A Database for the Analysis of Cross-Lingual Pronunciation Variants of European City Names

Stefan Schaden

Institut für Kommunikationsakustik (IKA)
Ruhr-Universität Bochum
D-44780 Bochum, Germany
schaden@ika.ruhr-uni-bochum.de

Abstract

This paper reports on a speech database that includes non-native pronunciation variants of city names/town names from several European languages. The database is designed as a research tool for the study of pronunciation variants in this specific domain that occur in different groups of non-native speakers. The ongoing data collection currently comprises 20 to 27 native speakers of 3 languages each who pronounce material from 5 languages. The languages covered are English, German, French, Italian, and Dutch. All languages are examined as the source language (L1) and as the target language (L2). For the first stage of the data collection, the targeted status is a collection of 5 x 5 language directions with at least 20 speakers per native language.

After a brief overview of related studies and an outline of some specifics of proper names (place names in particular) in the context of speech technology applications, the database design and the current stage of the data collection is described.

1. Introduction

In speech technology applications such as voice-driven travel information services or car navigation systems, the pronunciation of place names – e.g. city names, street names, railway stations, 'points of interest' – plays a central and often critical role. Like proper names in general, place names show a number of irregularities with regard to their orthographic and phonological structure and pronunciation. For instance, it was found in the development of grapheme-to-phoneme converters that in many languages, proper names are not always consistent with the regular grapheme-to-phoneme correspondences and therefore require specialised rules (see Belhoula, 1996; Gustafson, 1994).

A rather new problem related to proper names arises when Automatic Speech Recognition (ASR) systems are used in cross-lingual (i.e. international) environments that involve non-native speakers as potential users. Even if the recogniser itself is multilingual (see e.g. Schultz & Waibel, 2001), allowing to adapt to the speaker's native language by selecting the appropriate acoustic models, language models and vocabularies, proper names remain a critical issue as they are not 'translatable' in the same sense the standard vocabulary is. Thus, non-native speakers who are not aware of the correct pronunciation will use their individual knowlegde of the target language to pronounce the foreign names, producing a variety of non-standard variants that may differ significantly from the canonical transcriptions provided in the recogniser's pronunciation dictionary. Although present-day recognisers can handle a great deal of allophonic variation through statistical modeling by HMMs, non-native variants frequently involve uncommon phoneme realisations and even phoneme shifts that were not covered in the recogniser's training material, thus leading to a significant decrease in recognition performance.

One possible approach to handle this problem is *lexical adaptation*, i.e. the inclusion of multiple or alternative pronunciations that match the most characteristic non-native realisations in the recogniser's dictionary (for an overview of lexical adaption techniques for intra-lingual variation see Strik, 2001). For a standard vocabulary recognition task including Italian and German speakers who pronounced English material, Goronzy et al. (2001) could improve the recognition rate significantly by adding non-native pronunciation variants to the lexicon.

Lexicon adaptation, however, requires detailed knowledge on the potential pronunciation variants occuring in different non-native speaker groups: While in the case of intra-lingual or dialectal variation this knowlegde may partly be drawn from the linguistic literature or be inferred from existing monolingual databases, the situation is not as straightforward for non-native pronunciation variants. A small number of non-native speech databases do exist, e.g.

- the *Translanguage English Database* (Lamel et al., 1994),
- the *Strange Corpus* I & II (distributed by the Bavarian Archive for Speech Signals – BAS[1]),
- the *ISLE Corpus* (Menzel et al., 2000),
- the *NATO Native and Non-Native (N4) Speech Corpus* (Benarousse et al., 2001)

However, none of them explicitly covers place names. Moreover, the existing databases are limited to one target language (English; German for the *Strange Corpus*). An exception – although not in the domain of spoken language – is the *Onomastica* project, which is briefly sketched in section 3 below.

