language measures, was the rhythm subscore, closely followed by the score for pitch discrimination and then by the self evaluated singing scores (liking singing and singing capacity). The score complex instrument playing (number of instruments played and self reported ability and liking for this instrument) provided fewer interactions with the linguistic measures, except for MLAT4 and the pronunciation talent score (see Fig. 1). Within the last mentioned interactions, the strongest correlation was found between aptitude for grammatical sensitivity and the degree of enjoyment with which one plays an instrument (liking to play the instrument; $r = .26**$, $P = .005$). Furthermore, our instrument measures did not result in any significant correlations to the pure performance score of pronunciation, the phonetic coding ability (a perceptive aptitude
measured by the **MLAT3**), *vocabulary learning* (**MLAT5**) and the **TOEFL**
test.

No correlations were found between the language measures and self
reported *dancing* score (*capacity* and *liking for dancing*). In our opinion, this result is not unexpected, since *liking to dance* was not hypothesized to be related to pronunciation ability in L2, and largely depends on social environment and conventions. *Liking for dancing* was also the only measure that yielded highly significant differences between the sexes.

In short, the conclusions we can draw from this ongoing research are that musicality, ideally in the form of a well developed rhythm perception ability together with a good pitch perception ability and an enhanced ability and liking for singing, are the best ingredients for achieving talent and expertise in foreign language pronunciation from the experimental point of view based on our current studies.

### 4. Neuroscientific evidence

#### 4.1. Overview

In the previous sections, we have reviewed empirical contributions which demonstrate the existence of a relationship between language and music processing. There is evidence that musical training and expertise influence language processing and that certain musical abilities (above all rhythm and pitch discrimination) are associated to a certain degree with language abilities, especially in the phonetic/phonological domain.

In the last decade, the issue of a common neural substrate of language and music has become central. The recent advancements in neuroimaging techniques (PET, fMRI, MEG) have allowed scientists to investigate both the neural structures and the functioning underlying higher-order cognitive processes in a new way. There is an increasing number of studies which give empirical support to the hypothesis that language and music are processed, at least partially, in the same brain structures.

The first studies that investigated the neural substrates of music processing on the basis of neuropsychological evidence (Milner 1962; Kimura 1964), pointed at a difference in the hemispheric lateralization, suggesting that language processing was lateralized to the left, whereas music processing was processed on the right side. However influential, such findings have been proved wrong by subsequent studies. First, it was found that the lateralization of music processing was affected by the level
of expertise (Bever & Chiarello 1974), which is able to modify the networks on the basis of neural plasticity (for expertise and plasticity in L2, see also chapter by S Reiterer). Second, it was claimed that music (like language) is no single entity, but should be decomposed into different components (or levels of processing) which the literature has shown to be processed in different brain structures (Besson & Schön 2001). Modern concepts emphasize the modular organization of music cognition, that different aspects of music are processed in different (although partly overlapping) neuronal networks of both hemispheres (Altenmüller 2001).

In an influential paper, Patel and Peretz (1997) have focused on the music-language relation, criticizing the literature reporting cases of amusia without aphasia (and vice versa) as evidence of no cognitive overlap between music and language. In fact, according to them, such an argument does not take into consideration the subcomponents of music and language. Moreover, cases of aphasia without amusia are generally found in exceptional individuals such as conductors and composers. Finally, aphasia does not include all disorders of language. Patel and Peretz suggest that music–like–language is a confluence of interacting cognitive processes rather than an indivisible whole, and report various studies investigating different aspects of musical structure and their relationship to linguistic structures (Patel & Peretz 1997). Aspects under consideration include melody (melodic contour, pitch, and tonality) rhythm (tempo,

5. Modularity of mind is a theory of mental architectures proposed in 1983 by Jerry Fodor. According to him, some psychological mechanisms (typically perceptive and language systems) would be organized as mental modules, having at least more than one of the following characteristics: rapidity of operation, automaticity, domain-specificity, informationally encapsulation (that is independency from other modules and from the central processing), neural specificity and innateness.

6. Amusia refers to a disorder which consists in the inability to recognize musical pitches, melodies or rhythms or to reproduce them. Amusia can be congenital (if present at birth), or acquired, (i.e. following a brain damage).

7. Melodic contour is the general shape of a melodic line, that is its patterns of ups and downs in pitch directions over time, without regard to the exact pitch intervals, a very salient feature in melodic perception.

8. Tonality can be defined as a system of organizing pitch in which a single pitch (the tonic) is made central and serves as a reference point for the others. It is referred to as “the musical syntax” because it involves orderly structural relations embodied in the implicit knowledge of an experienced listener.

9. Tempo refers to the rate of auditory events in music.
grouping\textsuperscript{10}, and metre\textsuperscript{11}) and song. Results suggest an association between performance on musical contour and linguistic intonation tasks, as well as common mechanisms between grouping in language and music.

4.2. Music and language processing meet in the brain

Previous functional imaging studies have reported that musical tasks activate language areas and vice versa, suggesting that music and language share neural substrates (Gaab et al. 2003; Gaab & Schlaug 2003; Koelsch et al. 2003; Reiterer et al. 2005, 2008). The majority of studies investigating the common neural substrate of music and language processing have predominantly focused on syntax. However, in a study on basic acoustic processing (discriminating subtle differentiations in timbre) performed by our own research group (Reiterer et al. 2008), we found evidence that timbre or quality of tone presented in isolated synthesized tones (neither in the context of music, nor language) activated left Broca’s area.

The notion of a musical syntax has been proposed (Swain 1997; Koelsch et al. 2004), but its rules are difficult to define concretely. Although music consists of discrete elements, its organization is largely relational, especially in Western culture. In fact, in the Western tonal system, the listener’s interpretation of a given note is substantially influenced by both the preceding and the simultaneous notes, and each note contributes to form the framework within which the subsequent notes will be interpreted (Limb 2006). This principle of contextual influence or “embeddedness” into a surrounding frame is similar to the theory of co-articulation in the field of phonetics (see chapter by H. Baumotte). This feature of Western music leads to the notion of musical key (i.e., the relational characteristic of musical pitches), which allows the transposition of a melody into different keys, where although the absolute frequencies of pitches are altered, the contour of the melody is preserved. Similar relational organizations are also valid for rhythmic and harmonic principles.

\textsuperscript{10} Grouping is the clustering of adjacent elements into larger units (phrases) while listening to music.

\textsuperscript{11} Metre is the periodic temporal-accentual scheme, or the number of pulses between the more or less regularly recurring accents. In music grouping boundaries are not predictable from the metrical scheme, that is, metre and grouping are separate though interacting aspects of rhythm (see Lerdahl & Jackendoff 1983).
One of the most important consequences of such relational nature of music is the creation of strong expectations in the listener, based on the internalization of certain variables as a consequence of exposure and enculturation. This way, the vast majority of music listeners are accustomed to hear notes that fit the melodic, rhythmic, or harmonic contextual reference. In music, these expectancies are considered as a sort of vague but robust syntax (Limb 2006) and can be exploited by violating them in a controlled way in order to provoke cerebral responses capable of revealing what happens in the brain when musical “syntax” is violated. For instance, if the last note of a melody played within a single key is out of key, the listener immediately detects a syntactic aberration. This procedure has proven to be very effective also with non-musicians, who are very sensitive to this kind of violations (Koelsch & Friederici 2003). In language studies a similar paradigm would be semantic or syntactic mismatch studies, where anomalies in syntactic or semantic structure have to be detected.

Patel (2003) has pointed out the contradictory findings of the research on the neural correlates of syntax in language and music. In fact, whilst neuropsychological evidence shows that linguistic and musical syntax can be dissociated (Peretz 1993; Peretz et al. 1994; Griffiths 1997; Ayotte et al. 2000, 2002), neuroimaging data support the idea of an overlap in the processing of syntactic relations in language and music (Patel et al. 1998; Maess et al. 2001; Tillman et al. 2003; Koelsch et al. 2002). According to the author, this fact can be accounted for by claiming that syntax in language and music share a common set of processes (instantiated in frontal brain areas) that operate on different structural representations (in posterior areas).

Support to the idea of a common neural substrate for syntactical processing in music and language mainly rely on two kinds of finding, the evidence of a recruitment of the same neural structures (especially Broca’s and Wernicke’s areas, and their homologues on the right side), and the evidence of similar brain wave responses. A study with musicians by Patel et al. (1998) compared ERPs elicited by syntactic structural incongruities in language and music. By employing the violation of principles of phrase structure and principles of harmony and key-relatedness, the authors constructed sequences where an element was either congruous, moderately incongruous, or highly incongruous with the preceding structural context. Results showed that both linguistic and musical incongruities elicited the same component (P600), previously considered to be language specific.
Another study (Maess et al. 2001) investigated the relational (syntactical) properties of Western tonal music with MEG. In this study, non-musicians were presented a series of key musical chord sequences that occasionally contained so-called Neapolitan or sixth chord (which contains two out-of-key notes while being both major and consonant in character), allowing the examination of responses to musical chords that vary according to the musical expectancies created by the preceding chords. Results showed the formation of an early right anterior negativity (ERAN) during the Neapolitan chord presentation, generated in left Broca’s area and its right homologue, well-known key regions for syntactic processing of language.

An fMRI study (Levitin & Menon 2003) examined the brain responses of participants who listened to classical music and scrambled versions of that same music (the latter disrupting the musical structure while holding psychoacoustic features). Comparing music to its scrambled counterpart, the authors found an activation in the left inferior frontal cortex (Brodmann area 47), a region closely associated with the processing of linguistic structure in spoken and signed language, and its right hemisphere homologue, suggesting that this region may be responsible for processing fine structured stimuli that evolve over time, and are not merely linguistic.

A series of studies have systematically employed the violation of expectations in chord sequences with various groups of subjects (male and female, children and adults, musicians and non-musicians). An fMRI study by Koelsch et al. (2002) revealed that unexpected chords activated Broca’s and Wernicke’s areas, superior temporal sulcus, Heschl’s gyrus, planum polare and temporale, and anterior insula, structures previously thought to be domain-specific for language processing. In another study (Koelsch et al. 2003), where ERPs were recorded in 5- and 9-year-old children, it was found that the degree of inappropriateness of the chords modified brain responses according to music-theoretical principles in both age cohorts. Moreover, gender differences were found, resembling lateralization patterns typical of language processing (left predominant in boys, bilateral in girls). Finally, another fMRI study (Koelsch et al. 2005) confirmed and extended previous findings in three groups of subjects: 10-year-old children, adults non-musicians, and adult musicians. In adults, irregular chords activated structures mediating cognitive aspects of musical syntax processing, such as the inferior frontal gyrus, anterior insula, superior temporal gyrus and sulcus, and supramarginal gyrus. Whilst in the right hemisphere the activation pattern of children and
adults was similar, on the left, adults showed larger activations in pre-frontal and temporal areas. Moreover, in both adults and children, musical training correlated with stronger activations in the frontal operculum and superior temporal gyrus.

In a further study by our own research group (Reiterer et al. 2005) on basic pitch and duration discrimination, we found that task difficulty and not the stimulus characteristics per se, was another modulating factor affecting hemispheric involvement within the classical auditory processing areas. We found more right hemispheric involvement for both, the pitch and the duration discrimination task, when the task was easier, i.e. the discrimination between two stimuli could be achieved easily.

Speaking of semantics in the music domain seems strange, given that music is rather abstract and has little explicit reference to the external world. However, although still under debate, the question has been posed whether a musical phrase can convey meaning, and whether this can be proven.

According to Koelsch (2005), music can transfer meaningful information and is an important means of communication. Theorists distinguish between four different aspects of musical meaning: i) emerging from common patterns or forms (e.g., musical sound patterns that resemble sounds or qualities of objects); ii) arising from a particular mood; iii) inferred by extramusical associations; and iv) stemming from combinations of formal structures that create tension (e.g., an unexpected chord) and resolution (Meyer 1956). The emergence of this latter kind requires an integration of both expected and unexpected events into a meaningful musical context. The processing of such musical integration seems to be reflected in a late negative component evoked by unexpected (irregular) chords, which is substantially similar to the N400 component elicited by the processing of semantic integration during the perception of language. The N400 amplitude correlates with the amount of semantic integration required by a word and, similarly, with the amount of harmonic integration required by a musical event (Koelsch et al. 2000).

In an electrophysiological study with healthy subjects, Koelsch et al. (2004) examined whether the priming effect caused by presenting semantically related words in sequence (which evoke the N400 component) could also apply to music. Results showed that a semantic priming effect was also observed when target words (i.e., semantically unrelated to a preceding musical excerpt) followed musical excerpts. The N400 component did not differ between the language condition and the music condition with respect to amplitude, latency or scalp distribution, and the ef-
fect was observed for both abstract and concrete words. Moreover, in both conditions the main sources of these effects were localized bilaterally in the posterior part of the medial temporal gyrus (BA 21/37), in proximity to the superior temporal sulcus, regions implicated in the processing of semantic information during language processing (Friederici et al. 2000; Friederici 2001, 2002; Baumgaertner et al. 2002). Such findings demonstrate that music can activate representations of meaningful concepts, and that the cognitive operations underlying meaning decoding can be identical in language and music processing.

On the other hand, Besson and Schön (2003) have carried out an experiment which shows that lyrics and tunes seem to be processed in an independent way, giving support to a domain-specificity of semantic processing. Conceptually similar, research with fMRI by our own groups (Riecker et al. 2000) investigating overt singing and speaking found that singing is predominantly lateralized to the right and speaking to the left hemisphere.

4.3. The neural substrate of musicality

Unfortunately, the neural correlates of musicality have been very poorly investigated to date. An interesting study (Norton et al. 2005) has compared children who were about to begin an instrumental training with controls who were not, in order to determine whether there are: i) a priori structural neural differences (i.e., innate markers of musical ability) between the groups; ii) differences in other cognitive skills between the groups; iii) correlations prior to music training between perceptual musical skills (as measured by Gordon’s PMMA) and other outcomes (cognitive, motor, or neural) possibly associated with music training. Results showed no pre-existing neural, cognitive, motor or musical differences between the groups, as well as no correlations between music perceptual skills and brain or cognitive measures. However, correlations were found between music perceptual skills and phonemic awareness, suggesting the existence of a common neural substrate for language and music in the phonetic domain.

Recent findings suggest that Heschl’s gyrus\(^\text{12}\) could constitute a possible marker of musicality (as well as linguistic talent, see chapter by S Reiterer). Schneider et al. (2002) have conducted a magnetoencephalo-

---

12. Heschl’s gyrus is the primary auditory cortex in the human brain, located within the Sylvian scissure and corresponding to Brodmann area 41.
graphic (MEG) and structural imaging study in which professional musicians, amateur, and non-musicians performed an auditory processing task. Professional musicians showed a significantly greater increase in MEG activity within primary auditory cortex compared to non-musicians, which in turn was found to correlate with increased volumetric measurements of gray matter within Heschl’s gyrus in musicians compared to non-musicians. Moreover, psychometric testing revealed a positive correlation between the size of Heschl’s gyrus and musical aptitude as assessed by Gordon’s AMMA. These results were confirmed in a subsequent study (Schneider et al. 2005), and although the question of causality could not be addressed, these findings suggest a fundamental link between musical exposure, musical aptitude, and the physiologic and anatomic development of Heschl’s gyrus.

5. Conclusions

In this chapter we have examined the concept of musicality and reviewed the major experimental contributions concerning its relationship with language talent, ability and linguistic processing in general. We have started with definitions of musicality and related concepts in order to introduce and disclose the complexity of the topic. We have referred to musicality as a multi-faceted and fuzzy concept conveying the meaning of a collection of musical abilities which rely on both innate predispositions and experience. We have defined “talent” (or “aptitude”) as the innate component, and “musicality” (or “musical ability”) as a complex skill stemming from the interaction between innate and acquired factors.

To date, the “nature vs. nurture” problem seems far from a solution, because both standpoints are supported by evidence. Clearly, both of them play an important role in the shaping of the actual musical abilities. However, we suggest that the debate and the efforts towards a precise determination of the relative contribution of nature and nurture could be misleading, and possibly out of reach. Thus, future research could start from the assumption of the necessity of both factors, and focus more on their interrelation rather than on their relative pre-eminence.

We have seen that both theories and evidence support the idea of musicality as something different from general intelligence. However, some authors (i.e., Wing, Gardner, Reimer) use the term “musical intelligence” to refer to the cognitive component implied in musical abilities. Yet such musical intelligence is by no means meant as something related to the
logical-abstract ability which we normally associate with general intelligence, but rather with an independent skill. On the basis of the reviewed literature and our own current research, we can assert that, far from being a unitary entity, musical ability is made up of several sub-components which align themselves along a continuum ranging from most basic psychophysical skills (i.e. pitch discrimination, rhythm perception, timbre and intensity sensitivity, etc.), to the highest cognitive abilities (tonal representation, aesthetic appreciation, ability to create or improvise, etc.), also including motor skills (from tempo-tapping to accurate performance and improvisation). These sub-components are probably not isolated, but interact with each other, as well as with abilities in other domains (i.e. sub-components of language processing like phonetic perception and production, memory, imagery, creativity, etc.). We suggest that future research about musicality and language processing should take into account the complexity of these phenomena and would benefit from considering their sub-components in major details.

As regards the tests of musical aptitude, we have reviewed the most popular ones. To date, Seashore’s test has been heavily criticized for its atomistic and psychophysical approach. However it is still employed in those studies aimed at investigating the basic levels of musical talent. On the other hand, Wing’s and Gordon’s tests are more “cognitive”, and several versions of the latter make it particularly suitable for research with different age and expertise cohorts.

Our own research as well as the literature reviewed consistently show that language and music are not independent phenomena, but perhaps two sides of one coin with a lot of similarities, yet not being exactly the same. We have seen that music practice improves language processing, and that some aspects of music and language processing are correlated, especially rhythmic processing and phonetic aspects. In our own research we found that there are strong links between rhythm and pitch perception and singing capacity on the one hand, and pronunciation talent, pronunciation performance/proficiency, phonetic encoding ability and even grammatical sensitivity and proficiency on the other.

Moreover, when considering the neural substrates, it clearly emerges that the syntactic, semantic and phonetic aspect of both linguistic and musical processing share the same networks in a substantial way. This evidence contrasts the idea of a strict domain-specificity and suggests that language and music share some common cognitive operations which are involved in both processes. Conversely, the investigation of the neural correlates of musical talent has just begun, and will certainly in-
crease in next years, hopefully taking into account the complexity of musicality.

Finally, on the basis of the present work, the important arising question is “what is the exact relationship between musical skills, language skills, and cognitive processes?” Future research is still needed to investigate the nature of this relationship, in order to determine which cognitive operations underlie the various musical and language abilities, to bring forth the commonalities between both. This would advance our understanding of how the single components could best be exploited to improve one another.

References

Altenmüller, E. O.


Arellano, S. I. & Draper, J. E.

Arleo, A.

Atterbury, B. W.

Ayotte, J., Peretz, I., Rousseau, I., Bard, C. & Bojanowski, M.

Ayotte, J., Peretz, I. & Hyde, K.

Baily, J.

Bamberger, J.

Baumgaertner, A., Weiller, C. & Büchel, C.
Benson, D.F., Dobkin, B., Gonzalez-Roth, L. J., Helman-Estabrook, N. & Ker-
tesz, A.

Besson, M. & Schön, D.
2001 Comparison between language and music. Ann NY Acad Sci, 930:
232–58.

Besson, M., Schön, D., Moreno, S., Santos, A. & Magne C.
2002 Influence of musical expertise and musical training on pitch processing

Besson, M. & Schön, D.
2003 Comparison between language and music. In Peretz, I. & Zatorre, R.J.
(Eds.). The cognitive neuroscience of music. New York: Oxford Uni-
versity Press.

Bever, T. G. & Chiarello, R. J.
1974 Cerebral dominance in musicians and non-musicians. Science, 185:
537–539.

Brutten, S. R., Angelis, P. J., & Perkins, K.
1985 Music and memory: Predictors for attained ESL oral proficiency. Lan-
guage learning, 35: 299–313.

Carroll, J. B.
1993 Human Cognitive Abilities. A survey of factor-analytic studies. Cam-
bridge, New York: Cambridge University Press.

Chan, A. S., Ho, Y. C., & Cheung, M. C.

Clarke, E. F.
1993 Imitating and evaluating real and transformed musical performances.

Costa-Giomi, E.
1999 The effects of three years of piano instruction on children’s cognitive de-

Davidson, J.
1993 Visual perception of performance manner in the movements of solo musi-

Deutsch, D.
Music Perception, 8: 335–47.

Dexter, E. S., & Omwake, K. T.
1934 The relation between pitch discrimination and accent in modern lan-

Douglas, S., & Willatts, P.
1994 The relationship between musical ability and literacy skills. Journal of

Eterno, J. A.
1) Foreign language pronunciation and musical aptitude. Modern Lan-
Ewers, D. W.  

Fish, L.  

Fodor, J.  

Fonseca Mora, C.  

Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B. & Oertel, U.  

Friederici, A. D.  

Friederici, A. D.  

Gaab, N., Gaser, C., Zaehle, T., Jancke, L. & Schlaug, G.  

Gaab, N. & Schlaug, G.  

Gabrielsson, E. F.  

Gardner, H.  

Gilleece, L. F.  

Gordon, E. E.  

Gordon, E. E.  
1979  Primary Measures of Music Audiation. Chicago: GIA.

Gordon, E. E.  

Gordon, E. E.  
1989a  Advanced Measures of Music Audiation. Chicago: GIA.
Gordon, E. E.
1989b *Audie: A game for understanding and analyzing your child’s musical potential*. Chicago: GIA.

Green, L.


Harrison, N.
1979 *A study to find out if a correlation exists between the study of music and aural-oral ability in French*. M.A.T. of F. Simon Fraser U.

Holmes, J. A.

Jaffurs, S.

Jakobson, L. S., Cuddy, L. L., & Kilgour, A. R.

Jørgensen, H.
2008 *What is a talent, how is it identified, and how is it developed?*. Paper presented at the conference “Precollege education and talent development”, European Association of Conservatories (AEC, Oslo May 30, 2008).

Jusczyk, P. W.

Karimer, L.
1984 *Can southeast Asian students learn to discriminate between English phonemes more quickly with the aid of music and rhythm?*. Language Key to Learning, 5: 41–48.

Kimura, D.

Koelsch, S., Gunter, T. C., Friederici, A. D. & Schroger, E.

Koelsch, S., Gunter, T. C., Cramon, D.Y., Zysset, S., Lohmann, G., Friederici, A. D.

Koelsch, S., Grossmann, T., Gunter, T.C., Hahne, A., Schröger, E. & Friederici, A. D.
Koelsch, S., Maess, B., Grossmann, T. & Friederici A. D.

Koelsch, S. & Friederici, A. D.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T. & Friederici A. D.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D. & Schlaug, G.

Koelsch, S. & Siebel, W. A.

Kratus, J.

Kratus, J.

Lerdahl, F. & Jackendoff R.

Leutenegger, R. R., Mueller, T. H., & Wershow, I. R.

Levitin, D. J. & Menon, V.

Limb, C. J.

Lowe, A. S.

Maess, B., Koelsch, S., Gunter, T. C. & Friederici A. D.

Mainwaring, J.
McPherson, G. E.

Meyer, L. B.

Milner, B.

Milovanov, R., Tervaniemi, M., Takio, F., & Hämäläinen, H.

Milovanov, R., Huotilainen, M., Välimäki, V., Esque, P. A., & Tervaniemi, M.

Morgan, C.

Music Educators National Conference (MENC).

Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D. J. & Schlaug, G.
2005 Are there pre-existing neural, cognitive, or motoric markers for musical ability? Brain Cogn, 59(2): 124–34.

Palmer, C. & Kelly, M. H.

Pastuszek-Lipinska, B.

Patel, A. D. & Peretz, I.

Patel, A. D, Gibson, E., Ratner, J., Besson, M. & Holcomb, P. J.

Patel, A. D.

Peretz, I.

Peretz, I. Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., Belleville, S.

Pinel, S.

Reimer, B.

Reiterer, S. M., Erb, M., Droll, C. D., Anders, S., Ethofer, T., Grodd, W., & Wildgruber, D.

Reiterer, S. M., Erb, M., Grodd, W., & Wildgruber, D.

Révész, G.
1953 Introduction to the psychology of music. London: Longmans, Green.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W.


Seashore, C. E.

Seashore, C. E.

Shuter-Dyson, R., & Gabriel, C.
1981 The psychology of musical ability. London: Methuen.

Shuter-Dyson, R.

Slevc, L. R., & Miyake, A.

Sloboda, J.

Spearman, C.
Spychiger, M.
1993  *Music makes the school*. Die Blaue Eule.

Sternberg, R. J.

Stevenson, R.

Stokes, J. D.

Sutton, J.

Swain, J. P.

Swanick, K.

Swanner, D. L.
1985  *Relationships between musical creativity and selected factors, including personality, motivation, musical aptitude, and cognitive intelligence as measured in third grade children*. Unpublished dissertation, Case Western Reserve University, Cleveland, Ohio.

Threub, S. & Trainor, L.

Tillmann, B., Janata, P., Bharucha, J. J.

Tomatis, A.

Tucker, A.

Vaughan, M. M.

Wang, C.
1985  *Measures of creativity in sounds and music*. Unpublished manuscript, University of Kentucky School of Music, Lexington.
Webster, P. R.
1983 *Measures of creative thinking in music*. Unpublished manuscript, Northwestern University School of Music, Evanston, IL.

Webster, P. R.

Wheeler, L. R., & Wheeler, V. D.

Wing, H. D.

Wing, H. D.
Sociolinguistic factors in language proficiency: phonetic convergence as a signature of pronunciation talent.

Natalie Lewandowski

1. Adaptation processes in dialog

Research on adaptation processes in dialog has developed within the framework of Communication Accommodation Theory (CAT), which was established in the 1970s, originally under the term “speech accommodation theory” (SAT, Giles 1973). SA theorists then started proposing new explanations for some aspects of the Labovian paradigm (Labov 1966) concerning the formality-informality of context. According to them, a lot of details could be explained by interpersonal accommodation processes. Furthermore, they called for a more detailed concept of context, which was a major term in sociolinguistic theory at that time. One of the specific goals was increasing the attention for specific dimensions of context, like language itself and the role of the receiver in the interaction (Giles, Coupland & Coupland 1991). Since then, many topics have been under the focus of SA researchers, involving the clarification of the motivations underlying speech, the constraints imposed on it and finally their consequences for social interaction (Giles et al. 1991). Under the new name of communication accommodation theory, the range of topics has been extended to nonverbal behavior and general dimensions of discourse (Giles et al. 1987, cit. after Giles et al. 1991).

Communication accommodation researchers, being nowadays placed at the interface between sociology, social psychology, sociolinguistics and communication, have also adopted a broader perspective of macro-variation to their studies, instead of solely concentrating on interindividual exchange. It has been found that language used in interaction can be a remarkably good indicator of status differences, ethnic boundaries or can as well define ingroup or outgroup boundaries and also impose conformist behaviors (Shepard et al. 2001).

Language, as claims the crucial thesis of CAT, is used to achieve a desired degree of social distance between the self and an interacting partner.
Accommodation is perceived as a complex, context-sensitive set of alternative behaviors available to interacting partners in a face-to-face situation. Given a dynamic setting and online feedback, partners in dialog can achieve solidarity with or dissociation from one another (Giles et al. 1991).

Pickering and Garrod (2004, 2006) contrast this view of an underlying social dimension in accommodation processes with an automatic tendency for alignment in dialog. Without assuming any strategic components involved, they present arguments for graded automaticity in speech, discussing Levelt’s speech production model for monolog speech (1989) and aspects of it they consider being subject to automatic processes. They argue that dialog speech should be affected by automatization even to a greater extent than monolog since dialog imposes certain constraints on choices a partner can make (e.g. speech act planning), people show an apparent ease of conversation and most of them have overall much greater experience with dialogs than monologs (e.g., prepared talks; Pickering & Garrod 2006).

This chapter uses both terms – accommodation and alignment – for adaptation processes in dialog, regardless of the automatic vs. strategic distinction between them. I will argue for a partly controlled and partly aware process, affected by a number of internal features not subject to control, enhancing or limiting the degree of accommodation (supposedly talent, personality traits), as well as external factors imposed on the participants through the situational setting.

1.1. Accommodation strategies

Convergence, one of the approximation strategies within the CAT framework, stands for the adaptation of communicative behaviors towards those of a conversational partner both in verbal and nonverbal features. Adaptation can cover features as, e.g., gestures, smile, facial affect, head nodding, information density, voice quality, speech rate, utterance length, pausing frequencies and response latency (Giles et al. 1991).

The set of behaviors subsumed under the notion of divergence allows speakers to distance themselves from their partners in terms of verbal and nonverbal behavior. Possible realizations most commonly involve an explicit accentuation of differences in speech style and/or facial expressions and gestures (Giles et al. 1991). Possible discrepancies between the performance and the actual perception of a conversational partner have acquired the name of “perceptual/subjective divergence” (Shepard et al. 2001).
In case no convergence or divergence is observed, the speaker might be trying to keep her own speaking and behavioral style. Such intended *maintenance* may, however, often be interpreted as diverging from the interlocutor (Shepard et al. 2001).

Different social roles and the situational context have also been proven to influence partners’ behavior in interaction. Men were, e.g., observed speaking with a lower voice while talking to women (Hogg 1985) and the power structure within a dyad influences the choice of up- or downward accommodation (towards a higher or lower socially rated speech variety; Street 1982; Giles et al. 1991).

1.2. Research on phonetic convergence

Despite the relatively large number of studies devoted since the 1970’s to accommodation in lexical, syntactic and phonetic aspects of speech, it has been only recently that phonetic convergence has been defined in terms of an increase of segmental and suprasegmental similarities between the interacting speakers (Pardo 2006).

Phonetic convergence so far has been studied by extracting features such as speaking rate (e.g., Street 1984), fundamental frequency and amplitude contours (e.g., Gregory 1990), voice onset time (Sancier & Fowler 1997), amplitude, utterance duration and rate within a human-computer interface (Oviatt, Darves & Coulston 2004) or by conducting perceptual similarity experiments (Pardo 2006). Most of the studies though investigated phonetic convergence in a strictly native language environment. Sancier and Fowler (1997), as one exception, concentrated on the gestural drift between the two languages of an American English-Brazilian Portuguese bilingual speaker. The essential role of accommodation in a second language acquisition context has been recognized by Zuengler (1991) in her analysis of data on native-nonnative or fully nonnative conversational interactions.

Research on phonetic accommodation involuntarily raises the question of the nature and role of the perception-production link. Sancier and Fowler (1997) propose that speech perception automatically yields relevant linguistic (in this case phonetic and phonological) parameters that cause production and lead to imitation in a direct way. Given this assumption and the fact that there indeed is a very close link between perception and production, it should favor a fairly exact imitation of items at an articulatory and acoustic level. However, even for a single speaker no two productions of the same utterance are acoustically or articulatorily identi-
cal (Pardo 2006), which speaks against the first account. Supposedly far more probable is an intended imitation with a moderate degree of exactness, influenced by a number of limiting features to be discussed further on. As Pardo (2006) states, none of the accounts having been under discussion within accommodation and convergence literature before explained the influence of obvious perception and production limitations and other factors yielding discrepancies in the imitative responses.

An interesting new perspective on convergence research is presented by Wedel and Van Volkinburg (to appear). According to their study, convergence and divergence can also be part of a simultaneous process, with some features underlying an automatically driven positive accommodation procedure and others functioning as identity markers for a group membership and therefore subject to divergence. Such apparently countervailing and still simultaneously observed strategies might be best accounted for in a dynamical systems framework of language that incorporates usage-induced changes in that system.

2. A usage-based account of language

In order to study convergence we assume a usage-based account that allows for the storage and processing of fine phonetic detail and additional social information, i.e., the exemplar theoretical account (e.g., Johnson 1997; Goldinger 1996, 1998; Bybee 2002, 2006; Pierrehumbert 2001, 2006). Exemplar theory first emerged as a model in psychology and was further developed for speech processing and subsequently re-modelled by Goldinger (1996, 1998), Johnson (1997, 2006) and Pierrehumbert (2001, 2006). At present exemplar-based accounts are used in phonetics and phonology, as well as in semantics, lexicology, typology (Bybee 2002, 2006), syntax (Bod 2006) and language acquisition (Abbot-Smith & Tomasello 2006).

It is one’s experience with language that is taken to be central and it is the cognitive organization of this experience that eventually builds up grammar (Bybee 2006). Bybee (2006:711) proposed in detail that the input “the general cognitive capabilities of the human brain [get], which allow it to categorize and sort for identity, similarity, and difference” are the specific linguistic events a person encounters, which are then categorized and stored in memory. She also raises doubts about the strict separation between grammar and lexicon, given the assumption that both are involved in a dynamic process of updating through language experience. Implications following as a natural consequence from this view
include the change of the description models for linguistic categories, away from traditionally used abstract rules, processes and structures to actual patterns of occurrence of those linguistic categories (Wade et al. to appear; Bybee 2002, 2006). Exemplar-based models provide for exactly such a formal means of description, assuming that all various level categories (be they phonemes, syllables or words) consist of a collection of actually experienced instances of those categories. The processes underlying perception (or identification) and production then only operate on an exemplar level by comparing the items within and between collections. Further specification of occurrence regularities or surface forms of the exemplar categories is not necessary (Wade et al. to appear).

Usage-based accounts like exemplar-based models do also have the explanatory power to deal with discrete and gradient phenomena (e.g., phonetic neutralization, word frequency- or gender- and speaker-dependent acoustic differences), and, moreover, are suggested to “provide the most accurate, parsimonious description of linguistic competence and performance” (Wade et al. to appear:1; Bybee 2002, 2006; Johnson 2006; Pierrehumbert 2001, 2006).

2.1. An exemplar theoretic model of speech perception

Johnson (1997) proposed a model of speech perception where exemplars were seen as associations between a set of auditory properties and a set of category labels, the former defined as being output from the peripheral auditory system and the latter as including any classification of possible importance to the perceiver available at the moment of storage in memory (such as gender, speaker name, etc.). All speaker-specific details are retained in the set of exemplars, which allows for comparing and categorizing new items with reference to appropriate earlier stored exemplars also on speaker-specific dimensions. In case a new item is encountered, the process of categorization involves:

- comparing the new item’s auditory properties with each exemplar’s auditory properties,
- assigning each exemplar an activation level according to its similarity to the new item → the better the match, the higher the activation level,
- summing up the overall activations of all exemplars of a given category.

The last step serves as a basis to decide whether the newly encountered item should be categorized as an instance of that category or not (Johnson 1997).
Johnson’s exemplar-based model (1997) differed from previous perception models in several points. All speaker-specific details are retained in the set of exemplars, which allows for comparing and categorizing new items with reference to appropriate previously stored exemplars also on speaker-specific dimensions. Johnson also added an attention weight parameter to his model that controls the degree of sensitivity to particular auditory properties. Furthermore, it has been suggested that no further speaker normalization processes are needed in this kind of perception model because “the model retains the variability encountered in speech [and] it is able to cope with the variability that it encounters in new tokens” (Johnson 1997:162).

2.2. Frequency effects in exemplar models

After the process of identification a new token is categorized in a cognitive map such that similar exemplars are close to each other and very dissimilar ones far apart. As Pierrehumbert (2001) proposes the exemplar system then works through mapping between points in a phonetic parameter space and the corresponding labels of the categorization system. An important emergent property of exemplar models in general is related to word/syllable frequency. Given that every linguistic experience is categorized and stored in the exemplar space, more frequent categories will automatically have a larger representation of tokens and less frequent categories a less numerous representation. Assuming further that linguistic memory decays and that more recent memories will be more vivid than those from several years ago, Pierrehumbert (2001:141) argues that each exemplar be assigned an associated strength, or in other terms, a resting activation level. Exemplars of newly stored frequent experiences have higher activation levels than exemplars of temporally remote and infrequent experiences.

It plays a crucial role in the classification process of new tokens, since it is not only the distance from any given exemplar in the parameter space that contributes to computing the similarity to a new token but also the strength of that exemplar. After perceptual encoding, the new token is placed in the relevant parameter space, where the computation of distance and the most probable labeling take place (Pierrehumbert 2001). The classification is only influenced by the set of exemplars located in a fixed size neighborhood of the token. The last step consists of calculating the summed similarities to the surrounding exemplars for each label present in that neighborhood, with the similarity to the exem-
plars weighted by their activation level (Pierrehumbert 2001; Lacerda 1997). The label favored in this process is the one having more or higher activated exemplars in the neighborhood of the newly encountered token. This predicts that high frequency categories that are represented by more numerous exemplars with on average also higher resting activation levels will have an advantage in the labeling process.

2.3. The perception-production loop in Exemplar Theory

Following that line of explanation and assuming that exemplars of newly stored frequent experiences are assigned higher activation levels than exemplars of temporally remote and infrequent experiences, it seems to provide a natural explanation for alignment as it is understood by Pickering and Garrod (2004, 2005). It has been suggested that speech patterns which are heard recently and frequently also automatically guide the typical productions within a speech community and therefore lead to the adaptation of the prominent speech patterns (Pierrehumbert 2001).

Bybee (2002) also takes an exemplar model of phonological representations that allows for gradual changes in both the phonetic and the lexical dimension, to account best for phonetically conditioned changes affecting to a different degree high-versus low-frequency words. Her model provides a more in-depth description of vowel reduction and deletion processes in relation to their frequency of occurrence in categories. Based on an exemplar account of speech production, she argues that reductive changes show a tendency to appear earlier and to a greater extent in high frequency words and phrases. This seems to be a natural consequence of the distribution of exemplars within the cognitive space. Exemplar clouds of high frequency categories show a greater density and since their strength depends also on their recency, they display on average higher activation levels. Those exemplars in turn have a higher chance of being chosen for production, during which any existing acoustic variation (such as lenition or deletion) can be strengthened or a new mutation can be initiated. Changes introduced at the level of the individual production are accumulating over time and, considering the relatively quick re-use of exemplars within high frequency categories, occur also more rapidly. The exemplar clouds thus are subject to constant changes and updating processes while language is used (Bybee 2001, 2002; Pierrehumbert 2001; Wade et al. to appear).

While high-frequency categories are more prone to leniting changes as a natural consequence of automation of production and fluency, Bybee
(2002) also suggests different effects applying to lower frequency words. Less frequent words thus seem to be more susceptible to changes towards conformity with the stronger patterns of the language (such as the regularizing of verb patterns) which affect them earlier than their high-frequency category neighbors.

Although exemplar-based production models provide a straightforward explanation for the nature of the perception-production loop, with both processes relying on the same pool of exemplars (e.g., Pierrehumbert 2001), this does not imply that a phonetic target is necessarily realized as a perfect match. Even for the same speaker some random deviations from the acoustic target caused by noise in the motor control and execution seem to be probable. Pierrehumbert (2001:145) speaks about the process of adding new items to an exemplar cloud as a random sampling from that cloud with added noise, which means that recovering an exemplar for production would not assure an identical production of that item. Additionally it could be assumed that the steps of paying attention, recognition and storage are not passed through identically each time an exemplar is being chosen, neither for the same person nor for two different persons. Not every experience raises attention to the same extent—it has been suggested that people are focusing on events classified as being “most informative”. That of course influences the further steps of recognizing and encoding of an exemplar (Pierrehumbert 2006:525). Attention, recognition and storage can thus be seen as relevant sources of disturbances within the perception-production loop and are further adding to the overall level of noise.

2.4. Level of description in exemplar models

As pointed out by Bybee (2002:272), her analysis of frequency grounds on the assumption that words are the standard units in exemplar models that must be present in memory storage in order for the described changes to happen. Other accounts have posited the syllable, morpheme or even a multi-word string as the unit to be categorized in memory, understanding the introduction of variability as a top-down process moving in a hierarchical manner towards the lower level. Newer accounts (such as Pierrehumbert 2006) suggest that the lowest level of description should be a parametric phonetic map instead of any set of discrete categories.

An exemplar-based production model recently developed in Stuttgart assumes that frequency effects could be based on the constituent articu-
Sociolinguistic factors in language proficiency

2.5. Memory in exemplar models

Inherent to many exemplar-based models (e.g., Pierrehumbert 2001) is the assumption of gradual memory decay. Exemplars that have been stored at a distant point in time become blurred and their resting activation level is much lower than that of recently encountered tokens which in turn decreases their chances of being chosen for production.

In the memory sequence model underlying the simulations in Wade et al. (to appear) it has been proposed that everything that has been encountered is stored in full context and that perception and production take as much of this context into account as needed to find a match. As opposed to static models (such as the model described in Pierrehumbert 2001), this is a dynamic model which takes context and time into consideration. A temporal match with a stored sequence is therefore as important as a spectral match. The form the memory sequence thus takes is assumed to be of a spectro-temporal nature, meaning that it considers...
not only the units themselves and their spectral features but units as they are encountered and encoded in time. The memory sequence is suggested to take the form of separate frequency bands, where in case of a match-finding both contexts of the unit under consideration (the left preceding context of an acoustic nature and the right following context with linguistic information) are being compared to the target unit (Wade et al. to appear).

Due to the preservation of detailed information in exemplars, exemplar-based models are taken to be especially well-suited for describing gradient phenomena, as in lexical diffusion and language change in general (e.g., Bybee 2002, 2006). The elaborations of the model have also proved to be suitable for research in the area of sociophonetics (Labov 2006) and sociolinguistics in general, where its advantages compared to the Varbrul program\(^1\) have already been analyzed (Pierrehumbert 2006). Pierrehumbert pointed also to the future direction of advanced exemplar-based models in sociolinguistic research, which can be applied to the modeling of extremely detailed language-specific or dialect-specific patterns, describing probability distributions and capturing their dynamics in cognitive systems and in populations, or in other words, the advanced model “holds promise for explaining how cognitive and social factors interact to form language” (Pierrehumbert 2006:528).

3. Relevance of convergence for the assessment of language talent

Language talent or aptitude is no longer being questioned as one of the sources of individual variation in second language acquisition. Studies on individual differences in general have shown that variation in language talent, personality traits and attitudes (motivation) leads to significant contrasts both in the rate of learning and in the eventually acquired proficiency in a second language (e.g., Dörnyei 2005). Despite this recognition as one of the influential factors amongst others as motivation or learning styles and strategies, there have been hardly any focused and in-depth studies dealing with talent alone so far.

There have been many attempts to capture the notion of talent and explain its nature. Skehan (1998), on the one hand, defines talented learners simply as people with greater levels of aptitude and therefore likely to make quicker progress in language learning. As has been elaborately shown in other papers in this volume, language aptitude can be further divided into sub-areas. Strong evidence supports the fairly independent
functioning of the phonetic subcomponent of talent from other sub-skills (e.g., the well known Joseph-Conrad- or Henry-Kissinger-phenomenon). Especially the pronunciation part in acquiring a second language thus seems to cause substantial difficulties for some language learners but, surprisingly, not for all of them.

We hypothesize that for accommodation in pronunciation towards a foreign language speaking partner not only proficiency in that language is crucial but also a certain degree of phonetic talent. Besides obvious social factors, for phonetic convergence to happen in a native-nonnative interaction, the interacting person must display a certain talent for paying attention to fine phonetic detail and for its quasi-immediate recovery for production. Regardless of possible interfering personality factors, convergence within a short conversation might be an indicator of the speed of acquisition and perhaps even of the final success of acquiring the intended pronunciation.

4. Method

The study recruited 8 native speakers of German (3 female, 5 male) ranging between 20 and 36 years old (average 25.9 years) with no reported hearing or speaking impairments. All subjects had an academic background and learned English as a foreign language mainly in an instructional environment (in German schools regularly since fifth grade) with only few stays in English speaking countries, the vast majority not exceeding a few months. Two additional subjects were recruited for the dialog part of the experiment – one female Southern Standard British English (SSBE) speaker (age 56) and one male speaker of General American (GA) English (age 32).

4.1. Experimental design

The German participants were involved in two dialogs with the two English native speakers carried out within one experimental session. Data were elicited using a Diapix-task (Bradlow et al. 2007), a picture matching game in which participants have to identify ten differences between their pictures without seeing the partner’s picture (see Figure Lewandowsk 1 in the Colour figure section). The control task consisted of three parts inserted before and after the dialogs, each involving reading out a list of words from the two picture-sets together with unrela-
ted filler words. The whole experimental set-up thus was designed as follows: first reading of the word-list, Diapix-dialog A with the GA speaker (or with the SSBE speaker), second reading of the word-list, Diapix-dialog B with the SSBE speaker (or GA speaker), third reading of the word-list.

Describing and comparing the scenes on the pictures allows obtaining a wide range of utterance types and at the same time provides balanced talker roles as opposed to many other techniques supposed to elicit quasi-spontaneous speech from both partners. On the other hand, it also allows controlling for the appearance of a sufficient number of the same target words needed to carry out analyses on the convergent behavior during the dialog (Bradlow et al. 2007). The control task was added for the investigation of the persistence of the convergent effect after the dialog was finished, and the participants were confronted with a not overtly related reading task without any conversational character.

The pronunciation talent and the English language proficiency of the subjects had been rated beforehand in extensive perception and production tests (see Jilka this volume for details), which allowed a classification into two groups – of talented and less-talented speakers of English as an L2.

4.2. Analysis

Both the conversations and the word list reading tasks were recorded in an anechoic chamber using an AKG C520 head-mounted microphone and converted to a digital signal at a sampling rate of 48 kHz and 16 bit resolution. The dialogs were recorded as a stereo wave file with two separate channels, the word lists as a mono wave file.

The raw data set was automatically annotated at segment, syllable and word level by an Aligner tool for English and manually corrected. The participants’ and the native speakers’ utterances (at single word level) were then manually extracted from the dialog and the word lists. Since most studies on convergence in the phonetic domain so far concentrated on individual parameters as utterance duration, $F_0$ contours or speaking rate, ours is the first attempt of finding a global measurement for convergent behavior in dialog. Amplitude envelope signals are such a global means comprising a smoothed picture of the energy present in separate frequency bands.

The procedure of extracting amplitude envelopes is based on the method described in Wade et al. (to appear). The extracted words were equated for root mean square amplitude and then separated into four lo-
garithmically spaced frequency bands from 80 to 7800 Hz using 4th-order Butterworth filters. After high-emphasizing to give more weight to the (lower-amplitude) high frequency range of the sounds, amplitude envelopes were estimated using the Hilbert transform. In order to compare the envelope signals of two words, the envelopes with a cut-off frequency of 60 Hz were again normalized and the match was estimated by a cross-correlation function. The match value ranges from 0 to 1, with 1 indicating a perfect match and 0 no match. The match values were post-hoc transformed $x^3$ in order to fit best the scale.

The dialogs were divided into three parts – early, middle and late. The first 40% of the conversation is labeled “early”, the last 40% is labeled “late”. Only words from these two parts of the conversations are subject to analysis of accommodation effects, leaving out the middle part of the dialogs. Only comparisons within a speaking style: read-to-read and spoken-to-spoken were allowed. That resulted in two types of comparison in the analysis: words from one person in the dialog were only compared to utterances of the respective dialog partner from the same conversation and the read word lists from a native German subject were compared with utterances from word lists of the native speakers of English. A separate variable for analysis was the talent score of each participant (see Jilka this volume and Jilka et al. 2008).

5. Results

The German native subjects were classified into two groups – those of phonetically talented and less-talented speakers of English. Among the talented participants were two female and two male subjects, the other group included one female and three male participants.

5.1. Dialog task

Table 1 shows the differences of the match values between an early and a late point in the dialogs for the two English variety conditions (GA and SSBE) of all eight speakers. The second column displays their talent score and the third their preferred English variety (the one they consider themselves to use most of the time while speaking English or the one they are striving for – British or American English).

A t-test for dependent samples between the two measurements (early & late) in the dialog confirmed a significant increase for the first condition
(GA native speaker) with \( p < .01 \) \((t = -6.001\) and \(df = 7)\) and only a marginally significant increase for the second condition (SSBE native speaker) with \( p < .05 \) \((t = -2.677\) and \(df = 7)\) across the whole test group. A separate analysis of the values of the less-talented group could not detect any significant difference between the early and late measurements for both conditions \((p > .01\) two-tailed, \(df = 3)\). The repeated measurements \(t\)-test for the talented group though confirmed a significant convergent effect for the first condition \((p < .001, t = -31.235, df = 3)\) but no significant increase for test condition number two (SSBE speaker).

Table 1. Convergence in dialog. The differences of the match values between an early and a late point in the conversations for all eight speakers in the GA and SSBE condition.