2. Purpose of the database

In order to address the situation outlined above, a cross-lingual speech database (SDB) for the analysis of non-native pronunciation variants of European place names is currently built up at IKA. It is designed to provide initial answers to the following questions:

- Which processes of linguistic transfer from the native language to the target language are typically involved in the pronunciation of foreign place names?
- What is the particular influence of *spelling pronunciation*, i.e. of the transfer of native grapheme-to-phoneme correspondences onto the target language?

[1] http://www.phonetik.uni-muenchen.de/Bas/

- Are the observed pronunciation variants primarily based on individual knowlegde and proficiency, or is it possible to detect 'typical' variants for speakers with a common native language background?
- Is it feasible to derive non-standard pronunciation lexica for particular groups of non-native speakers on the basis of their linguistic origin and/or foreign language proficiency?
- Is it possible to derive phonological rules to convert standard pronunciation lexica into adapted lexica for specific non-native speaker groups?

It should be stressed that the database is not designed as a large-scale speech resource that can be applied e.g. for speech recogniser training or the extraction of units for a concatenative speech synthesis system. Rather, it should be viewed as a research tool for the phonetic/phonological analysis of particular linguistic variables and their realisations by different groups of non-native speakers. Yet it is expected that the data collection and analysis will provide useful insights for the development of adaptation methods for speech technology. Moreover, the SDB can be used to test the robustness of speech recognisers against non-native speech.

3. Related work

A number of studies related to the pronunciation of foreign proper names have been carried out during the last ten years. The following section provides a brief outline of two studies/projects which are most closely related to the present study. Other studies – not presented in detail – dealing with non-native pronuncation of names in the context of speech technology are Cremelie & ten Bosch (2001) on the combination of several grapheme-to-phoneme converters for the recognition of foreign names, van den Heuvel (1994) on generating "Dutchised" pronunciations of foreign names using a Dutch grapheme-to-phoneme converter, Trancoso et al. (1999) on foreign place names in car navigation queries, and Eklund & Lindström (2001) on the use of "Xenophones" (foreign speech sounds) in Swedish.

(i) The aim of the European *Onomastica* cooperation project (Trancoso & Viana, 1995; Gustafson, 1994) was to produce a large-scale collection of machine-readable pronunciation dictionaries of proper names (city names and person names) for 11 European languages. This huge data collection (up to one million names per language) is supplemented by the so-called *Onomastica Interlanguage Pronunciation Lexicon* (IL) (Onomastica Consortium, 1995). This lexicon contains a subset of 1.000 words per language, where for each word so-called *interlanguage* pronuncations, i.e. potential non-native pronuncation variants are given according to the pronunciation rules of the remaining languages. The IL lexicon thus contains a matrix of a total of 121.000 entries.

Although the basic idea of the SDB presented here is quite comparable to the Onomastica IL, there are clear differences in focus and scope: Since non-native pronunciations strongly depend on particular speaker characteristics such as proficiency in the target language, age, education, etc., it appears to be a plausible hypothesis that there will be a considerable amount of variation *within* one speaker group of a common native language background. Thus it is rather unlikely that only one potential pronunciation variant per word for each native speaker group will represent this group appropriately. Therefore, it was decided in the present study to strictly limit the number of names covered by the SDB in favour of a broader coverage of potential inter-speaker variation.

(ii) While the Onomastica IL contained hypothesised transcriptions, Fitt (1995, 1998) collected actual speech data from the place names domain. She examined the pronunciation of city names taken from six European languages by speakers of Scottish English. It is one important observation in Fitt's work that speakers rarely apply the unmodified spelling pronunciation of their native language to pronounce foreign names. Even speakers with no formal background in the target language seem to be aware of at least some of the pronunciation regularities of these languages (some remarks by Trancoso et al. (1999) point to the same direction). This implies that a transcription according to the unmodified rules of the speaker's native language will probably not represent actual speaker behavior very well – even though this approach proved to be useful in speech recognition applications (Cremelie & ten Bosch 2001). Thus further investigations on the partial knowledge that speakers apply are necessary in order to model non-native variants appropriately.