<table>
<thead>
<tr>
<th>No.</th>
<th>Talent group</th>
<th>Variety</th>
<th>Gender</th>
<th>Diff_GA</th>
<th>Diff_SSBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>AE</td>
<td>m</td>
<td>0.107</td>
<td>0.257</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>AE</td>
<td>m</td>
<td>0.105</td>
<td>0.127</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>AE</td>
<td>f</td>
<td>0.118</td>
<td>0.051</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>BE</td>
<td>f</td>
<td>0.119</td>
<td>0.088</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>BE</td>
<td>m</td>
<td>0.072</td>
<td>0.032</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>BE</td>
<td>f</td>
<td>0.034</td>
<td>0.003</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>BE</td>
<td>m</td>
<td>0.023</td>
<td>0.041</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>BE</td>
<td>m</td>
<td>0.067</td>
<td>0.025</td>
</tr>
</tbody>
</table>

The Levene test between the two groups talented and less-talented confirmed equality of variances, allowing a \(t\)-test for independent samples. The results of the \(t\)-test are displayed in Table 2. The mean difference in match value across the two conditions (GA and SSBE) is significantly different for the two groups, indicating that the better accommodation of the talented group is definitely above chance level.

Table 2. \(t\)-test for the groups “talented” and “less-talented” across the two test conditions (GA and SSBE).

<table>
<thead>
<tr>
<th>Diff_mean</th>
<th>Talent</th>
<th>n</th>
<th>mean</th>
<th>standard deviation</th>
<th>(t)-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>talented</td>
<td>4</td>
<td>.12133</td>
<td>.042329</td>
<td>(T = 3.762) (p = .009)</td>
</tr>
<tr>
<td></td>
<td>less talented</td>
<td>4</td>
<td>.03695</td>
<td>.014843</td>
<td></td>
</tr>
</tbody>
</table>
There were no significant effects for convergence regarding the preferred English variety of the German subjects and their performance in the two conditions. There were also no significant gender effects in convergence within the dialogs. Figure 2 displays box plots with the variance for the variable gender.

5.2. Control task

The analysis of the control task (read word list containing target words from the dialogs and filler words) before and after the respective dialog revealed no significant difference between the match values ($p > .05$) for neither of the two talent groups. The changes in match values for all eight subjects are displayed in Table 3.

Although the differences in match values are not significant, two of the four less talented speakers seem to move away from their dialog partners’ pronunciation or, in other words, diverge from it (subjects no. 7 and 8). Subject number one, belonging to the talented group (talent score 1.0), also seems to diverge from the pronunciation of the GA speaker after the dialog task. It might be worth noticing that the participant’s preferred and used variety of English is also American English. Figure 3 displays
two diagrams with the mean and the standard deviation of the match value differences for the two variety conditions grouped according to the talent groups for the reading task.

### 6. Discussion

Results for the Diapix-task overall confirmed a significant change in the match values early and late in the dialog, indicating measurable changes in pronunciation of the subjects within a task-oriented dialog in a laboratory setting. The repeated measurements $t$-test confirmed a significant increase in the match values for the talented group, whereas no such significant changes were found for the less talented group. Talented German native speakers in this experiment thus are likely to have converged in their pronunciation to their English native conversational partners.

An interesting pattern is found in the detailed analysis of the two experimental conditions, the dialog with the GA and the SSBE speakers. As a repeated measurements $t$-test for early and late in the dialog showed, the increase in match values for the GA conversation was significant but not for the SSBE conversation. Since the preferred variety of the participants does not explain this tendency, this bias might be caused by the fixed experimental design, in which the GA dialog always came first. By the time the second dialog took place, internal factors as fatigue and a subsequent lack of concentration or difficulty in switching to another variety of English in a relatively short time could have influenced the re-

---

**Table 3.** Table with the match value differences for the word list reading task before and after the respective dialog

<table>
<thead>
<tr>
<th>No.</th>
<th>Talent group</th>
<th>Gender</th>
<th>Diff_GA_read</th>
<th>Diff_SSBE_read</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>m</td>
<td>-.064</td>
<td>.009</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>m</td>
<td>.035</td>
<td>.020</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>f</td>
<td>.053</td>
<td>.001</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>f</td>
<td>.008</td>
<td>.056</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>m</td>
<td>.009</td>
<td>-.075</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>f</td>
<td>.014</td>
<td>.049</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>m</td>
<td>-.027</td>
<td>-.037</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>m</td>
<td>-.059</td>
<td>-.100</td>
</tr>
</tbody>
</table>
Figure 3. Differences in match values before and after the dialog for the two conditions (GA and SSBE)
sults to some extent. Fatigue and lack of concentration are very likely to induce disturbances at the aforementioned stages of attention, recognition and storage, adding further noise to the output later in the task.

As opposed to other studies (e.g., Namy et al. 2002; Pardo 2006) here we could not measure any significant differences in the performance of male and female subjects, neither for the dialog nor for the control task. The before and after measurements for the word list task showed no significant difference, speaking against a longer lasting effect induced by the accommodation during dialog. As described before, the control task was a read speech task, consisting of reading out a list with target and filler words. Monolog and dialog speech differ in many characteristics (see e.g., Pickering & Garrod 2005:9), the most important of which for our experiment would undisputably be the directedness towards a partner. Since this is the main factor necessary for any kind of accommodation, a control task that forces a change into a read speech mode might interfere with the accommodative mode of the speaker and therefore not show any carryover effects. Speaking in exemplar theoretic terms and bearing in mind also the differences in acoustic properties between read and conversational speech, one might think about it as choosing “read speech exemplars” from memory rather than relying on exemplars stemming from a dialog, even a recent one. Alignment thus does not seem to generalize to all speaking styles and remains a strictly conversational phenomenon, not observable within other speaking modes.

Comparing amplitude envelopes in the described form of course has its limitations. These limitations could be responsible for some biases in the obtained data. One important fact is the possible distortion in the match values induced not by different spectral features but by timing and durational effects. Cross-correlating the amplitude envelopes of two items functions ideally only if the envelope curves match both in amplitude and in time. One word being pronounced considerably slower than the other, regardless of their equal amplitude envelope shapes, would result in a lower match value. The present analysis did not include any length normalization of the target words, leaving a chance of speaking rate effects distorting the possibly successful alignment of spectral properties. If this proved to be true after introducing length normalization at segment level, it could be a hint on timing being a property considerably more difficult to accommodate than spectral features.
Notes

1. The Varbrul program has been the dominant approach towards variation in sociolinguistic research of the past two decades. The model that allowed both – to output probability distributions of predicted results and estimate underlying parameters from observed data sets – is based on a Chomskian formalism (Chomsky 1965; Chomsky & Halle 1968) that used transformational rules organized in a modular feed-forward way. Since it relies on categorical variables, predicting dependent from independent variables, Pierrehumbert (2006:519) concludes it being not as well suited for handling gradient data as exemplar theoretical models.

References

Abbot-Smith, K. & Tomasello, M.

Bod, R.

Bradlow, A. R., Baker, R. E., Choi, A., Kim, M. & van Engen, K. J.

Bybee, J.

Bybee, J. & Hopper, P. (Eds.)
2001 *Frequency and the emergence of linguistic structure*. Typological studies in language, 45. Amsterdam: Benjamins.

Chomsky, N.

Chomsky, N. & Halle, M.

Dörnyei, Z.

Giles, H.
Giles, H., Coupland, N. & Coupland, J.

Goldinger, S.

Hawkins, S.

Hogg, M.

Jilka, M., Anufryk, V., Baumotte, H., Lewandowski, N., Rota, G. & Reiterer, S.

Johnson, K.

Johnson, K. & Mullennix, J. W. (Eds.)

Labov, W.

Lacerda, F.
1997 Distributed Memory Representations Generate the Perceptual-Magnet Effect. [manuscript]

Levelt, W. J. M.
Namy, L. L., Nygaard, L. C. & Sauerteig, D.  
2002 Gender differences in vocal accommodation: The role of perception.  

Pardo, J. S.  
2006 On phonetic convergence during conversational interaction.  

Pickering, M. J. & Garrod, S.  
2004 Toward a mechanistic psychology of dialogue.  
*Behavioural and Brain Sciences*, 27, 169–225.

2004 Automaticity of language production in monologue and dialogue.  
Meyer, A., Wheeldon, L. & Krott, A. (Eds.).  
*Automaticity and Control in Language Processing*, pp. 1–20.

Pierrehumbert, J. B.  
2001 Exemplar dynamics: Word frequency, lenition and contrast. In  
J. Bybee & P. Hopper (Eds.),  

2006 The next toolkit.  

Port, R.  
2007 How are words stored in memory? Beyond phones and phonemes.  
*New Ideas in Psychology*, 25(2), 143–170.

Sancier, M. L. & Fowler, C. A.  
1997 Gestural drift in a bilingual speaker of Brazilian Portuguese and English.  

Schütze, H., Walsh, M., Wade, T. & Möbius, B.  
2007 Accounting for phonetic and syntactic phenomena in a multi-level competitive interaction model.  
*ESSLI Workshop on Exemplar Based Models of Language Acquisition and Use*. Dublin.

Shepard, C. A., Giles, H. & Le Poire, B. A.  
2001 Communication Accommodation Theory. W. P. Robinson, & H. Giles  
(Eds.),  

Skehan, P.  
1998  

Wade, T., Dogil, G., Schütze, H., Walsh, M. & Möbius, B.  
(to appear) Syllable frequency effects in a context-sensitive segment production model.

Walsh, M., Schütze, H., Möbius, B. & Schweitzer, A.  
2007 An exemplar-theoretic account of syllable frequency effects.  
*Proceedings of the 16th International Conference of Phonetic Sciences*. Saarbrücken.

Wedel, A. B.  
2006 Exemplar models, evolution and language change.  
Wedel, A. & Van Volkinburg, H.  

to appear  Modeling simultaneous convergence and divergence of linguistic features between differently-identifying groups in contact.  

Zuengler, J.  

Segmental factors in language proficiency: Coarticulatory resistance as a signature of pronunciation talent

Henrike Baumotte

1. Introduction

This chapter describes an acoustic study of the production of English /l/ by speakers of Standard German. In particular, we investigated velarization degree and coarticulatory resistance for /l/ in German learners of English. Previous research (Jilka, this volume) categorized these speakers as either non-proficient, average or proficient with regard to their phonetic abilities. Coarticulation differences might be amongst the reasons for less proficient speakers to be unable to overcome their foreign accent. Ashby and Maidment (2005:132) define coarticulation as the influence of vocal tract configuration at any point in time on more than one segment. In running speech, articulatory gestures overlap in time, leading, consequently, to an interaction of successive phonetic segments. In turn, phonological processes vary in different languages. Farnetani (1997:376) assumes that coarticulation is a universal phenomenon because it has been found in all languages analyzed, e.g. in French (Benguere, Hirose, Sawashima & Ushijima 1977a, 1977b), in English (Lehiste 1964; Bladon & Al-Bamerni 1976; Bladon & Nolan 1977; Majewski, Rothman & Hollien 1977), in Catalan (Recasens 1984a, 1984b; Recasens, Fontdevila & Pallarès 1995; Recasens & Pallarès 2001), in German (Recasens, Fontdevila & Pallarès 1995) and in Polish (Majewski, Rothman & Hollien 1977). Coarticulation is one of the sources of variation in phonological processes and differs both according to language (Öhman 1966; Manuel 1990; Recasens, Fontdevila & Pallarès 1995) and to speaker, resulting in allophonic variation and individual speech variants (Baumotte et al. 2007). Figure 1 represents the occurrences of coarticulation accounting for the appearance of foreign accent.

Coarticulation can be examined with the help of different experimental techniques: acoustic analysis such as spectrograms, electropalatography (EPG), imaging techniques such as x-ray studies (Ashby & Maidment 2005:129), electromagnetic articulography, electromyography
(EMG), transducers for investigating velopharyngeal function, as well as techniques for investigating laryngeal articulation, investigations of devoicing gesture and techniques for analyzing the voice source (for further information see Hardcastle & Hewlett 1999:229ff). Acoustic analysis has been shown to be especially reliable in the investigation of changes in acoustic characteristics. In this paper we use spectrograms to determine vowel formant frequencies for comparison of coarticulation/coarticulatory resistance patterns in proficient vs. average vs. non-proficient learners of L2 English.

This work is situated within the theoretical framework of Coarticulatory Resistance (CR) which was first employed by Bladon and Al-Bamerni (1976). The authors conducted a spectrographic investigation of the production of the extrinsic allophones of /l/ in Received Pronunciation (RP), analyzing coarticulatory resistance, while also including the dichotomy of the to some extent velarized (“dark”, here symbolized as [l]) or non-velarized (“clear” [l], with some degree of palatalization) types of the consonant. If velarization takes place, the tongue body and

![Figure 1. The different aspects of coarticulation/coarticulatory resistance](image_url)
root move from their neutral position in the direction of the vowels [u] and [ɔ], towards the soft palate (Clark et al. 2007:64, 96). Normally, the clear variant appears before a vowel (e.g. lend, alight, believe) in syllable-initial and/or phrase-initial position, and the dark one before a consonant or syllable- or word-final (e.g. wild, halt, will, hall) (Clark et al. 2007:96). In some varieties of English the degree of velarization is more extreme than in others; for example, Londoners and South Australians velarize very strongly. Recasens (1985) found for Catalan that syllable-final or word-final [l] is produced with more velarization than its non-velarized, syllable-initial or word-initial counterpart [l], as is the case in some American English dialects (Kenyon 1950). But this kind of assimilation, i.e. variation in the pronunciation of /l/, cannot automatically be transferred to every other language. It is not found in many of the world’s languages, e.g. German kalt (“cold”) and Italian caldo (“warm”) are produced with the clear variety of the consonant [l] (Clark et al. 2007:96). Bladon and Al-Bamerni (1976) found that the influence of adjacent vowels on F2 and of adjacent voiceless plosives on the degree of voicelessness in the British English laterals decrease from clear [l] to dark [l] to dark syllabic [l]. Their experimental results agree with the “articulatory syllable” theory of R-L coarticulation. Bladon and Al-Bamerni (1976:146) synchronously state their evidence of L-R coarticulation as also appearing constricted by the same notion of an articulatory syllable. In this case their data do not particularly confirm the existence of different mechanisms for each coarticulatory direction (Bladon & Al-Bamerni 1976:148).

Different views have been proposed concerning the notion of coarticulatory resistance. Bladon and Al-Bamerni (1976) define CR to be the degree of variation or similarity of the same speech sound across contexts (Bladon & Al-Bamerni 1976:137). It is the property associated with phonetic specifications for speech segments that varies according to their magnitude. The term indicates the extent to which a particular segment is permeable for coarticulation. Contrasting vowels and consonants can differ in the extent to which they allow context-dependent effects to occur, and so they can be categorized in terms of stability (Stevens & House 1963:119) or resistance (Bladon & Al-Bamerni 1976:137). According to Farnetani and Recasens (1999:32), it is the degree to which a given segment resists potential interference from the neighbouring segments.

In their Degree of Articulatory Constraint (DAC) model of lingual coarticulation, Recasens et al. (1997) and Recasens (2004) have included
quantification in a theory of coarticulation developing coefficients of CR. On the basis of grouping languages into velarized (“dark”) or non-velarized (“clear”) varieties of the consonant /l/, Recasens, Fontdevila and Pallarès (1995) analyzed velarization-pattern and vowel-to-consonant coarticulation for the German /l/ in the sequences [ili] and [ala] in contrast with Catalan /l/. The consonant /l/ has been categorized as “dark” or velarized in Catalan, as opposed to “clear” and non-velarized in German (Recasens & Farnetani 1990). At the same time the authors looked for articulatory and acoustic attributes of /l/ in German to ensure that they are indeed distinct from those in other languages. Low or high F2 values serve as an indicator for the velarization distinction. The authors compared German [l] with Catalan [l̃] production by collecting acoustic and linguopalatal data (by means of electropalatography [EPG]) and observed greater dorsal contact at the palatal zone for German [l] than for Catalan [l]. Those consonants which require conflicting tongue fronting, such as palatals, manifest a high second formant frequency. In contrast, those which have a pharyngeal constriction, such as [l], show a low F2. The lowest F2 for [a] can be found in combination with [w] and is probably associated with lip rounding (Recasens 1985:109). The tongue dorsum is more constrained for German non-velarized [l], and thus less sensitive to coarticulatory effects from, for example, [i] or [a]. Because it is produced with antagonistic tongue dorsum gestures, i.e. tongue dorsum lowering and retraction vs. raising and fronting, velarized Catalan [l̃] is highly resistant to coarticulation with [i] (Recasens 1985:109; Recasens, Fontdevila & Pallarès 1995:38). In line with the surrounding formant frequencies for the vowel [a], consonantal effects on F2 for [ə] are large because no defined vocal-tract shape is necessary for the production of [ə]; this is why schwa is highly sensitive to coarticulation (Recasens 1985). Recasens (1984a), as well as Recasens, Fontdevila and Pallarès (1995:38) suggest the existence of different mechanisms of articulatory control. Farnetani (1997:388) underlines this statement by arguing that the differences in coarticulation across consonants stem from consonant specific production constraints imposed on the tongue body. Data based on V-to-C coarticulation on the tongue body suggests no complete coarticulation of alveolar consonants with adjacent vowels, which hints at the presence of functional and/or physical coupling between tip/blade and body. It is possible that coarticulatory resistance at the acoustic level does not result exclusively from resistance in lingual activity but also from labial or mandibular activity (Recasens, Fontdevila & Pallarès 1995:50).
Bladon and Al-Bamerni (1976) for RP, Lehiste (1964) for American English, as well as Recasens (1984a, b) for Catalan, describe consonants extending beyond larger regions of the vocal tract due to tongue-dorsum activity (e.g. velarized apicoalveolar, bilabiodorsalveolar, dorsal consonants [alveolo-palatals, palatals]), which are highly resistant to vowel coarticulation. During the articulation of alveolars the tongue dorsum is not directly involved and is therefore free to coarticulate with the neighbouring vowels. In the case of palatals the tongue dorsum makes the vocal tract narrower, which inhibits possible coarticulatory gestures with the surrounding vowels (Recasens 1984a:62). Recasens (1985:103) analyzed Catalan CV sequences, discovering that dentals and alveolars (including [r] [l]) permit less F₂ variability than velars and labials. In comparison to [t d s z], [l] and [r] are produced with more backing of the tongue body, leading to a lower value for the second formant (Recasens 1985:104). Similar to other languages, small V-to-C coarticulatory effects for the Catalan labiovelar [w] may be connected with the necessity to form a bilabial and a dorsovelar constriction at the same time (Recasens 1985:104).

Dentals and alveolars (including [l] and [r]) also permit a small degree of F₃ variability at vowel onset, which is consistent with a highly invariant place of articulation and a considerable resistance to conflicting jaw-lowering effects. The small degree of F₃ coarticulation for [p b f] and [w] at vowel onset makes it clear that the degree of bilabial closure for labials and of labiovelar constriction for labiovelars depends on highly constant conditions across vowels (Recasens 1985:105).

“[…] coarticulatory resistance has been found to be positively correlated with large degrees of dorsopalatal contact, the formation of a double place of articulation, and tongue-body-tongue-tip coupling.” (Recasens 1985:105)

Recasens (1985:105) concluded that consonants can be differentiated according to the extent of articulation concerning one or more of these constraints, and thus, to the degree of V-to-C coarticulation.

In this study we assume that a high ability to break and loosen the language specific coarticulation/coarticulatory resistance leads to a smaller degree of foreign accent. Our experiments investigate the degree of coarticulation/coarticulatory resistance in English /ɔlV/-sequences in proficient vs. average vs. non-proficient German speakers to test if talented learners of L2 English develop strategies of breaking and loosening their language-specific coarticulatory constraints.
2. Experiment

2.1. Analysis 1: Degree of velarization

2.1.1. Methods

2.1.1.1. Subjects

29 (11 non-proficient [7 f, 4 m], 11 proficient [8 f, 3 m], 7 average [5 f, 2 m]) native German speakers (20 female, 9 male) have been recruited for the analysis. They were 21 to 41 years old with a mean age of 25.45 y. Most of them grew up in the Swabian region, i.e. south of Germany having in general a less dialectal manner of speaking because of exposure to Standard German speakers and having a high level of education. Bavarian and Swabian dialect speakers are known to produce a clear variety of the consonant /l/. All of them had an academic background. Subjects had beforehand taken part in extensive tests of phonetic language ability, which assessed their pronunciation accuracy in English (Jilka, this volume).

2.1.1.2. Materials

Formant frequency data were collected for the logotoms “gelate, gelite, gelüte, gelute”¹ spoken as close as possible to the Standard Southern British English. The target non-words were embedded in carrier sentences (I have said [gelate/gelite/gelüte/gelute] twice.) with stress on the second syllable. This speech material was read five times by each speaker resulting in 575 tokens (1 consonant /l/ × 4 vowel contexts × 5 repetitions × 29 speakers = 580). Five tokens were dismissed due to imprecise articulation and hesitation errors.

2.1.1.3. Procedure

At the beginning of each session, subjects were instructed to repeat a small text presented by a female university lecturer (56 y) speaking with a British English accent, to help speakers switch into the target language, and particularly to prime them in the foreign language.

Digital recordings were made at a 16 kHz sampling rate in a sound-proof recording room in the phonetics laboratory of the Institute for

---

¹. To control for pitch, duration and phonemes a non-word has been chosen.
Natural Language Processing, Universität Stuttgart, Germany. The data were then segmented at the phone level by automatic forced alignment (Rapp 1995) and checked manually afterwards. Formant frequencies were measured every 10 ms with the \textit{ESP5 formant} program. $F_1$, $F_2$ and $F_3$ were extracted from the middle of the steady state in /ə/ following a standard procedure (Recasens 1999:327).

According to Ziegler and von Cramon (1986:35) the formant frequencies in /ə/ are the most appropriate parameters to disentangle vowel-to-vowel coarticulation. $F_2$ frequency and the frequency distance between $F_2$ and $F_1$, $F_v$, served as indicators for the degree of consonantal velarization.

In order to determine significant formant-dependent differences, one-way ANOVAs were submitted separately for $F_2$ and $F_v$, with proficiency level as the independent variable. Further, the interactions between the degree of velarization and each proficiency level were calculated by means of a correlation analysis (Scheffé-tests) between the formant frequency values in /ə/. The purpose of this analysis was to investigate whether an increase for a given proficiency level was matched with an increase in degree of velarization, and less importantly, if velarization and proficiency level were strongly correlated.

Based on previous studies we predicted that velarized native-like English /l/ should show weaker coarticulatory effects on /ə/ before /l[aɪ, eɪ, ɪ, ʊ\textsuperscript{\text{lengthmark}}]/ than non-velarized less proficient German-accented English /l/ (no active tongue dorsum gesture). Following these assumptions, we predicted $F_2$ and $F_v$ in low proficient non-native speakers to be higher than in high proficient speakers. Average speaker values should lie between those of proficient and non-proficient speakers.

\footnote{According to Fant (1960) $F_2$ correlates positively with tongue dorsum fronting and raising and is thus inversely related to the degree of velarization. $F_v$ considers also the contribution of $F_1$ known to be inversely related to velarization (Recasens, Fontdevila & Pallarès 1995:41). $F_2$ and $F_v$ decrease, if degree of velarization rises.}
2.1.2. Results

Table 1. Mean $F_2$ and $F_v$ values (in Hz) for /ə/ before /a/, /e/, /y/, /u/ across /l/ for all less proficient vs. proficient speakers. Standard deviations have been included in parenthesis.

<table>
<thead>
<tr>
<th>Across all speakers</th>
<th>$F_2$</th>
<th>$F_v$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Less proficient</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_2$</td>
<td>1967.79</td>
<td>1578.23</td>
</tr>
<tr>
<td>(282.987)</td>
<td>(324.351)</td>
<td>(298.484)</td>
</tr>
<tr>
<td><strong>proficient</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_2$</td>
<td>1868.10</td>
<td>1467.13</td>
</tr>
<tr>
<td>(324.351)</td>
<td>(338.731)</td>
<td></td>
</tr>
</tbody>
</table>

Mean $F_2$ and $F_v$ values for /ə/ across /l/ are shown in Table 1. The vowel exhibits a higher $F_2$ for less proficient speakers (mean value: 1967.79 Hz) than for proficient speakers (mean value: 1868.10 Hz) with a difference between means of 99.69 Hz. The $F_v$ value is also higher for less proficient speakers (mean value: 1578.23 Hz) than for proficient (mean value: 1467.13 Hz), with a difference between means of 111.1 Hz. $F_2$ (mean value: 1951.68 Hz) and $F_v$ (mean value: 1575.69 Hz) mean values for average speakers fell between those of proficient and non-proficient speakers.

One-way ANOVAs across proficiency levels revealed significant differences in $F_2$ ($F(2)=7.408, p < .001$), as well as $F_v$ ($F(2)=9.052, p < .000$) which could be displayed by Scheffé-tests. $F_2$ and $F_v$ allow a high percentage of significant degree of velarization effects between average/non-proficient vs. proficient speakers, while non-proficient vs. average groups exhibited no significant results.

Table 2. One-way ANOVA (Scheffé) test results ($df_{bg} = 2$, $df_{wg} = 575$, $p < .05$) after comparison of $F_2$ and $F_v$ in non-proficient (1) vs. average (2) vs. proficient non-native (3) English speech

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Groups</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_2$</td>
<td>1 vs. 3</td>
<td>$p &lt; .002$ *</td>
</tr>
<tr>
<td></td>
<td>2 vs. 3</td>
<td>$p &lt; .026$ *</td>
</tr>
<tr>
<td></td>
<td>1 vs. 2</td>
<td>$p &lt; .880$</td>
</tr>
<tr>
<td>$F_v$</td>
<td>1 vs. 3</td>
<td>$p &lt; .001$ *</td>
</tr>
<tr>
<td></td>
<td>2 vs. 3</td>
<td>$p &lt; .004$ *</td>
</tr>
<tr>
<td></td>
<td>1 vs. 2</td>
<td>$p &lt; .997$</td>
</tr>
</tbody>
</table>
Figure 2. Distribution of $F_2$ (left) as well as $F_v$ (right) in less proficient vs. average vs. proficient speakers.
2.1.3. Discussion

Significant $F_2$ and $F_2-F_1$ differences across /l/ are in conformity with our hypothesis and indicate that the consonant is more velarized in proficient L2 English learners, and less velarized for less proficient speakers. The data reported in Analysis 1 brings up results as predicted, we can assume the possibility that consonantal realization, i.e. the degree of velarization, depends on the level of proficiency because of significant Scheffé-test-results for non-proficient vs. proficient speakers, as well as for proficient vs. average speakers. Mean vowel $F_2$ and $F_v$ values in /ə/ for only female speakers and for each individual female subject are given in Tables 3 and 4.

The effect obtained with degree of velarization might be proficiency-dependent and could also almost hold for coarticulatory resistance. Its effects will be discussed in a second analysis.

**Individual differences in degree of velarization for less proficient vs. proficient female speakers**

A separate analysis of female subjects’ data should give a closer look at individual differences and account for presumable sex-related differences in vocal tract size.

*Table 3.* Mean $F_2$ and $F_v$ values (in Hz) for /ə/ before /a/, /e/, /y/, /u/ across /l/ in less proficient vs. proficient speakers (female subjects only). Standard deviations have been included in parenthesis.

<table>
<thead>
<tr>
<th></th>
<th>$F_2$</th>
<th>$F_v$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>less proficient</td>
<td>proficient</td>
</tr>
<tr>
<td>$F_2$</td>
<td>2027.77</td>
<td>1908.83</td>
</tr>
<tr>
<td></td>
<td>(294.049)</td>
<td>(356.115)</td>
</tr>
</tbody>
</table>

In line with the results of Analysis I, less proficient female speakers exhibit a higher $F_2$ (2027.77 Hz) and $F_v$ (1614.37 Hz) than female proficient speakers with differences in means of 119.44 Hz for $F_2$ and 126.86 Hz for $F_v$. $F_2$ and $F_v$ formant frequency values for proficient female speakers were clearly smaller than those for non-proficient speakers.

One-way ANOVAs (Scheffé) across female speakers reached significant $F_2$ ($F(2)=5.395, p < .005$) and $F_v$ ($F(2)=5.510, p < .004$) differences
between proficient and non-proficient speakers ($F_2: p < .005$, $F_v: p < .005$) ($df_{bg} = 2$, $df_{wg} = 396$). Male speakers also differ in $F_2$ ($F(2)=4.931$, $p < .008$) ($df_{bg} = 2$, $df_{wg} = 179$) and $F_v$ ($F(2)=7.131$, $p < .001$) ($df_{bg} = 2$, $df_{wg} = 175$), mean values for proficient speakers being always lower than those for non-proficient speakers (mean values $F_2$: proficient: 1787.16 Hz (230.943 Hz), non-proficient: 1855.78 Hz (228.003 Hz); mean values $F_v$: proficient: 1431.18 Hz (268.851 Hz), non-proficient: 1493.90 Hz (217.933 Hz)), but they could not be differentiated significantly in those groups ($p < .148$).

The $r$ (Pearson correlation coefficient) value for $F_2$ and $F_v$ in /ə/ is .9837 (adjusted) (see Figure 3).

![Figure 3. $F_2$ and $F_v$ plot for regression line including less proficient (squares) and proficient female speakers (points).](image-url)
Table 4. Mean $F_2$ and $F_v$ values (in Hz) for /ə/ before /æ/, /e/, /y/, /u/ across /l/ in less proficient vs. proficient speakers (female subjects only) for individual speakers. Standard deviations have been included in parenthesis.

<table>
<thead>
<tr>
<th>Individual speakers</th>
<th>$F_2$</th>
<th></th>
<th></th>
<th>$F_v$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Less proficient</td>
<td>proficient</td>
<td></td>
<td></td>
<td>Less proficient</td>
<td>proficient</td>
</tr>
<tr>
<td>DB 1734.00 (288.435)</td>
<td>AJ 2141.48 (115.671)</td>
<td></td>
<td></td>
<td>DB 1313.35 (307.099)</td>
<td>AJ 1705.40 (126.683)</td>
<td></td>
</tr>
<tr>
<td>JM 2074.15 (85.248)</td>
<td>AT 2316.58 (237.872)</td>
<td></td>
<td></td>
<td>JM 1662.45 (78.587)</td>
<td>AT 1952.68 (221.483)</td>
<td></td>
</tr>
<tr>
<td>KM 1984.85 (328.7)</td>
<td>BR 1905.64 (310.431)</td>
<td></td>
<td></td>
<td>KM 1575.05 (368.771)</td>
<td>BR 1554.75 (323.182)</td>
<td></td>
</tr>
<tr>
<td>SW 2206.65 (317.140)</td>
<td>BS 1912.88 (165.112)</td>
<td></td>
<td></td>
<td>SW 1733.90 (386.490)</td>
<td>BS 1485.85 (164.003)</td>
<td></td>
</tr>
<tr>
<td>TS 2051.00 (54.807)</td>
<td>CB 1844.35 (104.5)</td>
<td></td>
<td></td>
<td>TS 1650.50 (63.298)</td>
<td>CB 1389.05 (109.042)</td>
<td></td>
</tr>
<tr>
<td>UB 2355.40 (61.344)</td>
<td>CS 1893.35 (266.445)</td>
<td></td>
<td></td>
<td>UB 1989.45 (71.051)</td>
<td>CS 1465.50 (243.471)</td>
<td></td>
</tr>
<tr>
<td>JN 1788.35 (155.178)</td>
<td>KK 1972.65 (174.662)</td>
<td></td>
<td></td>
<td>JN 1375.90 (186.685)</td>
<td>KK 1523.45 (233.035)</td>
<td></td>
</tr>
<tr>
<td>KN 1254.2 (298.914)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A test for the homogeneity of variance gave non-significant differences between the two proficiency levels, thus refuting that the two speaker populations are comparable. Figure 3 shows that there are no clear clusters for proficient vs. non-proficient speakers. Proficient speakers do not consistently produce lower vowels than less proficient speakers and vice versa, because vowel formant frequencies vary a lot through proficiency levels. $F_2$ and $F_v$ vowel data for these individual female speakers were submitted to cluster analysis (using complete linkage or the “furthest neighbour” clustering method) to investigate whether they were classified according to proficiency level or not, which obtained the same indistinctive cluster results as can be seen in the figure above. Identical conditions hold for the male speakers.

A more detailed look was taken at the most proficient speaker, AT, and the least proficient speaker, DB (see Table 5). Figure 4 shows spectrograms produced by these speakers. In contrast to our initial hypothesis,
AT realized a non-velarized type of the consonant [l] leading to a high F₂ value, while there is a lot of velarization in [l] for DB. AT was categorized as RP dialectal speaker and therefore might be conscious of the fact that in RP the clear variant appears before a vowel in syllable-initial and/or phrase-initial position, and the dark one before a consonant or syllable- or word-final. DB might know about the overall rule claiming that /l/ is generally more velarized in English than in German which made her try to come close to those varieties.

2.2. Analysis 2: Coarticulatory resistance in VCV clusters

2.2.1. Methods

2.2.1.1. Subjects and Materials and Procedure

Subjects, materials and procedure were identical to the ones described in Analysis I (see p. 6ff).

Following Recasens, Fontdevila and Pallarès (1995:47), we predicted that coarticulatory sensitivity is inversely related to the degree of consonantal velarization, i.e. /l/ should be more context-dependent in German low proficient speakers of English than in high proficient learners. According to Recasens and Farnetani (1990, 1991), this general hypothesis is valid for languages with very different varieties of the consonant /l/, e.g. Italian and Catalan. In order to test the hypothesis, we analyzed coarticulatory resistance of /εu/ vs. /u/, /y/ vs. /ε/, as well as /y/ vs. /a/ on /l/ in symmetrical VCV-sequences (/[dʒəlæt] / [dʒəlæt] / [dʒəlt] / [dʒəlu:]) produced by non-native speakers of English. The vowel sequence /εu/ (lips are spread [Gimson 1980:129]) vs. /u/ (lips are rounded [Gimson 1980:121]) has been chosen since it presents two maxima along the articulatory dimension concerning roundedness/unroundedness, as well as backness/fronting. The vowel /y/ (lips are rounded [Pompino-Marschall 1995:212]) and the diphthong /εt/ (lips are spread [Gimson 1980:129]) are both front vowels produced with a constricted pharynx differing in roundedness. /y/ (lips are rounded [Pompino-Marschall 1995:212]) vs. /a/ (lips change from a neutral position to a loosely spread one [Gimson 1980:131]) integrate a loose difference in roundedness/unroundedness, as well as a slight distinction in jaw opening/closing. Within the production of the RP diphthongal glide “/a/” the lower jaw closes, coming closer to the position of RP associated /u/ (Gimson 1980:131), but usually not reaching a tongue level closer than C [ɛ]. The front-close rounded vowel [y] (Pompino-
Marschall 1995:212, 254) does not exist in any variety of English, neither in American nor in British (see IPA transcriptions [German: Kohler 1999; Mangold 2005; British English: Gimson 1980]). The vowel [y] has been included because we are interested in examining if it will continue to be articulated in the German specific manner bearing in mind that the English phonetic alphabet does not contain this vowel.

We expect proficient speakers to velarize more, therefore their speech should be more coarticulatory resistant, i.e. /ə/ should emerge closer to its formant frequencies as a neutral vowel (F₁ = 560 Hz, F₂ = 1480 Hz, F₃ = 2520 Hz [Gimson 1980:101]). In contrast, less proficient speakers should coarticulate more because of a lesser degree of velarization while producing the non-velarized allophonic variant [l].

Consequently, coarticulatory distance was calculated using frequency measurements made from spectrograms in Hz by means of the following formula:

\[ e.g. \ CD = F₂\overline{e}l\overline{e} – F₂\overline{e}lu; \]

We modified the MCD (mean coarticulatory distance)-calculation proposed by Bladon and Al-Bamerni (1976:142). The so-called degree of coarticulation or coarticulatory distance is an appropriate measurement to distinguish the deviation from the canonical vowel (see below) indicated by the decrease or increase of F₂ and F₂' in the neutral vowel /ə/.

F₂- and F₂'-values should be positive in the case of /əl/ (/əl/: F₂ = 2060 Hz, F₃ = 2840 Hz, F₂' = 2450 Hz; /l/: F₂ = 2220 Hz, F₃ = 2960 Hz, F₂' = 2590 Hz [Gimson 1980:101]) vs. /u/: (F₂ = 920 Hz, F₃ = 2200 Hz, F₂' = 1560 Hz [Gimson 1980:101]), as well as /y/ (F₂ = 1750 Hz [Delattre 1981:73]) vs. /au/ (/au/: F₂ = 1320 Hz, F₃ = 2500 Hz, F₂' = 1910 Hz; /l/: F₂ = 2220 Hz, F₃ = 2960 Hz, F₂' = 2590 Hz [Gimson 1980:101, 131]), but not within the /y/ (F₂ = 1750 Hz [Delattre 1981:73]) vs. /æl/ (/æl/: F₂ = 2060 Hz, F₃ = 2840 Hz, F₂' = 2450 Hz; /l/: F₂ = 2220 Hz, F₃ = 2960 Hz, F₂' = 2590 Hz [Gimson 1980:101])-comparison.

In our study, high positive/negative values indicate coarticulation, while values around 0 Hz indicate coarticulatory resistance. In general, the closer the frequency value comes to 0 Hz, the more coarticulatory resistance is present in the given stimulus.

3. According to Chistovich and Lublinskaya (1979) and the “center of gravity” hypothesis, F₂' ([F₂ + F₃]/2) was taken to make it similar to a single-formant stimulus because of F₂ and F₃ are closely spaced.
In order to measure whether the degree of coarticulatory resistance effects was significant, one-way ANOVAs were submitted separately for F₂ and F₂' in /əleɪ/ vs. /əlu:/, /əlʏ/ vs. /əleɪ/ and /əlʏ/ vs. /əlɑːl/, with proficiency level as the independent variable. Further, the interactions between degree of coarticulation/coarticulatory resistance and each proficiency level were calculated by means of a correlation analysis (Scheffé-tests) between the formant frequency value differences in /ə/.

The purpose of this analysis was to investigate whether an increase for a given proficiency level was matched with an increase in the degree of coarticulatory resistance; and less importantly, if the degree of coarticulatory resistance and the proficiency level were strongly correlated.

2.2.2. Results

Significant interactions between F₂ in /əleɪ/ vs. /əlu:/ (F(2)=3.059, p < .050 [dfbg = 2, dfwg = 142]) could be shown, whereas no significant correlations could be detected between F₂ and the proficiency level in /əlʏ/ vs. /əleɪ/ (F(2)=1.657, p < .194 [dfbg = 2, dfwg = 141]) and /əlʏ/ vs. /əlɑːl/ (F(2)=1.644, p < .197 [dfbg = 2, dfwg = 140]) and F₂' in /əleɪ/ vs. /əlu:/ (F(2)=0.337, p < .714 [dfbg = 2, dfwg = 142]), /əlʏ/ vs. /əleɪ/ (F(2)=0.848, p < .431 [dfbg = 2, dfwg = 141]) and /əlʏ/ vs. /əlɑːl/ (F(2)=0.795, p < .454 [dfbg = 2, dfwg = 140]) and proficiency level.

Table 5. One-way ANOVA (Scheffé) test results after comparison CD of F₂ in /əleɪ/ vs. /əlu:/ in non-proficient (1) vs. average (2) vs. proficient non-native (3) English speech.

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Parameter</th>
<th>Groups</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>/əleɪ/ vs. /əlu:/</td>
<td>F₂</td>
<td>1 vs. 3</td>
<td>p &lt; .828</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 vs. 3</td>
<td>p &lt; .054 *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 vs. 2</td>
<td>p &lt; .195</td>
</tr>
</tbody>
</table>

(See Figure Baumotte 5 in the Colour figure section)

Front close spread /əleɪ/ vs. back close rounded /əlu:/

For /eɪ/ vs. /u:/, i.e. the backness/fronting-, as well as the roundedness/unroundedness-distinction, the degree of velarization does not show a tendency to be related to proficiency, thus rejecting our initial hypothesis.
While less proficient speech is more coarticulatory resistant, more proficient English speech is less resistant in $F_2$ and $F'_2$ to coarticulatory effects from /eɪ/ vs. /u/. Taking into consideration $F_2$, reduced coarticulation can be observed for less proficient speakers ($mean$ $value$: $-72.72$ $Hz$) than for proficient ($mean$ $value$: $-110.22$ $Hz$). For $F'_2$, the mean $Hz$-value is lower ($mean$ $value$: $-26.817$ $Hz$) for proficient learners than for less proficient ($mean$ $value$: $-6.680$ $Hz$), but non-proficient vs. proficient comparisons turned out to be insignificant. We predicted the values for $F_2$ and $F'_2$ in proficient, average and non-proficient subjects to be positive, however they were negative in the cases of $F_2$ and $F'_2$ for non-proficient learners as well as proficient (see above). In contrast, average speakers reach positive CR values in both comparisons ($mean$ $value$ $F_2$: $55.03$ $Hz$, $mean$ $value$ $F'_2$: $7.414$ $Hz$), with average vs. proficient speakers differing significantly in $F_2$ comparisons.

Front close rounded /əly/ vs. front close spread /əleɪ/

The /əly:/ vs. /əleɪ/-comparison in $F_2$ for proficient vs. non-proficient speakers is expected to tend towards the predicted direction concerning coarticulation/coarticulatory resistance-values, but $F'_2$ values are reversed, a mirror image. Overall results are, contrary to the hypothesis, positive in all cases. No significant effect was found for any of the $F_2$ comparisons in non-proficient ($mean$ $value$: $126.54$ $Hz$) vs. proficient learners ($mean$ $value$: $113.75$ $Hz$) vs. average learners ($mean$ $value$: $23.82$ $Hz$); and in average ($mean$ $value$: $23.82$ $Hz$) vs. non-proficient ($mean$ $value$: $126.54$ $Hz$) vs. proficient learners ($mean$ $value$: $113.75$ $Hz$), even if the average speaker’s $F_2$ mean value is quite low in contrast to that of non-proficient/proficient speakers. In $F'_2$, comparison blocks for less proficient speakers are lower, closer to $0$ $Hz$ than those of proficient learners, while the average speakers’ mean values can be found between those of proficient and non-proficient speakers.

Front close rounded /əly/ vs. front half-close loosely spread /əlaɪ/

No significant effects occur for all of these comparisons. In the /əly/ – /əlaɪ/-comparison, average speakers’ $F_2$ is very low ($mean$ $value$: $1.79$ $Hz$), while less proficient ($mean$ $value$: $113.36$ $Hz$) and proficient ($mean$ $value$: $104.66$ $Hz$) speakers coarticulate a lot. Less proficient speakers coarticulate more than proficient speakers. $F'_2$ mean values (non-proficient: $5.260$ $Hz$, average: $7.132$ $Hz$, proficient: $53.644$ $Hz$) increase with profi-
ciency level, while non-proficient and average learners are quite resistant concerning the vowel formant frequency influence of /y/ and /aɪ/ on /ə/ across /l/.

2.2.3. **Discussion**

*Front close spread /ælɛl/* vs. *back close rounded /əluː/*

Contrary to our hypothesis, there is more resistance in F₂ and F₂’ for less proficient speakers than for proficient ones. At the same time, coarticulatory resistance values for F₂ and F₂’ in less proficient and proficient speakers go in the negative direction, in contrast to those for average learners’ values which are positive and therefore higher than all of the other values. If a constriction in the frontal part of the tract takes place, frequency values of F₂ increase. The larger the degree of constriction, the higher the F₂ frequency value. This is why F₂ reaches its maximum for front vowels like /i/. Constriction in the back of the tract leads to a decrease in F₂, dependent on the degree of constriction. As was expected, back vowels like /uː/ show low second formant frequencies. Proficient and non-proficient speakers might not have pronounced an /eɪ/ in most of the cases, but instead a vowel with less constriction in the frontal part of the tract. The influence of F₂ on /ə/ before /lɛl/ is higher because of a considerably high degree of coarticulation. Obviously, average learners realized an [eɪ] with constriction in the frontal part of the tract, but, on the other hand, low F₂ values for [uː], which resulted in rather high positive values.

F₂’ (mean value of F₂ and F₃) values include not exclusively the constriction of the tract, but also integrate lip rounding. Lip rounding leads to a constriction of the tract, and at the same time includes lip protrusion which is commonly more extended in back rounded vowels than in front rounded vowels. The more lip rounding there is, the lower the F₂ and F₃ formant frequencies. In contrast to average speakers, proficient and less proficient speakers might have produced /uː/ with a low degree of lip rounding (high F₂ and F₃ values). It can be concluded while proficient speakers velarize more compared to average and less proficient speakers, less proficient speech is more resistant to coarticulatory effects of F₂ and F₂’ on /ə/ from /eɪ/ vs. /uː/.
**Front close rounded /ɛly/ vs. front close spread /əlet/**

In the case of the front close rounded/unrounded vowel-distinction for F2, there is a bit less coarticulation from /y/ vs. /εt/ for proficient than for less proficient learners, but the differences are insignificant. In total non-proficient and proficient speakers coarticulate a lot, while average speakers are quite resistant to the influence of the formant frequencies of the vowel following /l/ on /ə./

Lowered F2 values for /εt/ might have led to positive F2 values taking into account the values for /y/ (F2 = 1750 Hz [Delattre 1981:73]) vs. /εt/ (/ɛ: F2 = 2060 Hz, F3 = 2840 Hz, F2’ = 2450 Hz; /ɛl/; F2 = 2220 Hz, F3 = 2960 Hz, F2’ = 2590 Hz [Gimson 1980:101]). F2’-results refute the hypothesis and support the prediction of proficient speakers being resistant to let through the roundedness characteristic than average and less proficient speakers, because of an increase in F2’ height depending on proficiency level. The influence of lip rounding in /y/ on /ə/ might not be high, or the vowel might be articulated in a lesser rounded manner (low F2 and F3), because F2’ goes in the positive direction as well.

Contrary to our hypothesis, F2 differences in coarticulation/coarticulatory resistance of proficient vs. non-proficient speakers on /ə/ across /l/ (/ɛly minus /əla/) are insignificant, while at the same time less proficient speakers coarticulate quite the same as compared to proficient speakers. The average speakers’ CR value is resistant in F2. In the case of the front close/front half-close rounded/unrounded vowel-comparison for F2’ values are in opposition to the hypothesis and show an increase depending on the level of proficiency. Calculation of coarticulatory resistance resulted in positive values in all six cases (/y/: F2 = 1750 Hz [Delattre 1981:73]; /a/: /a/: F2 = 2220 Hz, F3 = 2960 Hz, F2’ = 2590 Hz; /ɛl/: F2 = 2220 Hz, F3 = 2960 Hz, F2’ = 2590 Hz [Gimson 1980:101, 131]). Roundedness might not influence the frequency value of F3 in /y/ to a noticeable extent, if yes, coarticulatory resistance values for F2’ should be negative due to roundedness lowering the frequencies of F3. In the beginning of /a/-articulation, the jaws are considerably separated while the lips are opened neutrally (similar to the articulation of RP type of /ʌ/ [Gimson 1980:131]). The tongue is elevated just above the fully open position with no contact to the upper molars. During the production of /l/ the lower jaw closes continuously. In contrast, /y/ is characterized
through roundedness and a constricted pharynx. In proficient learners, coarticulation of /y/ vs. /ai/ across /l/ is quite irresistant, however, less proficient speakers show a large degree of resistance in F₂’ (active tongue dorsum gesture).

To summarize, significant correlation for F₂ in /əleɪ/–/əlu:/ in proficient vs. average speakers are in opposition to the degree of velarization results of Analysis 1. And the results for F₂’ in /əleɪ/–/əlu:, /əly/–/əleɪ/, as well as /əly/–/əlaɪ/, cannot be clearly attributed to the same mechanism of articulatory coordination pointed out for the degree of velarization results in Analysis 1, in which proficient speakers velarize more than average as compared with less proficient speakers.

2.3. General discussion

Oh (2008:381) investigated coarticulation differences in native vs. non-native French and also English speakers, and concluded that more experienced learners developed more native-like degrees of coarticulation than less experienced learners. In accordance with Oh’s results, these data might suggest that proficient speakers more adequately and better acquire the fine-grained language-specific patterns of coarticulation. Probably, subjects categorized as less proficient or average might not automatically be able to enlarge highly their stored phonetic features after having heard a sound which is not similar to those existing in their mother tongue, even if they perceive a certain difference to mother tongue frequencies. As a result, during L2 production fewer exemplars can be activated than it is true for proficient speakers (Pierrehumbert 2001). The use of tongue dorsum control, tongue dorsum fronting and raising might correlate with perception abilities. Keating (1990) proposed that the language-specific phonetic details of each language, coarticulation and its magnitude are specified separately in the grammar of each language. Therefore, the amount of /l/-velarization might not necessarily be exploitable for less proficient or average language learners.