4. Some general linguistic properties of proper names

Before the database itself is described in the following sections, it might be useful to point to some linguistic and pragmatic properties of place names (most of which hold for proper names in general) that distinguish them from 'standard vocabulary' and that may become particularly relevant in speech technology contexts involving non-native speakers.

First, place names are a word class that may be used even by speakers without any knowledge of the target language: While e.g. native speakers of French with no knowledge of German would generally avoid the use of German vocabulary, they may have to use it when dealing with proper names, e.g. in a travel information, route description or navigation query context. In these cases, severe deviations from the canonical pronunciations are likely to occur, while in 'standard vocabulary' the deviations will be less prominent as speakers usually have at least some basic knowlegde of the target language.

Secondly, even for speakers with an average knowledge of the target language, place names may turn out to be problematic due to their morphological intransparency: Place names frequently contain morphological constituents which are no longer part of the productive vocabulary of the target language ('frozen elements'). These elements are not stored in the speaker's 'mental lexicon' (i.e. they are not part of the learned vocabulary of the target language), and hence their pronunciation must completely be constructed by rule or by analogy.

Thirdly, place names frequently preserve orthographic features of some older historical stage of a language or contain traces of other languages the target language has come into contact with during its historical development. As a result, there may be uncommon grapheme-phoneme correspondences that deviate from the present-day pronunciation rules. Therefore, it has become necessary to include specialised rules for the grapheme-to-phoneme conversion of proper names, as mentioned above; and for the very same reasons, the pronuncation of place names is likely to raise difficulties for non-native speakers.

5. Database description

5.1 Languages, language pairs, and language directions

The targeted data collection will cover 5 languages: English (EN), French (FR), German (GER), Italian (IT), and Dutch (DT). Currently (Apr. 2002), the recordings of 20 native speakers of each German, English, and French are completed; recordings of Italian speakers are currently being conducted, and recordings of native Dutch speakers are scheduled for the near future.

Each language is examined both as the speakers' native language (L1) and as the target language (L2), i.e. native speakers of each language pronounce material from each target language. Thus, the database will cover 10 *language pairs* and 20 (5^2 - 5) *language directions*:

Language pair		Direction 1		Direction 2	
1	EN / FR	1	EN → FR	2	FR → EN
2	EN / GER	3	EN → GER	4	GER → EN
3	EN / DT	5	EN → DT	6	DT → EN
4	EN / IT	7	EN → IT	8	IT → EN
5	GER / FR	9	GER → FR	10	FR → GER
6	GER / IT	11	GER → IT	12	IT → GER
7	GER / DT	13	GER → DT	14	DT → GER
8	FR / IT	15	FR → IT	16	IT → FR
9	FR / DT	17	FR → DT	18	DT → FR
10	DT / IT	19	DT → IT	20	IT → DT

Table 1: Language pairs and directions

Since processes of phonetic transfer and interference are always language-specific (as they depend to a large extent – though not exclusively – on the interrelations of the phoneme systems of L1 and L2), each language direction must be considered as a separate unit of analysis. With any language pair added, the SDB will grow exponentially: Thus even the limited number of 5 languages yields as many as 20 units of analysis. Currently our own analysis of the data is restricted to the three languages GER, EN, and FR in different combinations.

5.2 Language material, experimental design, and types of speech

5.2.1 Selection of names for reading lists

The language material for the central part of the data collection consists of lists of 45 city names (including 2 duplicates, see 5.2.3 below) from each country – Germany, France, Great Britain, Italy, and The Netherlands. When compiling the lists, the following selection criteria were used as a guideline:

- *Coverage of 'critical elements'*. The majority of selected items contain orthographic or phonetic elements that were considered to be 'critical' for non-native speakers. Examples are: <th> clusters representing the interdental fricatives /D/ and /T/ [2] in British names (e.g. *Appleton Thorn*), nasalised vowels and orthographic vowel clusters in French names (e.g. *Langeais*), or complex consonant clusters and 'Umlaute' in German names (e.g. *Fünfstetten*). Although these 'critical features' may vary according to the speakers' L1, the same material was used for all native speaker groups. This will allow a comparison of the performance of speakers of different native languages. Note that it was not attempted to obtain a phonetically balanced corpus. While phonetic balance is a crucial feature e.g. of ASR training databases, it is not necessarily a useful criterion for the type of SDB presented here, since it is likely that especially the *less* frequent phonemes (which will be underrepresented in phonetically balanced sets) constitute problem cases for non-native speakers.
- *Focus on less familiar city names*. For some city names, standardised *exonyms* exist, i.e. variants of foreign place names which are lexicalised by a phonetically and orthographically nativised form (e.g. Engl. *Cologne* for Germ. *Köln*, Ital. *Monaco di Baviera* for Germ. *München*, Germ. *Mailand* for Ital. *Milano*). More often, however, large or well-known cities do not have a genuine exonym in a foreign language, but only a relatively stable standardised pronunciation (e.g. the English pronunciation of *Paris* [p{rIs]). Only very few of these names are present in the lists; most of the selected items are rather unfamiliar city names whose pronunciation by non-natives is less predictable.
- *Exclusion of names from non-standard varieties/languages*. In cases where place names in one country have distinct linguistic origins, only the 'standard' language is included. This criterion is relevant e.g. for the selection of British place names, where names of Welsh or Scottish origin (such as *Aberystwyth* or *Crianlarich*) were not included, or for French names of Breton origin (such as *Guingamp* or *Plogoff*).

5.2.2 Instructions to subjects

In any experimental study of non-native speech, subject instruction is an important issue. Subjects may feel uncomfortable with the task and may be concerned about being tested on their foreign language proficiency (see also Tomokyio & Burger, 1999).

Therefore, all recording sessions were introduced by an oral instruction and a clear explication of the goals of the experiment. Subjects were encouraged to apply any kind of foreign language knowledge, regardless of potential mistakes. It was emphasised that the experiment is not designed as a proficiency test and that correct pronunciation is neither expected nor desired. Even if these introductory remarks cannot completely compensate for the somewhat artificial experimental situation, it may help to reduce the subjects' feelings of anxiety and the resulting potential effects on the phonetic performance.

5.2.3 Task 1: Reading task (scripted prompts)

The first task is central to the overall data collection. Here, subjects read the lists of place names from all target languages as isolated words. Each reading list is arranged

[2] Phonetic symbols according to the SAMPA alphabet (Wells, 2000).

and presented in a fixed order and contains repetitions of two items at the end of each list in order to capture some effects of potential intra-speaker variation.

The city names reading task is obligatory for all subjects, regardless of their proficiency in the relevant target language(s). In addition, each subject was asked to read the names from their own native language, as suitable recordings of native speech were employed as acoustic prompts for the second task (see below).

Subjects are permitted and encouraged to read silently before pronouncing the names and to repeat individual names until they eventually arrive at their intended pronunciation. Since any delays and hesitations in the subject's responses are recorded, this information can be used to estimate the degree of difficulty of some words.

In contrast to the study carried out by Fitt (1995), where the subjects themselves had to determine the origin of the city names (by choice from a closed set of six languages), the linguistic origin of the material used in the present experiment is known to the subjects at each stage of the recording sessions.

5.2.4 Task 2: Repetition of native pronunciation (voice prompts)

In the second task, the subjects listen to the correct pronunciation of the names read by a native speaker and are advised to repeat this pronunciation. The acoustic prompts are presented via headphones; subjects may listen to the reference pronuncation repeatedly. The reference speakers were selected from former recordings of the reading lists (task 1) from which one speaker per language – preferably with clear articulation and standard pronunciation – was chosen. Thus the stimuli in this task are identical to the names used in the reading task, but presented in a different order to reduce the chance of identifying familiar words by their order of appearance.

This task was included to allow a comparison of the pronuncation variants produced in reading and repeating. A direct comparison is expected to yield conclusions about the particular influence of spelling pronunciation, which is effective only in the reading task, while the repetition task will involve purely phonetic interference phenomena.