Bladon and Al-Bamerni’s concept of CR (1976) suggested that the CR value is a property induced and stored within each extrinsic allophone to which the speech production mechanism has access to. Anticipatory or carry-over coarticulation can strongly influence the phones in question, but they are inhibited if CR specification on some segments is high. As shown in Figure 1, coarticulation and also coarticulatory resistance is determined by a plenitude of factors such as universality, context-sensitivity, speaker- and language-specificity. It might be that non-proficient
learners of L2 English were not able to store and therefore produce the language-specific degrees of velarization and CR values, and, as a consequence use the exemplars to which they have access, these being less precise than those of proficient speakers. Non-proficient as well as average speakers generally coarticulate differently as it is usual in their mother tongue. Looking at the non-proficient subjects who we investigated, we must take into consideration that only 4 of them have been classified as clear anti-talents, while 2 are quite bad and 5 not great/below average. This tendency to the average level might lead to ANOVA test results for the average vs. non-proficient speakers to be insignificant (s. Tab. 2). Figure 4 and Table 4 show less clear-cut results for the least talented speaker DB as compared to the most proficient speaker AT. This suggests that also dialectal differences should be taken into consideration. Proficient speaker AT might have learnt to produce RP-English and was at the same time able to perceive and imitate the RP dialect of the Standard British English speaker taken for the priming which resulted in higher vowel formant frequencies and therefore a low degree of velarization. In comparison to AT, DB might process inter-language coarticulation/coarticulatory resistance differences less sensitively which result in a degree of velarization in the Standard British English production task, due to the inability to correctly imitate and, as a consequence, an over-generalization of the general rule that the English language is more velarized than German. Tables 3, 4 and Figure 3 might give a hint to the use of a more detailed look at individual speakers or to smaller categories of subjects (very talented vs. very good vs. talented/above average vs. average ability vs. not great vs. quite bad vs. clear anti-talent).

Coarticulatory resistance results are mostly inversely related to the degree of velarization results of Analysis 1, but not in the cases of F2 in /əly/ vs. /ˈæli/ and /əly/ vs. /ˈæli/. F2 and Fv values for proficient speakers suggest that while being more velarized, proficient speakers are less resistant to coarticulatory effects in tongue dorsum activity (see Figure 2). It might be that coarticulatory resistance at the acoustic level is not only a consequence of tongue dorsum activity, but to a certain amount also of labial or mandibular activity (Recasens et al. 1995:50). Recasens and Farnetani (1990, 1991) state that the clear dependence of articulatory constraints on gestural production is the case for languages with very different varieties of the consonant /l/ such as Italian and Catalan. In this study we compared proficient vs. average vs. less proficient non-native speakers of English who might not be able to produce such different types of the varieties. Despite being categorized as proficient, none of
these speakers produced accent-free English. At the same time, it must be considered that we did not take into consideration speakers imitating different types of dialects. The amount of velarization differs highly over the varieties of British English (Bladon & Al-Bamerni 1976), whereas in American English dialects, tongue dorsum activity depends more on the syllable-position of /l/ (Kenyon 1950), in General American /l/ is velarized in all positions.

In future work we would like to have a look at allophonic variations in subjects speaking different varieties of English, and to make sure that allophonic variations of the consonant do not appear due to proficiency-dependent vocalic differences in the sequences /dʒəl[ai, ei, y, u:]t/, but to differences in articulatory control which can be associated with the consonant. Further, we would like to unravel whether coarticulation differences in non-proficient vs. proficient speakers occur due to perceptual distinctiveness constraints or to independent learning of coarticulatory patterns, or if it is talent in the sense of spontaneous adaptation to the coarticulation/perception loop characteristic of exemplar based speech learning mechanisms (Dogil 2007; Wade et al. [submitted]). Earlier studies (e.g. Recasens 1984a, 1984b, 1985; Recasens, Fontdevila & Pallarès 1995) indicated that the investigation of CVC- and CV-segments through phonological rules cannot successfully define the presence vs. absence of coarticulation, due to its graded nature and the linguistically relevant aspects of coarticulation which are connected to it. In other words, the same segments exhibit different degrees of coarticulation across languages. To account for these facts, it is also necessary to consider factors outside the world of phonological features, i.e. articulatory constraints and aerodynamic-acoustic constraints (Farnetani 1997:390). In future studies, we will take into consideration the linguistic suprasegmental structure such as duration, speech rate and speaking style (Farnetani & Recasens 1999:32; Kühnert & Nolan 1999:28), to avoid the comparison of inhomogeneous target stimuli. This comparison should also yield interesting insights with respect to differences in coarticulation/coarticulatory resistance between German vs. L2 English speech. However, we expect to find less language-specific and more speaker-specific coarticulatory resistance patterns and even inter-speaker variability of coarticulation as a clear signature of pronunciation talent. The measurement of CR is a very precise acoustic phonetic tool, which can detect individual co-articulatory profiles of speakers with varying degrees of language aptitude.
References

Ashby, M. & Maidment, J.

Baumotte, H., Lenz, M., Dogil, G. & Reiterer, S.


Bladon, R. A. W. & Al-Bamerni, A.
1976 Coarticulation resistance in English /l/. *Journal of Phonetics*, 4, 137–150.

Bladon, R. A. W. & Nolan, F. J.

Butcher, A. & Weiher, E.

Chistovich, L. A. & Lublinskaya, V. A. V.
1979 The ‘center of gravity’ effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli. *Hearing Research*, 1, 185–195.

Clark, J., Yallop, C. & Fletcher, J.

Delattre, P.

Dogil, G.

Fant, G.

Farnetani, E.
Farnetani, E. & Recasens, D.

Gimson, A. Ch.

Hardcastle, W. J. & Hewlett, N. (eds)

Hirose, H.

Jilka, M., Baumotte, H., Lewandowski, N., Reiterer, S. & Rota, G.

Keating, P. A.

Kenyon, J. S.

Kohler, K. J.

Kühnert, B. & Nolan, F.

Lehiste, I.
1964 Some Acoustic Characteristics of Selected English Consonants. Research Center in Anthropology, Folklore and Linguistic. Indiana University, 34.

Majewski, W., Rothman, H. B. & Hollien, H.

Mangold, M.

Manuel, S. Y.
Oh, E.

Öhman, S. E. G.

Pierrehumbert, J.

Pompino-Marschall, B.

Rapp, S.
1995 Automatic phonemic transcription and linguistic annotation from known text with Hidden Markov Models/ An Aligner from German. In: *Proceedings of ELSNET goes east and IMACS Workshop “Integration of Language and Speech in Academia and Industry”*, Moscow.

Recasens, D.

Recasens, D.

Recasens, D.

Recasens, D. & Farnetani, E.

Recasens, D. & Farnetani, E.

Recasens, D., Fontdevila, J. & Pallarès, M. D.

Recasens, D., Pallarés, M. D. & Fontdevila, J.

Recasens, D.
Recasens, D.

Shibata, S.
Shibata, S., Ino, A., Yamashita, S., Hiki, S., Kiritani, S. & Sawashima, M.

Stevens, K. N. & House, A. S.

Wade, T. et al.
(subm.) Syllable frequency effects in a context-sensitive segment production model.

Ziegler, W. & von Cramon, D.
Prosodic factors in language proficiency: intonational variation as a signature of pronunciation talent

Volha Anufryk

1. Introduction

An individual’s proficiency in a given language, be it native or foreign, is determined by his or her ability to select and employ the language means prompted by the corresponding communicative situation. And while the majority of people are capable of getting their message through to the interlocutor, it is the degree of suitability, expressiveness, as well as variation, which differentiates a gifted speaker from a less successful one. Indeed, the ability to vary lexical units and grammatical patterns has often been mentioned as a talent in its own right.

In a similar fashion, intonational variation seems to be one of the building blocks of general pronunciation talent, such that a greater degree of variability would correlate with a greater degree of talent. Simple as this statement may appear on the surface, there are still quite a few aspects to take into account, especially from a cross-linguistic perspective. First of all, both phonetic and phonological factors come into play, and the relation between them may not be linear, even within one language, let alone in contrasted systems. Apart from that, in a study where we mainly investigate groups of people based on their pronunciation ability, individual variation might affect the analysis to a considerable extent.

In view of these considerations, we will first try to locate the phenomenon of intonational variation within the multitude of the existing linguistic paradigms and theories of second language acquisition. Further on, results of an experimental study will be presented to test the above hypothesis about a correlation between prosodic variability and pronunciation talent.
2. Intonational variation as a language phenomenon

2.1. Variation in linguistic theories

The correlation between constant and variant features in the language system is a cardinal problem which is explored in the majority of modern linguistic investigations. And whereas the general idea, that a language system is a heterogeneous entity, is broadly accepted nowadays, views on language variation have undergone a considerable evolution. It has been a long way from the pure structuralist abstraction to the atomic exemplar matrix, i.e. from the language code to its dynamic speech representation.

The key to understanding language variation lies in resolving two major issues. Firstly, it is important to identify the inventory of language units and their features, as well as define variability limits for each unit. Secondly, determining the dimensions of variation is crucial: should it be restricted to the systemic level, i.e. categorical variation, or should it also account for the plurality of speech realizations within a category.

In structural phonology the former issue clearly took the focus of attention. Language was considered to be a relatively self-contained sign system, which is stored as a whole only in the collective consciousness (Saussure 1916). For this reason structuralists were mainly interested in the formal description of this system.

Although intonation was not the central notion in early structuralist studies, it did receive a basic description as a dichotomy of falling and rising movements, the former having the meaning of finality and the latter bearing the continuation function (Trubetzkoy 1958).

In the language system intonation was seen as sharing the same properties as the segmental units, cf. the terms ‘secondary phonemes’ (Bloomfield 1933), ‘suprasegmentals’ (Lehiste 1970). Prosodic variation is thus generated from an abstract set of features which may vary to a certain degree, but are still perceived as the same entity. Jacobson & Halle (1971) pointed out that “tone level, or tone modulation ... are always purely relative and highly variable in their absolute magnitudes from speaker to speaker, and even from one utterance to another in the usage of the same speaker.” Despite this obvious acceptance of phonetic variation as a speech phenomenon, it was only seen as a source of material for the higher, phonological, analysis levels. Structural linguists were not interested in idiolects and individual realizations as such; they strived to generalize the language patterns of a certain community. Equally irrelevant was phonetic variability in cross-linguistic research and studies of second
language acquisition: language comparison (Hjelmslev 1974), as well as learning (Shcherba 1974), only consisted in getting a full command of its categories and invariants, regardless of the speech peculiarities.

With all due respect for the significance structural linguistics had for the understanding of language variation in general, it has to be stated that structuralist descriptions were still very abstract and limited to the phonological norm. But without regard to the phonetic variation, it is impossible to get a full perspective of how language units function in speech. Even comparative analysis is hampered when identical categories in two languages differ exclusively in their distributional phonetic characteristics.

Generative phonology shared the primacy of the phonological level introduced in structural studies.

Chomsky & Halle (1968) described the prosodic contour as a multi-level structure of temporal, stress and pitch parameters, which, in its turn, predetermines its variability. However, Chomsky only underlined its correlation to syntax, i.e. “once the speaker has selected a sentence with a particular syntactic structure and certain lexical items, the choice of stress contour is not a matter subject to further independent decision” (p.298).

If the above assumption were true, there would be one and only one realization for each utterance. Yet, prosodic variation is obligatory in syntactically ambiguous utterances; ample evidence has also been found that the relationship between syntax and prosody is non-linear (e.g. Selkirk 1995). Furthermore, extra-linguistic context, alongside with the speaker’s individual characteristics, adds up to the plurality of the possible prosodic realizations.

As evident from the arguments provided, generative phonology did not differ significantly from the structuralist positions on prosodic variation. Nevertheless, the notion of mental syllabary (Crompton 1982; Levelt 1989) is of crucial importance for the subsequent research in the field, for it marks a shift in the understanding of the language deep structure as opposed to the classical generative descriptions.

According to Levelt’s theory, infrequent and frequent syllables follow a different production pattern: the former are composed of segmental units, whereas the latter are realized as units per se, i.e. by means of retrieving the stored syllable gestures. For the theory of variation that was perhaps the first implication that at least some productions come from a certain realization, or exemplar, stored in memory.

Within the optimality theoretic paradigm prosodic variation consists in retrieving the optimal variant resulting from the interaction of the three
major elements – generator, constraints and evaluator (Prince & Smolensky 1993), whereby in the general ranking schema prosodic constraints dominate the morphological ones (McCarthy & Prince 1993: 145). It occasionally happens that in the course of evaluation more than one candidate is selected as optimal, i.e. best (e.g. Selkirk 2000: 250). Thus, OT does allow of a considerable variation in prosody, though only phonologically incurred, whereas phonetic variants remain beyond the theory.

The next big step in phonetic variation research was marked by prototype theory, with the notion of the perceptual magnet effect, introduced by Kuhl (1991). Exploring vowel space, she found evidence that “the prototype of the category assimilated neighbouring stimuli, effectively pulling them toward the prototype”.

The objectivity of this effect was supported by some studies of intonation. Schneider & Möbius (2005) reported finding the prosodic prototype for the statement category in German, which is characterized by very low $F_0$ values. However, no perceptual magnet effect was registered for the question category in that investigation.

Prototypicality effects may occur in the perception of some intonational categories, but no studies dwell upon the production of prosody in that respect.

Arguably, the greatest contribution to the theory of variation up to date has been rendered by exemplar theory, which has taken a sensible account of frequency effects on speech perception and production building up a realistic language hierarchy.

One of the main issues under investigation is what constitutes a unit of speech perception and production. In this respect, it is possible to discriminate between “mixed” models, where exemplars and the categories they constitute are viewed as coexistent (e.g. Nosofsky 1986, Pierrehumbert 2001, Bybee 2005); “pure” exemplar models, which only classify separate exemplars (sound types, syllables, pitch accents etc.) according to their distribution (e.g. Walsh et al. 2007) and “extreme” models postulating the absence of any units or types. Under this approach an exemplar is an acoustic sample stored in memory in its global context (Wade et al. to appear).

It seems logical that the last two types of models are only a derivation from the first one. Even if we oust the notions of category, unit and class from the language system, we would still need rules and algorithms for exemplar identification and production. And that in fact means re-constructing the system.

Another important question in exemplar theoretical studies concerns frequent and infrequent classes. The main claim here is that the former
are characterized by a large number of exemplars and are most eagerly produced, whereas the latter decay with time if there is no activation (Pierrehumbert 2001; Bybee 2006). It is true that frequency and recency are important factors in speech production. And the above studies present comprehensive accounts of these processes. However, frequent categories make up a relatively small percentage in the system. And getting a full perspective presupposes the knowledge of mechanisms behind the less frequent categories as well.

Noteworthy in this regard are the findings by Goldinger (1998), who pointed out some recurring effects in rare categories, namely that of unusualness. An infrequent exemplar heard in an unusual context and/or produced by an unfamiliar voice is more likely to be stored in memory than otherwise.

Yet, questions remain regarding both classes: why are some very infrequent exemplars still stored in memory and others decay, just as the contexts in which they initially appeared; and why are some frequent instances not stored in memory? The response to the above questions may be that those exemplars are, or are not, behaviourally relevant. Apart from that, speakers vary in individual perception, as well as language ability, which may all serve as a filter and obstacle for exemplar perception and production.

A common assertion about category architecture in exemplar studies is that it is represented by the so-called exemplar “cloud”, whose size depends on whether the category is frequent or not. Pierrehumbert (2001) stated that, while categories develop over time, frequent categories accumulate a larger number of exemplars. A similar effect was found by Schweitzer and Möbius (2004) for syllable durations. However, Wade et al. (to appear) predicted an opposite trend on the same phenomenon stating that frequent categories are more stable than the infrequent ones. Granted these contradictory findings, it is relevant to find out whether frequency is the key differentiator behind those language events.

The general positions on the phenomenon of language variation are developed in second language acquisition research.

2.2. Variation and interlanguage

There are two governing approaches in cross-linguistic research. The first one positions L2 as a reduced system, characterized by a basic variety (Klein & Perdue 1997) of language means, which a speaker reproduces, once he or she has been exposed to it. This theory was supported by a
number of recent phonetic investigations. Thus, it was claimed, for example, that Finns vary within a narrower pitch range in Russian (Ullakonoja 2007) and Americans in Mandarin (Bent 2005) than the respective native speakers.

According to the second approach, an L2 speaker develops an interlanguage, which accumulates features of both L1 and L2, as well as includes certain additional characteristics (Selinker 1972). Indeed, a number of studies did support this view: more variation was found in the tonal structures of Americans speaking Japanese (Ueyama 2000), as well as in the vowel representations of Spanish speakers of English (Wade et al. 2007).

The absence of consensus on phonetic variation implies that the outcome largely depends on the languages and phenomena involved. In languages with similar phonological systems, like English and German, this leads to an even more complicated interaction of variation patterns. It is accounted for by the fact that additional speaking strategies are employed: L1 categories similar to those in L2 are claimed to be transferred (Flege 1995, 1997; Wode 1981), whereas unknown, or unperceivable, L2 units tend to be avoided (Cruz-Ferreira 1987). And the degree to which all of the above factors come into play is also related to a speaker’s individual characteristics.

2.3. Dimensions of variation and research objective

A critical overview of the approaches presented above has led to our conception of intonational variation as a complex multi-dimensional phenomenon comprising several indispensable elements.

Thus, we believe that a given intonation category has certain boundaries, within which concrete token production is realized. Furthermore, if this or that token, i.e. combination of parameters, enjoys frequent activation, it is fossilized in the category as an exemplar. The component units of such an established category are its temporal and pitch events.

In the language system, both phonetic and phonological variation appear equally important, as, dialectically, quantitative aspects develop into qualitative ones under the influence of (cross-)linguistic, as well as extra-linguistic factors, pronunciation talent being our main interest area in the latter domain.

In particular, this study is aimed to investigate the variability of intonation categories, both on the phonetic and phonological level, in speakers of varying pronunciation ability. This therefore presupposes comparing the respective variation patterns in order to see whether there are any
regularities between them and how, if at all, these regularities relate to the
general linguistic and second language acquisition paradigms described
in sections 2.1. and 2.2.

The experimental procedure and the initial results of the current inves-
tigation will be presented in the remaining sections of this chapter.

3. Method

3.1. Subjects

The subject pool consisted of 38 native German speakers, 16 males and
22 females. The majority is aged 20–29, and six speakers are between 30
and 40 years old. The test persons predominantly come from the Swabian
region of Germany.

All subjects learned English at school and most of them, except for
three people, visited an English-speaking country, either sporadically,
during short trips abroad, or over a longer period of time. The latter is
ture about a third of the informants.

A comprehensive series of tests had been performed by the subjects in
the course of the DFG funded “Language Talent and Brain Activity”
project (see Jilka this volume and Jilka et al. 2008 for a detailed account
of the procedure), prior to the given experiment. As a result, their pro-
nunciation aptitude was evaluated as being below average (6 females,
3 males), average (7 females, 7 males) and above average (9 females,
6 males) based on the talent and proficiency ratings introduced by Jilka
in Chapter 2 (German system grades 3.0–4.5, 2.0–3.4 and 1.1–1.8 in the
three groups, respectively).

Additionally, a control group of twelve native English speakers (4 fe-
males and 8 males) took part in the experiment for the comparison of the
experimental results.

All subjects were paid for their participation.

3.2. Corpus and data processing

The above-mentioned subjects were recorded in the anechoic chamber of
the Institute for Natural Language Processing, University of Stuttgart,
with a sampling rate of 48 kHz.

Their task consisted in reading the classical fable “The North Wind
and the Sun”. The German subjects performed the task in both German
and English; the native speakers read the text in their mother tongue only.

The recordings were then downsampled to 16 kHz/16 Bit for further processing. Word, syllable and phone boundaries were marked automatically by forced alignment using the Aligner software for German and English (Rapp 1995) and corrected manually for any resulting discrepancies.

Manual labeling of the intonation events followed, in accordance with the general principles of autosegmental metrical phonology and the ToBI convention, but with certain modifications, seen as indispensable in the light of the present investigation.

The syllables carrying ToBI accents and boundary tones were further analyzed in a parametric intonation model (Möhler 2001).

Both the intonation labeling procedure and the parametric model will be described in the following sections.

### 3.2.1. Intonation labeling

The German and the English prosodic systems are generally described as being similar on the phonological level – they both constitute sequences of high and low targets, i.e. falling and rising F0 movements. The main differences between the languages in this respect are said to be found in the phonetic realizations (e.g. Esser 1978:51; Scuffil 1982:72). To a certain degree this can be traced from the respective ToBI conventions, although the underlying classification is a phonological one (see Beckman & Elam 1997 for English and Mayer 1995 for German).

The following comparison table lists all the pitch accents and boundary tones singled out in the two labeling paradigms.

<table>
<thead>
<tr>
<th>Pitch accents</th>
<th>H*</th>
<th>L*</th>
<th>L*H</th>
<th>LH*</th>
<th>!H*</th>
<th>H*L</th>
<th>HH*L</th>
<th>L*HL</th>
<th>H*M</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>German</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Boundary tones</th>
<th>L-L%</th>
<th>L-H%</th>
<th>H-H%</th>
<th>H-L%</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>German</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>
As is evident from the above table, the systems indeed differ only in a few major parameters. In particular, LH* accent is absent from the German notation, whereas H*L is not to be found in English. Instead, the final falling movement is labeled as a sequence of the H* and the low boundary L-L%.

In general, there are more bidirectional pitch accents in the German system, unlike the English ToBI, where, in contrast, boundary tones can involve a complex F₀ movement (e.g. Gibbon 1984:35).

In the given study H*L was used to mark the falls realized already in the stressed syllable, whereas H* combined with L-L% singled out cases where the falling movement started in the post-nuclear part. We resorted to this distinction because data analysis was complicated by the very high overall frequency and variability of the H* pitch accent at the statistical processing stage. Furthermore, this subdivision reflects a frequent realizational, as well as language-specific, peculiarity of the falling pitch accents / nuclei (Grabe 1998:213).

In general, the choice of the label, best suited for a given intonation event, was determined by the ToBI descriptions of pitch accents and boundary tones, irrespective of the source paradigm. Thus, both English and German labels were equally considered in each case in order to determine the L₁ transfer processes by German speakers, as well as to revise the frequency and distribution of intonation patterns in the English realizations.

3.2.2. Parametric intonation model

Syllables carrying ToBI accents were analyzed using a parametric model (Möhler 2001), which describes the F₀ curve in terms of the following six parameters:

- \(a_1\) and \(a_2\) signify the steepness of the rising and falling sigmoids, respectively
- \(b\) stands for the alignment of the function within the syllable
- \(c_1\) and \(c_2\) model the amplitudes of the rising and falling sigmoids, accordingly
- \(d\) corresponds to the frequency of the function’s peak

\(a_1\) and \(a_2\) values are measured in seconds; \(b\) is a relative value from −1 to 2, for the syllable is taken as a unity; whereas \(c_1\), \(c_2\) and \(d\) parameters are given in Hertz. The model therefore takes into account both temporal
and pitch characteristics of an intonation contour, as well as renders the alignment details. All this provides an accurate description of intonation events, both individually and globally.

4. Results

4.1. Phonological findings

At the first stage of data analysis we investigated the distributional and frequency parameters of separate ToBI pitch accents and boundary tones. Observable differences between the groups were revealed from the following descriptive statistical analysis.

The rising $F_0$ contour, i.e. high boundary, had a much wider distribution in the German realizations of the English version of the text than it was employed by the native speakers. The frequency was highest for the below-average group (57.5%), followed by the average speakers (48.1%). Informants of above-average ability approached native-like performance (33.4% against 19.9%, respectively), but their values were still much higher. Worthy of mention is the fact that in the German version of ‘The North Wind and the Sun’, read by the German participants only, percentage values for the high boundary were almost equal, making up 56–59% of the cases in these groups.

The low boundary, on the contrary, predominated among the native speakers with 71%, and the above-average speakers once again came
closest with 53.9%. For the below-average and average groups the low boundary was much less frequent, only covering 32–35% of all utterances. Almost the same percentages (34–36%) could be found in the German language samples, as realized by the speakers of below-average, average and above-average aptitude.