5.2.5 Task 3: Short sentences (scripted prompts)

In addition to the city names tasks, lists of ten short sentences (semantically non-related) were compiled for each target language. These sentences, taken from several online newspapers, do not contain any place names. This task is not obligatory for all speakers and all target languages: Subjects were asked to read the sentences only if they have at least a basic knowledge of the target languages and feel comfortable with the task. In order to avoid hesitations, false starts, disfluencies or the like, the sentences do not exceed a maximum length of 60 characters and have a relatively straightforward syntactic structure.

This additional task was originally designed to evaluate and double-check the language proficiency information in the questionnaires, which is based on self-judgement (see below). However, for the study of non-native speech in general, the short sentences can be used as an independent data collection.

5.3 Speakers

5.3.1 Speaker information and questionnaires

Each speaker had to complete a questionnaire that includes the following types of information:

- *Personal data*: age, sex, name, address, profession
- *Native language* or language spoken during the first 10 years of life
- *Country and region of origin* and current place of permanent residence
- *Language proficiency level* according to self-judgement
  - (a) for languages covered by the database
  - (b) for any additional language
- *Miscellaneous*: bilinguality, hearing or speaking impairments, etc.

The information provided in the questionnaires is crucial to relate individual speaker characteristics to particular features of their pronunciations. For instance, the question how and to what extent native regional dialects contribute to the speakers' pronuncation of foreign languages is an interesting research topic of its own (see James & Kettemann, 1983). In the present context, the speakers' proficiency levels are probably the most important piece of information. They are rated on a range from 0 to 5, where the figures represent the following levels:

0	none
1	beginner/basic knowlegde
2	intermediate
3	good
4	excellent/near-native
5	native

Proficiency levels were not only queried for the languages covered by the database, but also for any additional language, since it is a common observation that speakers who have none or only a fragmentary knowledge of the target language apply pronunciation rules of other languages they are aware of. Transfer from a third language is most likely to occur if the target language is deemed similar to the third language (e.g. Spanish rules applied to Italian).

Self-assessment of foreign language proficiency is a critical issue for at least two reasons: First, speakers may assess their proficiency in terms of vocabulary, syntax, idiomaticity, overall fluency, etc., which does not necessarily coincide with native-like pronunciation. Second, self-judgement may be too optimistic or too negative either. Since self-judgement can only provide an approximate rating of the speakers' phonetic skills, the speakers' judgements can be double-checked using the sentences recorded in task 3 and adjusted if necessary.

5.3.2 Speaker recruitment and distribution

Up to now, all speech recordings were conducted at the University of Bochum (Germany). The majority of the non-German speakers involved so far are exchange students from France, the British Isles (Britain and Ireland), and Italy (currently being recorded) who stayed in Germany for a limited period of time. This recruitment strategy, though very convenient on the one hand, has on the other hand a negative effect on the overall speaker distribution: First, language proficiency in this group tends to be above average. For instance, Table 2 shows that the English and French speakers' knowledge of German is rated approx. 3 (on a scale from 0 to 5) according to self-assessment. It is also worth mentioning that some of the subjects are language students, who are likely to have an increased linguistic awareness, independently of the languages involved.

Secondly, it seems to be a plausible assumption that age is an important variable governing the speakers' performance (see also Eklund & Lindström, 2001). Speaker recruitment among students, however, will naturally lead to an unbalanced speaker distribution with regard to age.

The following table shows an overview of the German, English, and French speakers recorded so far, along with the average proficiency levels in the target languages:

Native lang.	No. of spkrs	male/ female	avrg. age	avrg . proficiency (0 <= n <= 5)		
				Engl.	**French**	**Germ.**
Engl.	27	14 / 13	22	(5.0)	1.4	3.0
French	20	5 / 15	30	2.5	(5.0)	2.9
Germ.	22	14 / 8	23	2.9	1.2	(5.0)

Table 2: Speaker distribution and average proficiency

Due to the current misbalance mentioned above, future extensions of the SDB will focus on speakers of the age group > 40 years as well as speakers with a low proficiency level in the relevant languages. Ideally, part of the recordings should be carried out in the native countries where it will be less problematic to recruit low-proficiency speakers.