Summing up the above results, all German speakers downgraded their typical mother-tongue variation pattern when speaking English, and in the above-average group the accommodation to the target L2 variation model was strongest.

Next we looked at the distribution of individual ToBI pitch accents. Similar to the boundary distributions, German language realizations were considerably homogeneous, as far as separate percentage values are concerned.

Overall, one of the most noticeable peculiarities concerned the distribution of the L*H accent. This ToBI event was much more typical of the German speaker samples in both versions of the text under investigation. In their mother tongue it covered over 40% of all ToBI accented syllables; in English the values were lower than that, nevertheless, at least twice as frequent as in the native speaker realizations. Not only was the frequency different, there also appeared language-specific distributional peculiarities: in German the L*H accent was used both in the pre-boundary and nuclear positions; contrary to that, in English the distribution was almost exclusively IP-final.

The simple low L* and the high H* targets were most frequent for the native speaker group, which is accounted for by the fact that in English plateau-type pre-boundary accents are generally more typical (Gibbon 1984). As noticed previously, informants of above-average ability were closest to the native speakers in the H* and L* accent distributions.

Still another peculiarity we ought to mention in this respect is the clear evidence for negative prosodic transfer, e.g. as was the case with the L*HL and H*M accents, which were found in the English language samples of the German speakers. The percentages were almost equal: ~1–2% for L*HL and ~1% for H*M. However, the former accents were slightly more frequent in the below-average realizations, whereas the latter, i.e. the stylized level contour, in the above-average group. Simultaneously, a few German productions could be classified as LH*, which might be a marginal variant of the predominant L*H. Due to the fact that these transferred pitch accents made up a very small overall percentage, it is necessary to test this effect for consistency on additional speech samples of the same speakers.
Figure 2 presents all the above-stipulated findings for the ToBI accents. The second stage of the phonological analysis consisted in exploring the distribution of ToBI events on the text level in order to see how uniform, or variable, the prosodic interpretation of the text by the native and the German speakers was.

For that purpose the Levenshtein distance (Levenshtein 1966) was applied to compare all ToBI transcriptions of “The North Wind and the Sun” across the informants. This measure calculates the number of corrections necessary to transform one string of ToBI events into another and therefore shows how similar they are. For example, the following two transcriptions have a Levenshtein distance of 1 because only one substitution is necessary to transform one into the other: 1) H*L H* H* L-L% 2) H*L L*H H* L-L%.

A striking observation was the homogeneity within the speaker groups. This supports their initial subdivision based on pronunciation aptitude. The only exception was constituted by one native and two above-average speakers, whose text interpretation was highly idiosyncratic – Levenshtein distance values over 60 as compared to all the other subjects’ transcriptions (in contrast to that, the majority of values lay below 45–50). However, even those outliers did not influence the subsequent statistical analysis.

In the English version of the text the greatest uniformity in the text interpretation, i.e. the lowest Levenshtein values, could be found within the native and the average groups (values ~18–35, see Figure 3). This could most probably be explained by their neutral reading manner. Apart from that, the context itself imposes the choice of intonation means to a certain degree, and that knowledge is obviously shared by speaker communities. The above-mentioned reasons might also account for a greater uniformity in all German realizations of ‘The North Wind and the Sun’, both on the pitch accent and the boundary level.

A greater degree of variation in the prosodic text interpretation by the above- and below-average speakers (values ~35–55) could have two possible explanations as well. On the one hand, these speakers might have aimed at an original way to read the text, which seems more typical of above-average informants. On the other hand, increased variation could be caused by the L2 speakers’ uncertainty about the intonation means appropriate in each particular context – characteristic for the below-average informants.

To corroborate the Levenshtein distance analysis, we explored the values by means of a paired samples t-test. The Levenshtein value distributions in each aptitude group were compared against each other.
Figure 2. ToBI accent distribution in English and German in all aptitude groups
Figure 3. Phonological variability in text interpretation
group, both in the English and the German versions of “The North Wind and the Sun”. The \( p \)-values for the majority of the comparison pairs lay below 0.01, which shows that there were significant differences in text interpretation across the aptitude groups. However, the following group comparisons did not yield any significant results:

1) on the pitch accent level
   - in English – below-average vs. above-average and average vs. native
   - in German – average vs. below-average

2) on the boundary tone level
   - in English – above-average vs. below-average
   - in German – above-average vs. average and average vs. below-average

In general, the \( t \)-test results supported the Levenshtein measure analysis presented above: statistical significance in that respect stands for a greater / lesser variability in the text interpretation between the groups, whereas the opposite signifies an almost equal degree of variation.

4.1. Phonetic findings

Similar to the phonological analysis of ToBI categories, described in the previous section, the investigation of the intonation model parameters was conducted within two stages: discretely – each separate parameter being explored across the aptitude groups; and globally – by means of a cosine similarity measure, which takes account of all the six parameters as a whole.

Prior to the investigation proper, we subdivided all the given ToBI events into boundary, or nuclei, and pre-boundary ones. This was performed mainly in order to take into account the phrase final declination and lengthening phenomena. This subdivision was also suggested by the initial statistical analysis.

As a result, we singled out the four most frequent intonation subcategories to be investigated in our study: namely the pre-boundary events H* and L*H; as well as the nuclei – the final fall (realized by a sequence of either H* or H*L, followed by a low boundary tone L-L%) and the final rise (L*H in combination with the high boundary H-H%).

It should be noted in this regard that due to the relative inequality of speaker groups (see Section 3.1.) and, consequently, sample sizes, all cases were weighted, with talent being the frequency variable. In addi-
tion, male and female realizations were treated separately because of the evident differences in pitch, described in earlier investigations (e.g. Loveday 1981), as well as supported by our preliminary statistical tests.

4.1.1. Investigation of individual temporal and pitch parameters

One-way ANOVAs were calculated for each parameter in all the four intonation categories mentioned above, with talent as the independent variable. The tests yielded significant differences \((p < .05)\) between the speaker groups in most of the comparisons. And whereas each separate sample pair in all of the conditions under investigation still calls for further analysis (the post-hoc Scheffé tests did not elicit any consistent pattern), it was possible to determine one steady trend.

The \(d\)-parameter, or \(F_0\) peak frequency, was consistently different \((p < .01)\) for all female speakers in both the English and the German versions of ‘The North Wind and the Sun’. The same holds true for the male realizations, with the only exception being made by the L*H category.

Not only were the samples statistically different, they also brought an interesting peculiarity in the above-average productions to the surface: both female and male speakers realized the \(F_0\) peaks on a higher pitch level than all the other participants (even in the aforementioned L*H case).

The other \(d\)-parameter effects were category- and gender-specific.

The English H* samples were characterized by an almost equal degree of variation in all German groups (see Figure 4), while the corresponding realizations by the native informants were visibly more variable. The German language realizations had a rising stepwise pattern with the female subjects, both from the pitch-level and variation perspectives; the male participants, on the other hand, did not follow that trend, although there was a doubled degree of variability in the above-average group.

The falling nuclei also had the highest \(d\)-values in the above-average group in the English language realizations, as mentioned previously, closely followed by the native speakers. The degree of variability was almost equal in those groups, as opposed to the below-average and average realizations of the final fall, which were significantly more concentrated, as well as located observably lower on the Hertz scale (females: 200–250 Hz for Germans, 170–230 Hz for native subjects vs. ~180–200 Hz in the average and below-average groups; males: ~100–140 Hz for both groups in question vs. 100–105 Hz as realized by the average and below-average informants).
Figure 4. $F_0$ peak frequency distribution in the H* accents in all aptitude groups
The L*H category in female realizations followed the above pitch-level and variability pattern. The $d$-parameter values rose as high as 280–300 Hertz in the above-average and native female productions, the degree of variability being almost equal in these two groups, whereas lower $F_0$ peak frequency (up to 250 Hz) and less variation were found in the below-average and average samples. Interestingly, average male participants produced $F_0$ peaks on almost the same level and with a similar degree of variation as the native speakers (~125–155 Hz). Nevertheless, the above-average L*H events were still more variable and higher in pitch (130–170 Hz).

As mentioned in the previous section, the rising nucleus was the most frequent boundary category in the German speaker samples. And here the degree of variation, as well as the pitch-level, was almost equal for all German female informants. The native speakers realized the given category within a significantly lower pitch range: 160–220 Hz vs. 200–260 Hz, accordingly. The corresponding male productions were somewhat unexpected in the light of the general tendency: the final rises were produced on almost the same $F_0$ level by the below- and above-average informants, with the degree of variation being observably smaller for the former. In the German language samples participants of below-average ability even surpassed all the other informants, as far as pitch-level is concerned (~165 vs. 150 Hz, respectively)\(^1\).

The remaining parameters did not reveal any obvious tendencies. Therefore, further investigation of temporal and pitch aspects is necessary, both individually and in their interrelations.

4.1.2. Investigation of similarity and variation in $F_0$ curves.

Because the exploration of individual parameters did not elicit any obvious trends, except in the $F_0$ peak frequencies, a cosine similarity measure was applied in order to compare each pair of $F_0$ curve tokens in the most frequent intonation categories – $H^*$, $L^*H$, falling and rising nuclei. For that purpose all the parameter values were scaled to the corresponding $z$-scores (calculated by subtracting the population mean from an individual raw score and then dividing the difference by the population standard deviation). As a result, six $z$-transformed parameter values were available for each token, in all the four intonation categories and for all

---

1. For the visualization of those phonetic peculiarities in L*H accents, as well as the rising and falling nuclei, please refer to figures in the Appendix.
the subject groups, gender-specific. These values were treated as six-di-
dimensional vectors, for which we extracted the respective cosine similarity
measures (calculated by dividing the dot product of two vectors by the
product of their magnitudes)\textsuperscript{2}. Each comparison yielded a score in the
range from –1 to 1: the former stands for high dissimilarity and the latter
for high similarity.

Further on, these cosine similarity scores were compared using the Kol-
mogorov-Smirnov test. Group differences in most comparisons were insig-
nificant, according to the test statistics ($p > .05$). However, average group
realizations did produce a significant effect ($p < .05$) in the English samples
of ‘The North Wind and the Sun’. Female values lay well above, whereas
male tokens visibly below the other density functions (see Figure 5).

Another unexpected peculiarity concerned the L*H accent in male
German language samples, where the below-average realizations were
significantly different from the corresponding average and above-average
values (see Figure 6). Bearing in mind the high frequency of the given
pitch accent, this finding requires further investigation as to factors / reasons
it was triggered by, especially since the corresponding female productions
were rather uniform in that respect.

Overall, the results imply that, as far as the F\textsubscript{0} curve similarity / dis-
similarity is concerned, the four aptitude groups realize various inton-
ation categories in a similar way, with the exception of the two cases de-
scribed above.

5. Discussion

The results of the present study confirm the initial hypothesis of a corre-
lation between prosodic variation and language ability. Significant differ-
ences between the groups were found both in the distribution of the ToBI
categories and in the realization of the individual F\textsubscript{0} curve parameters.

The phonological findings were expected and showed a clear trend to-
wards the native patterns in the German productions, i.e., first and fore-
most, a wider distribution of the L*H accents and high boundary, as well

\textsuperscript{2} The dot product of two vectors of equal size is the sum of products of the
corresponding values in these vectors. For example, the dot product of the
vectors a=[1,2,3] and b=[4,5,6] equals 32, i.e. $1*4+2*5+3*6$. The magnitude
of a vector is the square root of the dot product of the vector with itself. For
example, for vector a=[1,2,3] the magnitude is $\sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$, or $\approx 3.7$. 
Figure 5. Distribution of cosine similarity scores in the final fall category in the English version of “The North Wind and the Sun”
Figure 6. Distribution of cosine similarity scores in L*H accents in the German version of “The North Wind and the Sun.”
as a smaller percentage of falling contours, in contrast to the native speakers. That result confirms an earlier investigation on the problem: Anderson (1979) discovered a similar trend.

Not only was the frequency different, there also appeared language-specific distributional peculiarities: with English speakers the distribution of L*H was almost exclusively IP-final; contrary to that, German speakers employed the given accent both in the pre-boundary and nuclear positions.

Nevertheless, this L1 transfer was accompanied by another important phenomenon – accommodation to the English variation patterns. It can be accounted for by the fact that although the typical German patterns were prevalent in the below-average, average and above-average speaker samples, these values were still lower in English than they were employed in the German version of “The North Wind and the Sun”. And the greatest degree of accommodation to the L2 variation pattern could be found within the above-average aptitude group.

Another evidence of negative L1 transfer was the employment of the typical German pitch accents, namely, L*HL and H*M, in the English language samples of the German groups. The rising-falling “hat” accent L*HL was used most extensively by the below-average speakers (in about 1–2% of the cases) and H*M by the above-average informants (covering about 1% of the cases). There were a few instances of the rising-falling pattern in the native speaker samples as well, but, similar to L*H, they tended towards the end of the intonation phrase. In addition to that, there was a phonetic difference: the German L*HL was realized within the stressed syllable, whereas the English rise-fall extended onto the post-nuclear part. However, due to the overall infrequency of these intonation events, it is necessary to investigate additional samples of the same speakers in order to see whether this L1 transfer trend should persist.

The additional measure applied to compare the differences in the intonational text interpretation on both pitch accent and boundary tone level also produced statistically significant results. Firstly, similar ranges of Levenshtein values revealed the relative homogeneity within each aptitude group as opposed to the other ones.

Average and native informants were characterized by the greatest degree of agreement, i.e. the shortest Levenshtein distance. This could most probably be explained by their similar neutral reading manner. The above reason might also account for a greater uniformity in all German realizations of ‘The North Wind and the Sun’, both on the pitch accent and the boundary level.
A greater degree of variation in the prosodic text interpretation by the above- and below-average speakers could have the following explanations. The former might have aimed at an original, i.e. more expressive, way of reading the text, whereby in the case with the latter increased variation can be caused by the L2 speakers’ uncertainty about the intonation means appropriate in each particular context. Apart from that, as has been mentioned above, below-average productions were marked by the greatest degree of negative transfer from German, which, in its turn, increases variability due to combining the pitch accent inventories of the two languages.

This analysis of intonational text interpretation produced some expected, as well as unexpected results. The greater variability in above-average productions appears to support the hypothesis under investigation. On the other hand, an almost equal degree of variation in the below-average group calls for a subdivision of individual tokens into the typically German and validly English ones. This way a realistic picture can be drawn about “bad” and “good” variation, respectively.

The two-step phonetic analysis also yielded some interesting results, both as far as individual parameters are concerned and in the global similarity measure.

One steady trend was discovered in the $F_0$ peak frequency parameter, which was consistently realized on a higher pitch-level, and often with a greater degree of variation, by the above-average subjects as compared to the other groups. This finding in fact disagrees with the results of an earlier study by Anderson (1979) who claimed that English utterances are higher in pitch than the German ones. Indeed, native speaker productions came closest to the above-average group in the H* (for males) and L*H (for females) accents, as well as in the final fall category (for all speakers). However, the low rise had significantly higher $F_0$ peak values in all German groups. And in the female productions of the H* events below-average, average and native realizations were located on an almost equal pitch-level.

Below-average and average productions were marked by an almost equal lower degree of variability in most conditions, except for the H* category, where it approached the above-average and native patterns. So, in general, this finding seems to agree with the initial experimental hypothesis. However, it is necessary to look at the below-average and average groups more closely, perhaps subdividing the subjects into smaller groups, which might render more information on the variation / uniformity of the respective intonation patterns.
No obvious tendencies were discovered in the remaining parameters, although the majority of comparisons were significantly different, according to the respective ANOVA results and post-hoc Scheffé corrections. This therefore seems to be a valid enough incentive for further investigation. In particular, it would be interesting to explore the temporal parameters, i.e. $a_1$ and $a_2$, indicating the velocity of falling and rising sigmoids. Rather intuitively, we would assume that there should be less variation and lower value ranges in the above-average and native groups due to higher proficiency and fluency levels, but this speculation needs to be tested in sufficient detail.

The $b$-parameter, or the alignment of the function within a syllable, also has potential for further exploration, as some studies claim, for example, that there is a fundamental difference in the realization of the falling category in English and German (Grabe 1998:214). Thus, we could investigate the falling nuclei and pitch accents, as well as the other intonation categories, in order to see to what degree, if at all, this effect is present in all of the four aptitude groups.

Finally, $c_1$ and $c_2$, or the Hertz values of the rising and falling sigmoids, could render some fine-grained detail about the actual shape of each separate F0 curve. And that, in its turn, would make it possible to single out some more and less frequent pitch accent shapes typical of each group of participants.

Already at this relatively early research stage, we were able spot some indications as to those prototypical contours: in each intonation category there were recurring patterns of all six values, possible exemplars, so analyzing their regularities in a dynamic model seems a sensible step to undertake henceforth.

The second stage of phonetic analysis consisted in extracting cosine similarity measures for the four investigated intonation categories in order to compare all the six parameters as a whole. The Kolmogorov-Smirnov statistics only revealed a significant effect ($p < .05$) for the average group vs. the other subjects in the final fall category in the English version of “The North Wind and the Sun”. Bearing in mind the relative infrequency of this intonation unit in German samples, this finding could point out to its instability resulting in highly dissimilar values.

A surprising effect was found in the L*H accents of below-average male participants, whose values tended towards the dissimilarity region (cosine values below zero), whereas the other density functions were characterized by highly uniform distributions. The finding is unexpected
because of the very high frequency of $L^*H$ accents in all German groups, which should assume a more regular variation pattern.

On the whole, a closer exploration of the cosine similarity measure is required, such as expanding the corpus to spontaneous speech recordings, for instance, in order to draw more accurate conclusions about the uniformity/variation of separate $F_0$ curves. Apart from that, we need to study the remaining pitch accents to include the infrequent categories and their peculiarities into analysis.

From a general linguistic perspective, the findings of the current study appear to fit into the variation scheme proposed in Section 2.3.: all categories under investigation varied within certain boundaries (originally a structural linguistic concept); at the same time the variation phenomena within these categories can be best described in exemplar theoretic terms. In particular, frequent, or entrenched, intonation events were generally marked by a higher degree of variability and a more regular variation pattern, i.e. a relatively uniform distribution of exemplars, with the exception of the above-mentioned $L^*H$ cosine similarity distribution in the below-average male group. Infrequent categories were less stable. If their constituent exemplars had more or less similar values, as was the case with the $H^*L$ accents in the average and below-average groups, this resulted in a low degree of variation, which was supported by the $d$-parameter analysis. Should the tokens lay further apart, this might create an effect of “false” variability, especially if the corresponding values are averaged. The latter cases were not described in the given study, as they require further exploration for consistency.

As a final remark, we ought to point out that although the experimental results seem to concur with the initial hypothesis that the degree of variation increases with talent / proficiency, it is important to discriminate between the so-called “positive”, or L2 proper, and “negative”, or transfer of typical L1 patterns, variation. Whereas the former should be nurtured and considered a signature of a better proficiency and a greater talent, negative transfer and the resulting variation should vice versa account for a lower rating on the talent scale.
Appendix

Distribution of the $F_0$ peak frequency parameter in L*H accents (Figure A), falling (Figure B) and rising (Figure C) nuclei.
Prosodic factors in language proficiency

Figure B. Rising nuclei
Figure C. Falling nuclei
References

Anderson, K. O.  

Beckman, M. E. & Elam, G. A.  
1997 *Guidelines for ToBI Labelling*. Ohio State University. http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/

Beckman, M. E. & Pierrehumbert, J.  

Bent, T.  

Bloomfield, L.  

Bybee, J.  

Chomsky, N. & Halle, M.  

Crompton, A.  

Cruz-Ferreira, M.  

Esser, J.  

Flege, J. E.  

Flege, J. E.  

Goldinger, S. D.  

Gibbon, D.  
Grabe, E.
1998 *Comparative Intonational Phonology: English and German*. MPI Series in Psycholinguistics 7, Wageningen, Ponsen en Looien.

Hjelmslev, L.

Jakobson, R. & Halle, M.

Jilka, M., Anufryk, V., Baumotte, H., Lewandowska, N., Rota, G. & Reiterer, S.

Klein W. & Perdue C.
1997 The Basic Variety (or: Couldn’t natural languages be much simpler?). *Second Language Research*, 13(4), 301–347.

Kuhl, P.

Lacerda, F.

Levelt, W.

Levenshtein, V.

Lehiste, I.

Loveday, L.

Major, R. J.

Mayer, J.

McCarthy, P. & Prince, A.
Möhler, G.

Möhler, G. & Mayer, J.

Nosofsky, R.M.

Rapp, S.
1995  Automatic phonemic transcription and linguistic annotation from known text with Hidden Markov Models / An Aligner for German. In *Proceedings of ELSNET goes east and IMACS Workshop “Integration of Language and Speech in Academia and Industry”*, Moscow.

Scuffil, M.

Pierrehumbert, J.

Prince, A. & Smolensky, P.

Saussure, F. de
1916  *Cours de linguistique générale*. Payot.

Schneider, K. & Möbius, B.

Schweitzer, A. & Möbius, B.

Selinker, L.

Selkirk, E.

Selkirk, E.

Shcherba, L.
Trubetzkoy, N.

Ueyama, M.

Ullakonoja, R.
2007 Comparison of pitch range in Finnish (L1) and in Russian (L2). In Proceedings of the XVIth Congress of Phonetic Sciences, Saarbrücken, 1701–1704.

Walsh, M., Schütze, H., Möbius, B. & Schweitzer, A.
2007 An exemplar-theoretic account of syllable frequency effects. In Proceedings of ICPhs XVI, Saarbrücken, Germany, 481–484.

Wade, T., Jongman, A. & Sereno, J.

Wade, T., Dogil, G., Schütze, H., Walsh, M., & Möbius, B.

Wode, H.
1997 Perception and production in learning to talk. In S. J. Hannahs & M. Young-Scholten (Eds.), Focus on phonological acquisition (17–46), Amsterdam/Philadelphia: John Benjamins.
Direct brain feedback and language learning from the gifted

Giuseppina Rota

1. Introduction

The ability to communicate in one or more languages has gained considerable interest in recent years. We live in a mobile world that is becoming more and more cosmopolitan, where economical exchange and trade markets draw countries closer. Ours is a globalizing era, which brings us into contact with cultural diversity, an epoch of remarkable changes and challenges. This cosmopolitan world fosters the need to be able to communicate in more languages. In the reality of every day, an individual’s ability to communicate is becoming more and more important: it is the indispensable key to more choices, so as to take advantage of greater cultural opportunities and the benefits of diversity. That is why understanding the mechanisms that make it possible to learn second languages (L2) and disentangling the mysteries of phonetic talent should be among the main pursuits of neurolinguistics and cognitive neuroscience.

The first insights into brain specializations emerged from studies conducted on patients suffering from strokes or brain injury. Since then, major efforts of scientific research have been devoted to deciphering the shortcomings of the impaired brain and to understanding the neuronal basis of lost functioning. Over time, this approach has yielded important insights into the functioning of the intact brain as well. In recent years however, it has become clear that one should not be content with this approach only. We should devote our efforts to uncovering the mechanisms that enable some people to develop their skills to an extraordinary extent, thus reaching levels of great expertise. We should learn those strategies to make them accessible to people who need to refine their abilities and use this knowledge to promote individual enhancement. Given the social environment where we live, this holds true particularly for the gift of phonetic talent.

The conjunct efforts of our research group have led us to highlight psychological and cognitive factors that may help during the acquisition of L2. We investigated the “nature” of phonetic talent by analyzing the
ability to discern subtle phonetic differences thereby answering to the question as to whether a good ear helps in the learning of speaking. Moreover, we addressed the neuronal basis of talent and imaged the gifted brain through functional and anatomic Magnetic Resonance Imaging (fMRI/MRI). We hypothesized that the development of exceptional phonetic abilities leads to a more efficient use of the brain circuits responsible for speech perception and production. We speculated that this process would even reshape brain anatomy in particularly talented subjects.

In this chapter I address further topics. Are there ways of directly intervening on the brain, as we do when we take gym classes to exercise muscles and enhance our physical strength? Are there ways of using the knowledge that we have gained so far to boost skill enhancement?

2. Neuronal plasticity, cognitive flexibility and skills development

Until recently scientists have accepted the idea that the human brain cannot change with age, and that no new neurons can be generated during adulthood. During the last decades, new scientific tools have allowed us to take a closer look at the complexity of the nervous system, and to gain insights into its functioning. Only recently scientists have observed that brain plasticity is a life-long phenomenon. Studies have shown that the generation of new brain cells occurs in the hippocampus as an effect of environmental stimulation (Gould 2007; Gage et al. 2002; Kaplan and Hinds 1977). Exposure to environmental enrichment induces robust neuronal plasticity in the nervous system, which includes modifications in the thickness of the cortex, in the dimensions of the neuronal soma and the dendritic spines (Dhanushkodi and Shetty 2008; Johansson and Belichenko, 2002; Mohammed et al. 2002; Rosenzweig et al. 1962).

Neural plasticity induced by experience occurs all through life and affects loss and increase of dendritic spines (Hickmott and Ethell 2006; Segal 2001), as well as the recruitment of new neurons in the circuitry of brain areas such as the hippocampus and the olfactory bulb, which subserve critical learning functions (Gould 2007).

The brain naturally repairs itself in cases of endogenous damage or external injury. Evidence had shown that a cortical reorganization might follow brain damage, and might induce a partial or complete recovery of the lost functions (Di Filippo et al. 2008; Moucha and Kilgard 2006; Jenkins and Merzenich 1987). This process reflects the enormous ability
of cortical networks to adapt to traumatic and harmful events. Brain areas adjacent and/or contralateral to the injured tissues may take over functions previously carried out by those regions, thus recovering their functioning (Winhuisen et al. 2007, 2005). This phenomenon happens in cases of language loss and for a variety of impairments such as motor or sensory disruption (Masiero and Carraro 2008; Kleim and Jones 2008; Winhuisen et al. 2007, 2005; Sabel and Kasten 2000) and plays a major role in rehabilitation. Neurological interventions capitalize on one important property of the brain: its amazing adaptability.

Cortical regions responsible for the processing of sensory inputs remodel and enlarge as an effect of repeated stimulation of the sensory epithelium (Sterr et al. 1998; Elbert et al. 1995). The same concept holds true for another vital area of human growth and personal development: learning. When we learn a new skill and engage in the repetitive training of this ability, such as when playing tennis or skiing, we do not only increase our muscular power and the flexibility of our articulations. We also train our nervous system: we stimulate the brain networks that make it possible for us to coordinate our movements and engage in those activities. Visible changes in brain morphology have been reported for the exercise of motor skills, and interestingly for language learning and musical expertise (Mechelli et al. 2004; Callan et al. 2003; Lotze et al. 2003; Pantev et al. 2001; Karni et al. 1995; Raichle et al. 1994). Mechelli and colleagues (2004) found that learning a second language increases the density of grey matter in areas of the left inferior parietal cortex. These researchers documented a correlation between the extent of local structural reorganization in those cortical regions and the level of speech proficiency of their subjects. In line with these findings, Bialystok and al. (2005) and Craik and Bialystok (2005) suggested that bilinguals possess a higher degree of mental flexibility, a condition that is likely to result from the prolonged exposure to multiple languages. According to those authors such a constant training presumably affects brain plasticity. Those mechanisms reveal their remarkable value whenever an organism must meet a new environmental demand. Those are the circumstances when experience most shapes us.

Adults who face the challenge of learning L2 take advantage of interventions that rely on brain plasticity mechanisms. Often, adults have difficulties discerning subtle acoustic nuances and pronouncing the sounds of non-native languages (Kuhl et al. 1992). People can benefit from special training and improve their language performances (Louis et al. 2001; Pulvermüller et al. 2000; Mohr et al. 1998). It seems, for instance, that
auditory training helps to reduce or eliminate a foreign accent (Iverson et al. 2004).

Understanding the mechanisms that regulate plasticity in the brain is essential to be able in the future to control those processes, influence learning mechanisms and possibly enhance them. Is there a way of directly intervening on the brain? Can we consciously stimulate our brain plasticity? Thanks to modern neuroimaging methods and newly developed paradigms we can now answer: yes, we can.

3. Real time functional Magnetic Resonance Imaging

Can humans act upon the instrument of the brain, thereby changing their own behaviour? Increasing evidence from the 60’s onwards has shown that humans can voluntarily control different components of their electroencephalographic (EEG) signal through biofeedback (for a review refer to Birbaumer and Cohen 2007). Showing ongoing changes of the EEG signal to the subjects is an effective method to allow them to find useful strategies to learn volitional control over it, and be able to modify it at will. This method has proven efficacious for important clinical applications. Kotchoubey and colleagues (2001) showed that patients with refractory epilepsy can learn to prevent seizures from happening by self-regulating their slow cortical potentials. Strehl and collaborators (2006) observed a reduction of distractibility and impulsivity in children diagnosed with the attention deficit/hyperactivity disorder who learned to modulate their electrical brain activity. The same method has been applied to communicate with completely paralyzed patients through a process that incredibly recalls mind reading (Birbaumer et al. 1999).

Physiological self-regulation of the blood oxygen level-dependent (BOLD) signal through real-time fMRI (rtFMRI) is a new paradigm of neuroscience that might be suitable for studying and even inducing brain plasticity (Sitaram et al. 2007; Weiskopf et al. 2007). Compared to EEG, an fMRI-based system offers extraordinary possibilities. fMRI allows us to obtain information on the ongoing activity of every single point in the brain with an astonishing precision that is in the order of the millimeter (Hinterberger 2005, 2004, 2003; Weiskopf 2004). This means that an individual can observe in real-time his own brain functioning while he is performing all kinds of mental imagery, recalling past events or planning future ones. Through operant conditioning (i.e. by coupling successful behavior and reward) even more can be done: in theory an individual can
learn to control activation in any given area of interest, he can *play* with his own brain. The ability to increase the blood supply in circumscribed areas of the brain opens up to another important aspect: a self-induction of desirable behavioural effects that are modulated by the brain sites under control.

Studies that have addressed this topic so far have yielded interesting results. For instance, deCharms and colleagues (2005) investigated the effects of the regulation of the rostral part of the anterior cingulate cortex (rACC) on the sensitivity to pain. Chronic pain is a syndrome that affects a relatively high number of patients in the world. The physiological causes of chronic pain are mostly unknown and patients affected by this syndrome usually undergo a number of treatments without reporting any benefit. Imaging studies have suggested the involvement of rACC in the perception of painful stimulation (Apkarian et al. 2005; Ploner and Schnitzler 2004). In their study, deCharms and colleagues described a reduction of sensitivity in patients with chronic pain following the volitional decrease of activation in this brain area.

4. Operant conditioning of the language system

To date, only a few studies have addressed language processing through paradigms of operant conditioning. Mohr and colleagues (1998) investigated the effects of the self-regulation of slow cortical potentials (SCP) on language processing. In a lexical decision task on words, they showed that the speed of answers improves as an effect of the degree of self-induced negative shifts in cortical potentials. Using an analogous paradigm Pulvermüller and coworkers (2000) targeted SCPs recorded above left-hemispheric language cortices only, substantially confirming those results but showing a more pronounced effect for words versus pseudo words.

We decided to focus on an area, whose involvement in speech processing is not completely understood yet (Rota et al. 2008). In a recent study conducted in our laboratory we used real-time fMRI to investigate whether healthy volunteers can learn to improve linguistic performances by enhancing activation in the right inferior frontal gyrus (rIFG). Figure 1 presents a schematic illustration of the real-time fMRI system we used.

This study addressed the question whether the self-regulation of this brain site would trigger short-term measurable changes in speech percep-
Figure 1. Real-time fMRI system and data flow (adapted from Dogil and Rota in press). Systems for real-time fMRI are typically constituted of multiple components, which must work in a synchronized fashion to allow a subject to view and learn to control his brain activation. We acquired functional images using an MRI whole body scanner (Magnetom Trio, Siemens, Erlangen, Germany). fMRI scans of the brain provided information on the changes in local blood oxygenation that were recorded in our region of interest (ROI), the rIFG. We used the Turbo Brain Voyager software (Brain Innovation, Maastricht, The Netherlands; Goebel 2001) to correct the functional data for artefacts and perform statistical analysis on them (refer to Weiskopf et al. 2003 for a detailed description of this fMRI-biofeedback system). We presented time-courses of the signal extracted from the ROI with a delay of 1.5 s from image acquisition. The subjects could view the changes of this signal through a visual feedback, the constantly updated bars of a thermometer, as displayed in the figure. Self-regulation of circumscribed brain regions works through mental imagery techniques (for examples on mental imagery for EEG-biofeedback, refer to Neumann et al. 2003). All experimental subjects who took part in the experiment learned to voluntarily increase the level of activation recorded from in the rIFG during three sessions of training.
tion. According to recent findings, the right hemisphere plays a role for certain pragmatic aspects of language (Cutica et al. 2006) and among these for the expression of emotions in speech (Friederici and Alter 2004; Dogil 2003; Kotz et al. 2003). Lesions of the right hemisphere might compromise speech abilities, reduce the comprehension of metaphors and humour (Coulson and Williams 2005; Lehman Blake 2003; Brownell et al. 1990), and alter prosodic expression (Gandour et al. 1995). These deficits might reflect a high degree of insensitivity to the peculiarities of communicative situations, which results in a limited capacity for the patients to understand and reproduce emotional nuances (Pell 2007, 2006). Clinical data suggest a role for the right hemisphere during the recovery of language following aphasia, as well as during therapy based on melodic intonation (Code 1987; HelmEstabrooks 1983), which presumably relies on brain plasticity mechanisms.

Stimulated by these findings, we trained our subjects to enhance activation in frontal areas of the right hemisphere, and we investigated their performances during the identification of affective prosody. Furthermore, we tested grammaticality judgments and attentional processing to control for the specificity of the behavioural effects of up-regulation of this site. The results of our experiment show that an increase of activity in the rIFG correlates to an improvement in accuracy and reduced reaction times during prosody identification. Changing activation in the right IFG influenced prosody but did not affect syntax or attention. This result suggests that areas of the “language network” can be trained separately to improve specific performances.

5. Conclusions

In the last chapter we introduced the concept of real-time fMRI and suggested that this method is suitable for self-regulating activity in circumscribed regions of the brain. We argued that this imaging technique might be useful to exert specific behavioural effects that depend on the functions of the regulated areas. So far, the potential of real-time fMRI for language has remained almost unexplored, the only exception being the study conducted by our group. The results of our investigation suggest that self-regulation of local activation is feasible, and that language areas may represent good candidates for neuro-feedback training. These preliminary data are promising for a number of applications. Imaging and clinical evidence suggest the relevance of the right hemisphere for the
comprehension and production of emotional features in speech. Our findings corroborate this evidence: increasing activity in the right IFG affects the identification of prosody and improves its accuracy. As the capacity of decoding affective prosody is altered in disorders such as social phobia and depression (Quadflieg et al. 2007; Kan et al. 2004) this technique could be an effective adjuvant means during psychological and/or pharmacological treatment of those disturbances. Non-invasiveness and the potential absence of side effects might make this technique a choice preferential in the future.

It is plausible that self-regulation of specific brain areas could favour speech production as well. Among others, Seliger (1982), Krashen (1981) and Lamendella (1979) suggested that the right hemisphere plays an important role when L2 learning takes place in adulthood. Real-time fMRI could serve for the treatment of various language problems concerning language acquisition and re-acquisition, thereby improving skills that are important for communicative competences. In short, it could become an instrument for promoting phonetic skills in people who encounter difficulties in eliminating their mother-tongue accent, a tool for shaping plasticity and promoting enhancement.

References

Code, C.  

Coulson, S., Williams, R.F.  

Craik, F., Bialystok, E.  

Cutica, I., Bucciarelli, M., Bara, B.G.  
2006 Neuropragmatics: Extralinguistic pragmatic ability is better preserved in left-hemisphere-damaged patients than in right-hemisphere-damaged patients. Brain and language, 98 (1), 12–25.

deCharms, R.C., Maeda, F., Glover, G.H., Ludlow, D., Pauly, J.M., Soneji, D., Gabrieli, J.D., Mackey, S.C.  

Dhanushkodi, A., Shetty, A.K.  

Di Filippo M., Tozzi A., Costa C., Belcastro V., Tantucci M., Picconi B., Calabresi P.  

Dogil, G. Rota, G.  

Dogil, G.  

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., Taub, E.  

Friederici, A.D., Alter, K.  

Gage, F.H.  

Gandour, J., Larsen, J., Dechongkit, S., Ponglorpisit, S., Khunadorn, F.  
Goebel, R.  
2001 Cortex-based real-time fMRI. *Neuroimage, 13*, S129.

Gould, E.  

Helm-Estabrooks, N.  

Hickmott, P.W., Ethell, I.M.  

Hinterberger, T., Veit, R., Wilhelm, B., Weiskopf, N., Vatine, J.J., Birbaumer, N.  

Hinterberger, T., Weiskopf, N., Veit, R., Wilhelm, B., Betta, E., Birbaumer, N.  


Iverson, P., Bannister, K., Hazan, V.  

Jenkins, W.M., Merzenich, M.M.  

Johansson, B.B., Belichenko, P.V.  

Kan, Y., Mimura, M., Kamijima, K., Kawamura, M.  

Kaplan, M.S., Hinds, J.W.  

Kleim, J.A., Jones, T.A.  

Kotchoubey, B., Strehl, U., Ulmann, C., Holzapfel, S., Konig, M., Froscher, W., Blankenhorn, V., Birbaumer, N.  


Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B.  

Krashen, S.  

Lamendella, J.  

Lehman Blake, M.  

Lotze, M., Scheler, G., Tan, H.R., Braun, C., Birbaumer, N.  

Louis, M., Espesser, R., Rey, V., Daffare, V., Di Cristo, A., Habib, M.  

Masiero, S., Carraro, E.  


Mohr, B., Pulvermüller, F., Schleichert, H.  
Moucha, R., Kilgard, M.P.  

Neumann, N., Kübler, A., Kaiser, J., Hinterberger, T., Birbaumer, N.  

Pantev, C., Engelien, A., Candia, V., Elbert, T.  

Pell, M.D.  
2007  Reduced sensitivity to prosodic attitudes in adults with focal right hemisphere brain damage. *Brain and language, 101*, 64–79.

Pell, M.D.  
2006  Cerebral mechanisms for understanding emotional prosody in speech. *Brain and language, 96*, 221–34.

Ploner, M., Schnitzler, A.  

Pulvermüller, F., Mohr, B., Schleichert, H., Veit, R.  

Quadflieg, S., Wendt, B., Mohr, A., Miltner, W.H., Straube, T.  

Raichle, M.E., Fiez, J.A., Videen, T.O., MacLeod, A.M., Pardo, J.V., Fox, P.T., Petersen, S.E.  

Rosenzweig, M.R., Krech, D., Bennett, E.L., Zolman, J.F.  

Rota, G., Sitaram, R., Veit, R., Erb, M., Weiskopf, N., Dogil, G., Birbaumer, N.  

Sabel, B.A., Kasten, E.  

Segal, M.  

Seliger, H.  
Sitaram, R., Caria, A., Veit, R., Gaber, T., Rota, G., Kuebler, A., Birbaumer, N.  

Sterr, A., Muller, M.M., Elbert, T., Rockstroh, B., Pantev, C., Taub, E.  

Strehl, U., Leins, U., Goth, G., Klinger, C., Hinterberger, T., Birbaumer, N.  

Weiskopf, N., Sitaram, R., Josephs, O., Veit, R., Scharnowski, F., Goebel, R., Birbaumer, N., Deichmann, R., Mathiak, K.  

Weiskopf, N., Mathiak, K., Bock, S.W., Scharnowski, F., Veit, R., Grodd, W., Goebel, R., Birbaumer, N.  
2004  Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE transactions on bio-medical engineering*, 51 (6), 966–70.

Weiskopf, N., Veit, R., Erb, M., Mathiak, K., Grodd, W., Goebel, R., Birbaumer, N.  


2005  Role of the contralateral inferior frontal gyrus in recovery of language function in poststroke aphasia: a combined repetitive transcranial magnetic stimulation and positron emission tomography study. *Stroke*, 36 (8), 1759–63.
1. Talent and a lifelong commitment to learning languages

In this book we investigated as objectively as possible the factors that contribute to what might be described as language talent. We identified some general biological factors (e.g. functional connectivity of brain areas while processing language) and cognitive factors (e.g. empathy and memory span), as well as very specific sociolinguistic (e.g. ability of convergence in dialogue) and phonetic factors (e.g. degree of coarticulatory resistance and prosodic variance) that contribute to the emergence of talent. We have shown that these factors positively correlate with language achievement, which we have tested and investigated on a population of over 100 German subjects learning English. We can safely assume that the subjects, in whom these factors correlate positively, are able to learn to speak foreign languages without an accent. We believe that having discovered these factors, we can actually point to many practical areas that might have to be considered more thoroughly in the practice of teaching foreign languages.

However, there is a more subjective aspect to this investigation. If you meet all the criteria for talent, how many languages can you actually learn to speak (without a foreign accent)? Many …, well let’s not be modest, most of the authors of this book speak more than two languages fluently. Some speak three without any accent, but for us this is about the limit. How do you get over this limit, how many languages can you actually learn to speak? The short biography of someone generally considered a “language genius” will hopefully convince you that with a bit of talent, a structured learning method, and a lifelong commitment to learning languages you can learn many, very many of them.

Max Mangold is known to most people interested in language as a highest authority on German pronunciation. He has been seminal in the codification of the German sound system and for decades has had single-handed responsibility for the German pronunciation Dictionary (Duden Aussprachewörterbuch). In the introductory part to this dictionary
Mangold gives a cursory description of the sound systems of 26 different languages (Duden 1962: 80–92). Few people know that this is not second-hand knowledge! Max Mangold actually speaks all these languages without a noticeable accent. In his prime, he actively spoke almost 40 languages. How can one achieve such proficiency and diversity? Here is his short language biography.

1.1. Childhood and youth

Max Mangold was born on the 8th of May 1922 in a small Swiss village Pratteln which sits close to the German border not far from the affluent town Basel. This is an area in which hard to understand Swiss German alemannic dialects are spoken. The high alemannic dialect version spoken in Pratteln (hochalemannisch) was a native language of his father. This dialect is very different from standard German; however, most of its speakers can understand standard German. His mother spoke highest alemannic (hochstalemannisch), which is not mutually understandable with German standard. Max Mangold’s childhood was not characterized by a multilingual situation. Actually, the only language variant that he was confronted with at home and at school was a highly substandard variety of Swiss German.

The social situation of his family was also not stimulating for the open multilingual education. His stay-at-home mother, just like most of the inhabitants of a typical Swiss village, never learned a standard language. His father, who worked as a lorry driver, drove around the alemannic area (north Switzerland and south-west Germany, south-east France) but never learned any other language than German.

Mangold’s first interest in languages was awoken by the road maps that his father used for his business. Young Max Mangold went through all the routes that his father drove and learned all the names of cities, rivers and other locations marked on the map by heart. He soon figured out that many areas could be characterized by common names that are used there, and that these names were pronounced in a very similar way. Learning place names became really interesting when he got his first world atlas. Learning foreign names without knowing how they are supposed to be pronounced became Max Mangold’s first real challenge at an otherwise unenthrobbing school. An interesting by-product of learning how to pronounce English, French, Indonesian, Japanese, Russian, and African place names correctly, was an apparently astounding knowledge of world geography that a grammar school boy from a small village ac-
Beyond talent: a short language biography of Prof. Max Mangold

By grade 4 he could pronounce all the names from the world atlas and locate them on the map. Max Mangold’s amazing “talent for geography” was noticed by a village doctor, who convinced the parents and the local authorities to support the young boy’s schooling in natural sciences. This is how the skill at pronouncing foreign names got Max Mangold a place in a prestigious Basel Science High School (Naturwissenschaftliches Gymnasium).

Basel confronted Max Mangold with another linguistic challenge. This affluent swiss city is a dialectal island of a prestigious low alemannic version of German (Niederalemannisch) in the ocean of strongly marked high alemannic variants. Mangold’s high alemannic pronunciation stigmatized him in the elite surrounding of his low alemannic speaking peers. The German teacher actually wanted him to be removed from school whose up-scale dialect he could not speak. It took Max Mangold three years to master the prestigious pronunciation of “Niederalemannisch”. In a kind of revenge on a snob teacher he also mastered standard German (Hochdeutsch), a pronunciation variant that fewest Swiss German speakers ever accomplish.

Mangold was much more successful with other languages offered at school. French, which was offered from the first grade on, he mastered very fast. The method used at his high school concentrated on teaching pronunciation for seven hours a week in the first grade – words and sentences were not important, only correct pronunciation of French sounds and French prosody were taught. At the end of the year Max Mangold had mastered French pronunciation and had spoken his first foreign language without an accent, but often not really knowing what he was talking about.

1.1.1. Language learning with the IPA

A live changing experience was the language learning method practiced in the English class. English was taught only at the sciences schools in Basel at that time, and it was allowed to teach it with very modern methods (Latin, which was the language of art schools, was taught very traditionally). Dr. Klemm, the English teacher in Basel at that time whose name Max Mangold still cherishes, was a strict follower of a than hardly known British phonetician Daniel Jones. Jones, one of the principal moving forces behind the International Phonetic Alphabet and the author of the English Pronouncing Dictionary, proposed a radical method of teaching foreign languages from the phonetic alphabet and the phonetic transcrip-
tion. This method was diligently practiced in the 4th grade of science schools in Switzerland in the 1930’s of the previous century.

The discovery of the International Phonetic Alphabet (IPA) at the age of 14 opened a totally new window on languages for Max Mangold. All the phoneme-grapheme conversions that he had to work out for himself in order to learn how to pronounce foreign words became obsolete. The Gaspé/Otto/Sauer German transcription system (used by Langenscheidt Publishing Company since 1880 in lexica for language learners and travelers) which he mastered before, was now accommodated by a much more versatile and universal IPA. Max Mangold became serious about learning languages – spoken languages.

During his first years in Basel he shared a room with others, but after a short time (at 11) he moved to his own room, in the attic, where he could practice languages without disturbing anyone. He learned French at 12, English at 14, Spanish at 15–16, Portuguese at 16 – all using the phonetic approach of Daniel Jones and with lexical support from the Langenscheidt dictionaries. The talent for languages was noticed by his teachers, who sent him to a Latin class for gifted pupils of all of Basel’s schools. Max Mangold was 15 and it took him a year to become the best pupil in this elite class. Latin remained the only language that he learned from a textbook.

His success with Latin made his teachers aware that his real talent was not in geography and other sciences but in languages. At the age of 16 he was allowed to switch from the sciences school to the arts school. The experience with Latin taught him another two things. First, you can learn languages not only by trying to speak them, but also by reading texts; second, professional instruction makes language learning easier. He took lessons in Greek (20 lessons that he actually paid for) and mastered this language at 16–17, also by reading classical Greek texts.

Late 1930’s were a time of great political turmoil and Mangold devoured all the foreign language newspapers in languages which he already spoke. His Spanish got much better due to Franko’s war. He learned Italian in the summer of 1938 which he spent in Verona and Rovigo giving language lessons in German and English in exchange for language lessons in Italian. He also learned Albanian, after Italy invaded Albania, and learned Polish after the German invasion of Poland in 1939. The choice of a new language to learn was driven by the current political situation.
1.1.2. Language learning with the radio

In 1940 Max Mangold bought his first short wave radio. He already had a good phonetic knowledge of most of the languages spoken in Europe, and with a short wave he could listen to all of them at will. The frequency band was very limited (16–49 m) but the time of the Second World War was the radio time. BBC World Service, Voice of America, Radio Moscow were broadcasting around the clock in all possible languages (mainly propaganda) and neutral Switzerland had no restrictions on the choice of broadcasters you wanted to listen to. Max Mangold developed a habit of listening to languages on the short wave radio. He has not given up this habit since then. He possesses the most impressive collection of short wave radios of the last century. His flat is dominated by a huge-size satellite aerial (satellite radio of the 1990’s), and nowadays he obviously uses a modern hi-fi internet radio receiver.

Since 1940 until today Max Mangold has listened to foreign languages over the radio every day. Some days he has listened to as many as 20 languages a day. On other days he has followed individual languages and individual broadcasters. He has concentrated on individual words and expressions or on individual pronunciations and pronunciation variations. Short wave broadcasters repeat their news very frequently and for a language learner these “old news” are very good news, because the same words and expressions are repeated over and over again. If announcers change in the course of the day, you can experience different pronunciation variants and even different dialects.