5.4 Recording conditions and format

All recordings were performed digitally directly to hard disk using a high-quality sound card and an AKG C 414 microphone. The speakers were located in a low-noise booth, at a microphone distance of 0.3 metres. The file format is Microsoft PCM (*.wav), at a sampling rate of 22.05 kHz, 16 bit resolution, mono. All recordings are segmented into files containing one utterance (word or short sentence) each. Currently, the total number of utterances (files) is approx. 23.000.

All non-audio data (transcriptions, speaker and recording session information) is stored in an MS Access relational database. Separate label files can automatically be extracted from this database.

5.5 Lexicon and phonetic transcription

The SDB contains a reference lexicon including an orthographic transcription and a phonemic reference transcription (citation form) for the entire language material used in the recordings, i.e. for all place names and short sentences of each language. In cases where place names have two alternative pronunications in their language of origin (e.g. *Broughton* in England), a variant is included. For the reference transcription, the standard SAMPA inventory for the corresponding languages is used (Wells, 2000).

For the word-level phonetic transcription of the speaker variants, however, standard SAMPA is insufficient: Although principally designed as a language-independent phonetic alphabet, SAMPA symbols are always interpreted in terms of their language-specific phonemic values. For instance, the symbol `/r/` is used for the 'r' sound in both English and Italian, but represents quite distinct consonants in the two languages (an approximant in the former and a trill in the latter case).

In a cross-lingual situation like the one investigated here, speakers will typically produce a mixture of native, foreign, and intermediate speech sounds. It is thus necessary to extend the basic language-specific inventories and to interpret the symbols in terms of language-independent phonetic values instead of language-specific phonemic values. Therefore the X-SAMPA set proposed by Wells (2000) is applied as an extension to SAMPA for the transcription of individual speaker variants in cases where standard SAMPA would cause ambiguities.

5.6 Annotation of phonetic variables

A primary goal in building up the SDB is the detection of linguistic regularities that govern the pronunciation of foreign place names by non-native speakers. However, with 43 (+ 2 repetitions) place names for each language, the vocabulary covered by the SDB is rather limited. Even if it will be possible to deduce typical non-native variants for this specific vocabulary, results will not straightforwardly be transferable to new words.

For this reason, an additional level of annotation was introduced. It has been adopted from sociolinguistic or dialectologist studies, where the basic unit of observation is not the word, but the *phonetic variable* (see Hudson, 1980, Ch. 5). A phonetic variable (PV) is defined as an element – a phoneme, a cluster of phonemes, or a syllable – that is expected to undergo a certain degree of variation, depending on particular speaker characteristics such as regional origin, social class, or speaking situation. With each realisation of the variable by a speaker, the variable takes on a specific *value*, which constitutes its *variant*.

This concept has been adopted for the SDB: Each word in the reading list/dictionary is the carrier of one or more PVs. PVs are defined as phonemes in context; identical phonemes in different contexts constitute two distinct PVs (e.g. word-initial `/a:/` vs. word-final `/a:/`). Also, the PVs are not uniquely tied to only one word; instead, one PV may occur in more than one lexicon item. Figure 1 illustrates the corresponding annotation scheme for phonetic variables and their speaker-specific values using the example of the English city name *Northampton* `[nO:T{mpt@n]` with its PVs `/O:/`, `/T/`, and `/@/`.

In order to decide what constitutes a relevant PV, the pronunciation variants produced by a particular speaker group are assessed by a phonetically trained person. Whenever a striking deviation from the standard pronunciation is noticed, this is marked as a relevant PV. Once a PV has been identified, the PV value for all speakers in one group is determined.

In a subsequent step, statistical information on the PV variants can be gathered. The SDB supports this process by a semi-automatic summary, calculation, and update of all PV values. This feature helps to trace pronunciation regularities within the observed speaker groups on the

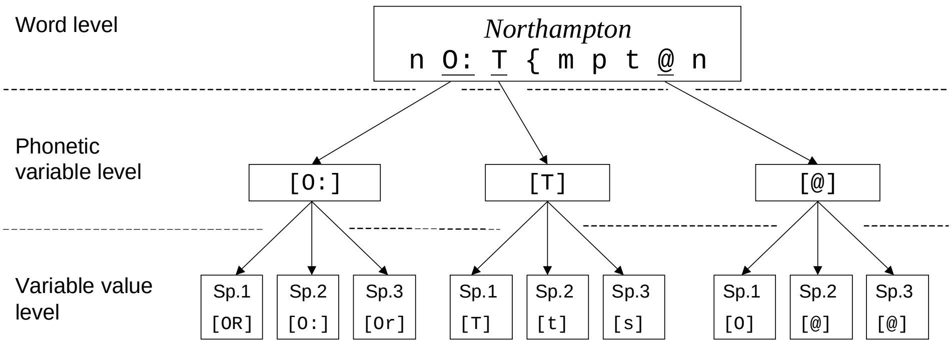

Fig. 1: Annotation scheme for phonetic variables

symbolic-phonetic level. Any generalisations can then be formalised as 'rewrite rules' of the form

$$X \rightarrow Y \ / \ \ Z$$

where X (in the context of Z) is transformed into the variant Y. The value of Y is determined by choosing the PV value which has the highest frequency within a particular speaker group. In this way, the information on PVs and their values can readily be exploited to derive pronunciation rules that can be applied to new vocabulary.

6. Future work, limitations, problems

Some limitations of the SDB in its current development stage have already been mentioned above. Most notably, the current speaker distribution is suboptimal with regard to age and proficiency and should be compensated for by adding speakers of at least another age and proficiency level group. These extentions to the SDB are planned for the second stage of the data collection, after having completed the current stage with its 20 speakers per native language.

While problems of speaker distribution can be settled in principle, a more serious challenge may be *intra-speaker variation*. As a number of studies suggest, it is a specific trait of non-native speech that speaking situation and individual psychological disposition have a significant influence on the speaker's phonetic performance (Dickerson, 1975). There may be a considerable degree of variation ranging from 'near-native' pronuncation in careful, formal speech to rather strongly accented speech in spontaneous, informal speech. Within the current experimental design, this characteristic of non-native pronunciation cannot be captured. Although the reading material is arranged in a way that allows to detect at least some instances of inconsistent speaker behavior (as some words and phonetic variables appear repeatedly and may be pronounced differently), effects of intra-speaker variation will probably be much more prominent in unsupervised speech.

This potential lack of naturalness, however, can be viewed as a general drawback of SDBs for which speech data is collected by any form of supervised elicitation. Thus, for further investigations into non-native speech, real-life speech data – e.g. telephone dialogues from the travel information domain – would certainly be a valuable resource.

7. Acknowledgement

This study is carried out at the Institute of Communication Acoustics, Ruhr-University Bochum (Prof. J. Blauert, PD U. Jekosch). It is funded by the *Deutsche Forschungsgemeinschaft* (DFG).

The author would like to thank Ms. Andreea Niculescu and Mr. Anders Krosch for their support in conducting the recording sessions and processing the data.

8. References

Belhoula, A., 1996. *Ein regelbasiertes Verfahren zur maschinellen Graphem-nach-Phonem-Umsetzung von Eigennamen in der Sprachsynthese*. Düsseldorf: VDI.

Benarousse, L.; Geoffrois, E.; Grieco, J.; Series, R.; Steeneken, H.; Stumpf, H.; Swail, C., and Thiel, D., 2001. The NATO Native and Non-Native (N4) Speech Corpus. *Proceedings Workshop on Multilingual Speech and Language Processing*, Aalborg, Denmark.