Over the years of listening to short wave Max Mangold established a ranking of the major broadcasters. Vatican Radio and Radio Moscow have been good for beginners, because their speakers have been instructed to use the slow familiar speaking style. BBC speakers all use clear speech, but they are substantially faster and use the rapid familiar style. The speakers of the Voice of America use rapid colloquial speech, which is useful for more proficient foreign language speakers. In the 1940’s Max Mangold learned a great number of languages over the radio – he just listened, transcribed, checked the meanings and pronunciations in his Langenscheidt lexica, and tried to repeat what he has heard. This non-communicative method of learning languages from the radio sharpened his pronunciation skills in most of the Germanic languages (he could speak English, Dutch, Swedish, Danish and Norwegian); Romance languages (Romanian, Italian, French, Spanish, Portuguese); Slavic languages (Polish, Czech, Russian, Bulgarian, Serbo-Croatian –
he never learned Slovenian, which he considers the “toughest” of all Slavic languages) and a couple of “odd” ones (Hungarian, Arabic, Persian, Chinese, Japanese, and Korean). In 10 of those languages he was really fluent and accent free! He was finishing school then, and was not really putting his knowledge of those languages into practical use. This was going to change during his academic and professional life.

1.2. Academic and professional life

Max Mangold joined the university for the first time in order to learn Russian. He still had one year before his final exams (Abitur) but was so fascinated by the language and the political situation in Europe (Russia was fiercely fighting the Nazi Germany) that he signed up for one semester of Russian. The class had 4 students and all of them were as strongly motivated to learn the language as Max Mangold was. At the end of the semester all of them were quite fluent in Russian. Mangold persuaded them to listen to Radio Moscow every day and to try to repeat everything they heard. They learned many of the Russian victory news and speeches, and picked up an accent of the main soviet propaganda announcer – Yuri Levitan. By the way, it was Levitan’s and not Stalin’s voice, which led Russia through the great patriotic war against Germany – the reason was purely phonetic – Stalin spoke Russian with a strong Georgian accent.

Max Mangold studied philology (general and classical) at the Universities of Basel (a German speaking university) and Geneva (a French speaking university). He was not really too much excited about the formal classes, but used the international and open atmosphere of the academia to practice languages in real communication. During the Second World War Switzerland attracted a large number of foreign students, who were lucky to escape the cruelties of the political madness. Max Mangold was hanging around with them most of the time and talked to them in their languages. His habit of listening to the radio made him one of the most thoroughly and universally informed students on campus, which also raised his attractivity to foreigners who wanted to have the most recent information about their home countries. Close contacts with foreign students made the Swiss secret police suspicious and, for a short time, Mangold’s contacts were closely monitored. Somebody who spoke standard German, perfect Russian and contacted every foreigner at the university was obviously very suspicious at that time. The fact that the foreign friends were exclusively male got Mangold to other types of problems. Those problems disappeared as “the gossip” realized that there
were practically no female foreign students in Switzerland at that time. Max Mangold left the Swiss academic system with a founded knowledge of over 30 languages and excellent communicative skills in almost 15 of them.

In 1945 he was accepted as a graduate student to the famous Institute of Phonetics of the University of Paris. Pierre Fouché, a famous experimental phonetician of the time, was to become his teacher. Fouché, whose expertise was in the descriptive and in the physiological phonetics, never made a real impact on Max Mangold. Fouché was Catalan by origin and never managed to learn to speak French without an accent. Fouché’s saying “a strong character needs only one accent” definitely did not win him Mangold’s respect.

After three years in Paris Max Mangold joined the phonetic department of the University College London. There he studied phonetics under the supervision of Gimson, Fry, and O’Connor, all of whom were pupils and strict followers of Daniel Jones. He also attended classes by Firth, Henderson, and Pregg at SOAS (School of Oriental and Asian Studies), where he got interested in further languages (mainly Siamese and Burmese). This awoke his interest in Chinese and Korean that he studied already before. He learned to speak, read and write Chinese really well, and has been using this language for the rest of his life.

The knowledge of practically all European and numerous Asian languages paired with a Swiss passport got him his only job outside of the academia. In 1953–54 he worked for the United Nations as an interpreter during the Korean conflict. The languages he interpreted in were French, German, English, Russian, Polish, Czech, Slovak, Swedish, and Chinese.

Back in the academia he worked as a lecturer and assistant professor at the universities of Basel, Zürich and Bonn, until he was appointed to full professorship of phonetics at the University of Saarland in Saarbrücken in 1957. As a professor he divided his interests into teaching of descriptive phonetics and to the study and mastery of further languages. His seminal contribution to the codification of the German phonetic transcription has already been mentioned. He also supervised almost 100 scientific theses and dissertations – many of which were the first and only descriptions of endangered languages. In the late 70’s and early 80’s foreign students who wanted to describe their endangered native language (most of them from Asia and Africa) were routinely sent to Saarbrücken by the German Academic Exchange Foundation (DAAD). Prof. Mangold was also substantial in the phonetic description of many local dialects of the Saarland area (Bonner et al. 1993).
However, all these scientific achievements were bleached by the personal contact. Prof. Mangold’s incredible knowledge of languages was apparent in every conversation and filled his colleagues and students with awe. Not only could he speak with almost anybody in the person’s native language, he also knew many details about the language’s dialects and areal variants. If your native language happened to be spoken somewhere in Europe, Mangold could reconstruct your language biography from the accent you had with great scrutiny and precision. Sometimes it was difficult to figure out if Prof. Mangold’s interest lay in the topic of the conversation or in the interlocutor’s pronunciation and language biography. Anyway, you can be sure that he is very attentive to what you are saying and to how you are saying it!

In preparing this biography I met Prof. Mangold in a Chinese restaurant in Saarbrücken. It was about a week before his 86th birthday. He ordered his meal in Chinese and talked for a while with a waitress about the Chinese transliterations on the menu. He then turned to me and asked if I understood their conversation (which was in Chinese). When I denied, he explained to me that the restaurant was apparently run by a Chinese from Shanxi (he figured it out from the transliterations) but the waiter was also speaking Mandarin: “as you know the dialects of Shanxi and Mandarin differ in tone-sandhi and glottalization.”

References

Bonner, M., Braun, E. & Fix, H.

Mangold, M.
1990
absolute pitch 221, 223
accommodation 257–260, 267, 269, 270, 274, 315, 326
acoustic analysis 279, 280
acquired accent 194
acted speech 25, 34
adaptation 200, 257, 258, 263, 299
aetiology 202
affective prosody 343, 344
age of onset of learning (AOL)/acquisition (AOA) 1, 5, 8–10, 17, 18, 78, 115, 156, 164, 165, 166, 169, 171
agrammatism 196
agraphia 131
agreeableness 20, 106, 115, 117, 120
alexia 131
alignment 258, 263, 274, 285, 312–314, 328
allophonic variant 292
allophonic variation(s) 22, 27, 279, 299
alpha rhythm 166, 167
amount of language use 1, 8, 17
amplitude contours 259
amplitude envelopes 268, 269, 274
amusia 239
angular gyrus 131–133, 141, 142, 172
anterior cingulate cortex (ACC) 76, 112, 113, 121, 136, 163, 175, 341
anxiety 71, 100, 103, 107–111, 114, 115, 118, 162, 217
aphasia 131, 157, 159, 160, 193, 196, 201, 239, 343
aphasic deficits 196, 200
approximation strategy 258
apraxia of speech 196, 200, 201
aptitude test 4, 11, 21, 72, 88, 225, 227, 228, 236, 237
aptitude test: 4, 11, 21, 72, 88, 225, 227, 228, 236, 237
arcuate fasciculus 131, 140, 145
articulation 25, 81, 82, 120, 133, 197, 280–284, 296, 339
articulatory undershoot 197, 201
associative memory 85, 237
attention deficit/hyperactivity disorder 80, 340
attentional processing 343
audiation/audiate 225, 227–229
auditory cortex 6, 138, 139, 142, 173, 174, 231, 243–245
autosegmental metrical phonology 312
avoidance 1, 25, 310
backward repetition 82, 83
basal ganglia 133, 136, 137, 142, 176, 207, 209–211
basic variety 309
behavioral inhibition/activation system (BIS/BAS) 20, 109–112, 116–120
belt region (auditory cortex) 139
benzodiazepine 70, 71
bilingual/polyglot aphasia 158, 159, 160
bilingualism 74, 76, 78, 155–157, 163, 170
bilinguals 74, 78, 155–157, 159, 161–165, 170, 175, 339
biofeedback 340
blood oxygen level-dependent (BOLD) signal 340
bodily-kinaesthetic intelligence 85
BOLD 134, 136, 137, 340
boundary tone 27, 312–314, 319, 326
brain damage 159, 194, 195, 201, 232, 338
brain plasticity  73, 78, 164, 338–340, 343
Broca  82, 131–133, 137, 142, 143, 145, 159, 163, 164, 174, 175, 207, 240, 241, 242
Brodmann (areas)  132, 242, 244
Carroll (John B.)  11, 21, 217, 218, 236
category transfer  23
caudate (nucleus)  207, 208, 211
center of gravity  292
central executive  79
cerebellum  82, 133, 136–138, 141, 209
cerebrovascular event  193, 201
chord(s)  216, 218, 224, 225, 242, 243
chronic pain  341
circumstances of language acquisition  17–19
cognitive control  74–79, 112, 113, 163, cognitive interference  81, 108, 111
cognitive neuroscience  142, 155, 177, 178, 337
cohesiveness  128
communication  1, 24, 68, 74, 85, 86, 107, 137, 243, 356
communication accommodation theory (CAT)  257
composing/composition  213, 217, 220, 223, 226, 229
connectivity  132–134, 144, 145, 179, 351
conscientiousness  20, 40, 107, 115, 117, 120
constriction  282, 283, 295
convergence  35, 257–260, 266–268, 270–271
core region (auditory cortex)  139
Corsi’s block tapping test  79
cortical efficiency (theory)  166, 170, 176, 177
cortical reorganization  338
creative abilities  216, 217
critical period  5, 6, 10, 169, 172
critical period hypothesis (CPH)  5, 6
degree of social distance  257
delayed imitation  32, 33, 53, 62
detention interval  81
Developmental music aptitude stage  226–228
dialectal  195, 284, 291, 298, 353
Diapix (dialog picture matching)  267, 268, 272
direct imitation  31, 32, 50, 54, 62
discourse  23, 25, 27, 28, 30, 31, 38, 39, 229, 257
divergence  258–260
DNA  157, 162, 167, 168
Don Quixote effect  69
dorsal stream  143, 144
dorsolateral prefrontal cortex  76
DTI (Diffusion Tensor Imaging)  134, 135, 144, 145, 173
dualism  162
dual-routes processing  75
duration  27, 28, 218, 243, 259, 268, 299, 309
DWI (Diffusion Weighted Imaging)  134
dysprosody  193
EEG-biofeedback  340, 342
ego boundary  102, 114, 120
ego-permeability  4, 69, 70
E-IQ  71, 72
electroencephalogram (EEG)  112, 113, 134, 144, 155, 165–167, 235, 340, 342
electroencephalographic (EEG) signal  340
Emil Krebs  174
emotional intelligence  68, 70, 108
emotional resonance  68
empathetic capacity  68, 102, 103, 114, 118, 120
empathy 1, 4, 11, 20, 40, 68–73, 116, 118, 120, 162, 351
encoding 81–83, 142, 246, 262, 264
enhancement 4, 337, 338, 344
environmental enrichment 338
E-Prime software 76
E-Scale 71, 72, 118, 120
ESOL 8, 11
evaluation methods 34
executive functions 75, 80, 163
exemplar theory 260, 261, 263, 265, 266, 274, 275, 299, 306–308, 329
experience (linguistic) 2, 8, 9, 18, 19, 25, 30, 35, 36, 41, 162, 167, 171, 195, 260–264, 297
experience (neurological) 76, 171, 338, 339
expert analysis 34, 35, 37, 38
expert knowledge 8, 9
extraversion / extroversion 1, 9, 11, 17, 20, 85, 97, 98, 100, 105–107, 110–112, 114, 115, 117, 120, 162
F₀ contour 29, 38, 198, 268, 314
F₀ curve 322, 323, 328, 329
F₀ peak 320–322, 327, 330
FAS, Foreign Accent Syndrome 193, 194, 201
fiber tracts 134, 144, 145, 231
fine phonetic detail 260, 267
Five Factor Model (“Big Five”) 100, 106, 107, 115, 116
fMRI-biofeedback 342
foreign accent(s) 5, 7, 8, 22, 23, 27, 36, 69, 73, 193–195, 199, 201, 207–211, 279, 283, 396
formal learning 10
formant 197, 280, 282–285, 288, 290, 292, 293, 295, 296, 298
fortications 197
forward repetition 82, 83
fractional anisotropy 145
frequency effects 262, 264, 265, 308
frontal operculum 142, 209, 243
fronto-temporal 145
functional magnetic resonance imaging (fMRI) 112, 115, 116, 134–137, 139, 142, 144, 145, 155, 165, 175, 176, 238, 242, 244, 338, 340–344
fundamental difference hypothesis 5
Gardner’s multiple intelligences 85, 86, 215
general factor “g” 85
general intelligence 81, 85–87, 214, 217, 221, 233, 234, 245, 246
generic 199
Geschwind 2, 132, 133, 141, 144, 145, 177
Gordon (Edwin E.) 20, 215, 219, 225–229, 234, 235, 237, 244–246
Gordon’s advanced measures of music audiation (AMMA) 228–230, 235, 237, 245
gradient phenomena 261, 266
grammar vs. accent 3, 4
grammaticality judgments 4, 22, 343
grey matter 78, 245, 339
handedness 75, 162, 167, 168, 177
harmony 216, 218, 220, 224–227, 229, 241

Index 361
Heschl’s gyrus 133, 138, 139, 173, 174, 242, 244, 245
hippocampus 175, 176, 340
imitation 7, 17, 23, 31–35, 50, 88, 234, 260
improvise/improvisation 216, 217, 225, 226, 228, 229, 246
individual differences 6, 67, 69, 102, 111, 112, 114, 117, 157, 158, 168, 172, 173, 175–177, 217, 220, 266, 288
inductive reasoning 85
ingroup/outgroup (boundaries) 257
instrument playing 72, 86, 213, 214, 216, 223, 226, 227, 229, 233, 235–237, 244
insula 137, 142, 163, 173, 175, 209, 210, 242
intelligence (general) 1, 9, 17, 20, 67, 74, 81, 84–88, 97, 109, 113, 119, 162, 163, 214–218, 221, 233, 234, 245, 246
intelligibility 23, 195, 201
intensity 36, 136, 197, 217, 218, 220–222, 224, 233, 246
interlanguage 309, 310
interpersonal intelligence 85
intonation contour 32, 135, 198, 314
IPA (international phonetic alphabet) 25, 292, 353, 354
Joseph Conrad phenomenon 3, 178, 267
Kainz’ law 159
kanji + kana 143
kinesthetic abilities 216
L2 learning 5, 7, 10, 155, 167, 344
L2 proficiency 77, 83, 113, 168, 234
language acquisition 1, 4–12, 17, 41, 97, 121, 155, 161, 167, 175, 177, 230, 259, 260, 266, 305, 307, 309, 311, 344
language anxiety 107, 108, 109, 111, 114, 118
language aptitude 1, 4, 11, 17, 21, 72, 88, 97, 98, 101, 102, 114, 162, 163, 171–178, 213–215, 234, 236, 237, 266
language ego 4, 69, 102
language network 173, 343
language processing 7, 22, 132, 135, 144, 161, 166–170, 173, 232, 238, 240, 242, 244, 246, 285, 311, 341
lateralization 133, 137, 139, 145, 165, 168, 169, 173, 177, 238, 242
lateralized readiness potential (LRP) 75
left hemisphere 131, 132, 136, 138, 144, 145, 159, 168, 174, 196, 207–211, 231, 232, 244
left inferior parietal cortex/lobe 78, 82, 175, 339
Left prefrontal cortex (PFC) 76
lesion site 158, 195, 202
less-talented 268–270
linguistic prosody 198, 201
locus of control expectancy 100, 104
logical-mathematical intelligence 86
long-term encoding 81
long-term memory 79, 81, 139
loudness 27, 197, 218, 219, 224, 233
maintenance 79, 81, 82, 140, 259
match value 269–274
mean coarticulatory distance 292
INDEX

MEG (magnetoencephalogram) 134, 144, 155, 238, 242–245
Mehrfachwahl-Wortschatz-Intelligenztest (VMLT) 20, 86, 88
memory sequence 265, 266
memory span 20, 79, 351
mental flexibility 20, 67, 73, 74, 78, 97, 339
metre 230, 240
micro-momentary expression device (MME) 70
mother-tongue accent 344
motivation 1, 2, 8, 9, 10, 11, 17, 18, 35, 40, 41, 104, 111, 162, 163, 266
motor cortex 6, 133, 136, 137, 140, 143, 195, 196, 210, 211
motor theory of speech 161
multilingualism 155
music achievement 225–227
music training 177, 230, 231, 244
music(al) ability(ies) 177, 213, 217, 220–225, 230, 231, 236, 238, 244–246
music(al) talent 86, 177, 213–215, 218–220, 223, 246
musical expertise 232, 339
musical intelligence 213, 222, 245
musical skills 232, 233, 235, 244, 247
musical syntax 217, 240–242
musicality 20, 213–219, 225, 235–238, 244–247
myelination 172
native language magnet theory 7, 8, 308
native-like 4, 5, 6, 10, 11, 18, 67, 69, 80, 102, 285, 297, 314
naturalistic intelligence 85, 86
nervous system 111, 114, 137, 338, 339
neural-processing efficiency 113, 114
neurofeedback 5
neuroimaging 17, 41, 79, 82, 113, 141, 143, 155, 171, 238, 241, 340
neurophysiological aspects of language acquisition 4, 160
neuropsychological models of talent 3, 7, 170, 241
neuroticism 20, 100, 106, 111, 112, 115, 117, 120
nonverbal features 258
nonverbal intelligence (IQ) 9, 20, 67, 74, 84–88, 109, 113, 115, 119, 162, 163
nucleus (plural nuclei) 313, 319, 320, 322, 328, 330–332
openness / openness to experience 20, 98, 106, 107, 115, 117, 120
operant conditioning 340, 341
Pair comparisons 27–30, 39, 40, 55–59
perception-production link/loop 259, 263, 264
perceptual assimilation model 7
perceptual correctness 85
perceptual distinctiveness 299
perceptual evaluation 35–37
personality trait 9, 17, 20, 100, 102, 105, 107, 109, 112, 115, 119, 120, 217, 226, 258, 266
personality type 100, 102, 105
person-situation debate 100, 107, 118
PET (positron emission tomography) 134, 155, 174, 238
phonetic feature(s) 7, 20, 73, 197, 199, 297
phonological loop 79, 80–82
phonological processes 279
phonological rules 299
phonological store 80, 83, 140
phonological working memory 7, 20, 140, 163, 175, 177
phrasing 46–48, 218, 224, 227, 229
pitch level 320, 322, 327
Pitres’ law 159
plasticity (cortical, brain) 73, 76, 78, 162, 164, 169, 239, 338–344
polyglot 83, 155, 157, 162–164, 194
polyglot aphasia 160
prefrontal cortex (PFC) 76, 111, 113, 132, 163, 166, 173, 243
premotor cortices 82
primary auditory cortex 138, 139, 244, 245
primary motor cortex 133, 195, 196
priming effect 75, 243
proficiency level(s) 17, 157, 158, 161, 164, 165, 168, 169, 170, 171, 285, 286, 290, 293, 296
pronunciation talent 1, 7, 9, 17, 18, 73, 88, 97, 114, 119, 120, 121, 236, 237, 246, 268, 299, 305, 310
prosodic skills 23
prosody interpretation 30, 49
prototype (theory) 308
pseudoaccent 193
psychiatric 194, 195
putamen 137, 207, 208, 210, 211

Raven’s advanced progressive matrices 20, 86, 87, 88, 109
read speech 25, 26, 46–49, 271, 274
realizational differences 27, 29, 31, 39, 54, 55, 57, 313
real-time fMRI (rtFMRI) 340
receiver role 257
recovery patterns 157–159, 193, 200, 338, 343
rehabilitation 339
rehearsal 80–83
retrieval 81, 139, 141, 142, 218
reverted accent 194
rhythmic/rhythm 20, 86, 199, 216–242
Ribot’s law 158
right hemisphere 82, 132, 134, 139, 143, 166, 168, 177, 232, 242–244
right inferior frontal gyrus (rIFG) 341
rostral part of the anterior cingulate cortex (rACC) 341
scoring/scorecard 34–40, 45–63
Seashore (Carl E.) 218, 219–222, 231–235, 246
Seashore’s measures of musical talents 218, 219, 222, 231, 233–235, 246
second language acquisition (SLA) 1, 4–7, 10, 12, 17, 97–99, 101–105, 107, 108–111, 113–115, 118, 120, 121, 155, 167, 175, 177, 259, 266, 305, 309, 311
segmental features 23, 26, 46, 196, 245
segmental level 5, 17, 31, 197
selective attention 74–76
self-awareness 86
self-efficacy 104, 114
sensitive period 5, 171
short-term memory 79, 81, 234
short-term storage 78
Shuter-Dyson (Rosamund) 216–219
Simon task (effect) 20, 74–78
singing 138, 139, 213, 216, 232, 235, 237, 238, 244, 246
skill enhancement 338
slow cortical potentials (SCP) 341
social psychology 257
sociolinguistic 1, 257, 266, 351
socio-psychological aspects of language acquisition 4, 9, 12
sound shift 194
spatial intelligence 86
Index 365

spatial visualization 85
speaking rate 22, 24, 36, 259, 268, 274,
Spearman, Charles & Spearman’s “g” 84–87, 218
spectrograms 279, 280, 290, 292
speech disorder 193, 201
speech learning model 7
speech output motor system 195, 196
speech perception 26, 27, 30, 133, 138, 140, 144, 232, 234, 259, 261, 308, 338
speech production 6, 22, 23, 25, 26, 67, 137, 138, 143, 176, 196, 198, 201, 258, 263, 297, 309, 344
speech timing 201
speech variety 259
speech/speaking style 25, 258, 259, 266, 269, 274, 299, 355
stabilized music aptitude stage 226–228
STG/STS (superior temporal gyrus/sulcus) 131–133, 138–144, 175, 176, 242–244
strength/activation level 262, 263, 265
stroke 158, 193, 208, 337
style 3, 25, 105, 109, 117, 162, 163, 227, 229, 258, 259, 266, 269, 274, 299, 355
sub-vocal rehearsal 82
sub-vocal repetition 81
supramarginal gyrus 82, 133, 139–142, 208, 210, 242
suprasegmental features 23, 27
suprasegmental level 5, 17, 31, 196, 198
suprasegmental structure 198, 230, 299
sylvian fissure/cortex 132, 137, 142, 143, 196, 207, 208, 211, 244
synchronization 165–167
talent (definition) 2, 3, 213, 214
talent vs. proficiency 2, 8, 19, 34, 35, 41
TBI, traumatic brain injury 193, 201
temporal cortex/lobe 131–133, 138–143, 163, 168, 173, 175, 176, 208–212, 242–244
temporoparietal 142, 143, 159, 166
test subjects 9, 18, 19, 35, 36
timbre 213, 218–222, 230, 233, 240, 246
ToBI 27, 57, 63, 312–317, 319, 323
TOEFL 8, 11, 22, 236, 237, 238
tonal imagery 218, 229
tonal memory 219, 233, 234
tongue dorsum activity 283, 298, 299
tractography 134, 144, 145
transfer 23, 281, 310, 313, 315, 326–329
translation disorder 158
types of tasks 22
up-regulation 343
usage-based account (of language) 260, 261
ventral stream 143, 144
ventrolateral frontal cortex 82
verbal comprehension 85
verbal intelligence (IQ) 6, 11, 20, 86–88, 115, 119, 162, 163
visual analogue scale 36, 37
visual word form area 131, 132, 141
visuo-spatial sketchpad 79
vocabulary learning 72, 88, 178, 236–238
voice onset time (VOT) 7, 259
voicing error(s) 197, 201
vowel quality 197, 199
Wechsler adult intelligence scale  82
Wechsler digit span test  82
Wernicke  131–133, 138, 141–145, 241, 242
white matter  131, 134, 138, 145, 172, 173, 207–212
Wing (Herbert)  219, 221, 222–225, 234, 245, 246

Wing’s standardised tests of musical intelligence  219, 222–224, 234, 246
word fluency  85
word frequency  261
word list  230, 268, 269, 271, 274
working memory  7, 9, 11, 17, 20, 67, 78, 82, 97, 112–114, 120, 136, 140, 162, 163, 174–177