Cremelie, N.; ten Bosch, L., 2001. Improving the Recognition of Foreign Names and Non-Native Speech by Combining Multiple Grapheme-to-Phoneme Converters. *Proceedings ISCA ITRW Workshop Adaptation Methods for Speech Recognition*, Sophia Antipolis, France.

Dickerson, L.J., 1975. The Learner's Interlanguage as a System of Variable Rules. *TESOL Quarterly*, Vol. 9, No. 4, 401–408.

Eklund, R.; Lindström, A., 2001. Xenophones: An Investigation of Phone Set Expansion in Swedish and Implications for Speech Recognition and Speech Synthesis. *Speech Communication* 35 (1-2), 81–102.

Fitt, S., 1995. The Pronunciation of Unfamiliar Native and Non-native Town Names. *Proceedings Eurospeech 1995*, Madrid, Vol. 3, 2227–30.

Fitt, S., 1998. *Processing Unfamiliar Words. A Study in the Perception and Production of Native and Foreign Place Names*. Ph.D. Thesis, University of Edinburgh.

Goronzy, S.; Sahakyan, M.; Wokurek, W., 2001. Is Non-native Pronunciation Modelling Necessary? *Proceedings Eurospeech 2001*, Aalborg, Denmark, 309–312.

Gustafson, J., 1994. ONOMASTICA – Creating a Multilingual Dictionary of European Names". In: *FONETIK ´94, Papers from the 8th Swedish Phonetics Conference*, Lund, Sweden, 66–69.

Hudson, R., 1980. *Sociolinguistics*. Cambridge: Cambridge University Press.

James, A.; Kettemann, B. (eds.), 1983. *Dialektphonologie und Fremdsprachenerwerb*. Tübingen: Narr.

Mengel, A., 1993. Transcribing Names – a Multiple Choice Task: Mistakes, Pitfalls and Escape Routes. *Onomastica Research Colloquium 1993*, London.

Menzel, W.; Atwell E.; Bonaventura P.; Herron D.; Howarth P.; Morton R.; Souter C., 2000. The ISLE Corpus of Non-Native Spoken English. *Proceedings 2nd International Conference on Language Resources and Evaluation (LREC 2000)*, Athens, Greece, Vol. 2, 957–963.

Onomastica Consortium, 1995. "The Onomastica Interlanguage Pronunciation Lexicon." *Proceedings Eurospeech 1995*, 829–832.

Lamel, L.F.; Schiel F.; Fourcin, A.; Mariani, J.; Tillmann, H.G., 1994. The Translanguage English Database (TED). *Proceedings ICSLP 1994*, Yokohama, 1795–1798.

Schultz, T; Waibel, A., 2001. Language Independent and Language Adaptive Acoustic Modeling for Speech Recognition. *Speech Communication* 35 (1-2) 31–51.

Strik, H., 2001. Pronunciation Adaptation at the Lexical Level. *Proceedings ISCA ITRW Workshop Adaptation Methods for Speech Recognition*, Sophia Antipolis, France.

Tomokiyo, L.M.; Burger, S., 1999. Eliciting Natural Speech from Non-Native Users: Collecting Speech Data for LVCSR. *Proceedings ACL-IALL Joint Workshop on Computer Mediated Language Assessment and Evaluation in NLP*, Maryland.

Trancoso, I.; Viana, M.C., 1995. Issues in the Pronunciation of Proper Names: The Experience of the Onomastica Project. *Proceedings Workshop on Integration of Language and Speech*, Moscow.

Trancoso, I.; Viana, M.C.; Mascarenhas, M.I.; Teixeira, C.J., 1999. On Deriving Rules for Nativised Pronunciation in Navigation Queries. *Proceedings Eurospeech 1999*, Budapest, Vol. 1, 195–198.

van den Heuvel, H., 1994. Pronunciation of Foreign Names by Dutch Grapheme-to-Phoneme Converters. *Proceedings of the Department of Language and Speech*, University of Nijmegen, 99–105.

Wells, J., 2000. *SAMPA. Computer Readable Phonetic Alphabet*. http://www.phon.ucl.ac.uk/home/sampa/home.htm