

DEVELOPING THE GERMAN PRONUNCIATION DATABASE (DAD) - AN ONLINE DICTIONARY FOR SPOKEN GERMAN

Alexandra Ebel¹, Johannes Förster¹, Mathias Walther²

¹Martin-Luther-University Halle-Wittenberg, Halle, Germany

²Technical University of Applied Sciences Wildau, Wildau, Germany

alexandra.ebel@sprechwiss.uni-halle.de

Abstract: The Department of Speech Science and Phonetics at Martin-Luther-University Halle-Wittenberg is traditionally linked with orthoepy research. Since 2017 the Department is running a Standard German Pronunciation Database, called *Deutsche Aussprachedatenbank (DAD)*, which contains more than 130,000 entries. This database is currently being further developed in cooperation with the Professorship for business informatics at Technical University of Applied Sciences Wildau. The Department of Speech Science and Phonetics published three pronunciation dictionaries since the middle of the 20th century, the last one in 2010 [1]. But databases have many advantages over books, so in the 21st century the development of a database was a logical step. Especially audio files for the individual entries offer great added value. At this point, however, there is also a great need for developing because the audio files have to meet very specific criteria. The reason for this is that the phonetic characteristics of all entries must be reproduced uniformly and unambiguously. In our contribution we will give an insight in the German Pronunciation Database (DAD) and show functions of the current version. At the same time, we are going to mention limitations that are linked to this version. Finally, we give an outlook on technical and content-related expansions that are planned in the future.

1 Introduction in orthoepy research at Martin-Luther-University Halle-Wittenberg

At the Department of Speech Science and Phonetics orthoepy research is one of the traditional research areas. Since the beginning of these researches, pronunciation of standard German is analysed empirically through sociophonetic investigations, acoustic analyzes and evaluations of acceptability [cf. 1; 2]. The results have been published in the form of several pronunciation dictionaries, recently the *Deutsches Aussprachewörterbuch (DAWB - "German Pronunciation Dictionary")* [1]. The orthoepy research at Halle university is traditionally characterized by a descriptive understanding of pronunciation norms defining the following characteristics of German standard pronunciation [cf. 1]: It doesn't contain any dialectal features or colloquial forms, it can be understood in all German regions and social groups, it is used and expected in official public situations and its use promotes prestige in these situations, it is characterized by different types of articulatory precision, it is codified and is by that able to fulfil regulatory functions.

The recommended pronunciation applies especially in fields in which people speak consciously of form, like audiovisual media, stage, education, speech therapy and German as a foreign language. People expect use of standard pronunciation from persons that work in these fields how Hollmach [3] showed in a representative survey.

2 Pronunciation Databases in common

2.1 Restrictions in pronunciation dictionaries

The book format of an orthoepic code has some restrictions that limit users as well as researchers in their work. The main reason for these restrictions is limited space. This means that many

additional information cannot be given. For example the phonetic transcription of standard pronunciation is done using International Phonetic Alphabet (IPA). Another phonetic spelling like XSAMPA wouldn't find place in a dictionary and would also reduce clarity there. The provision of further information is also restricted. In the DAWB the lexical meaning is only given, if an identically written entry has more than one meaning, like German *August* (month) [aʊg'ʊst] vs. *August* (name) ['aʊgʊst]. Further restrictions concern the naming of indications of origin and the representation of pronunciation variants.

2.2 Opportunities in pronunciation databases

While the book format is associated with many restrictions, as shown in the previous chapter, a database allows a wide range of uses and expansion options for the existing amount of data. In contrast to a linear dictionary a web application can be used hypertextually. This makes links possible between individual entries and also to other websites that provide for example etymological information. Pronunciation databases provide several search functions, which significantly simplifies use. If you are looking for words with certain combinations of letters or even phones in certain positions or syllables, you can find them using the search box syntax. Of course, it is possible to search for IPA characters as well as XSAMPA characters and their combinations. There are also filter options to sort entries by initial letters, the indication of origin or number of syllables. Furthermore, you can give multimedia additions like audio files that make the transcribed pronunciation audible. The addition of audio files is certainly one of the greatest advantages that pronunciation databases have over dictionaries. Because many users are not used to read phonetic transcription and have difficulties to realize the pronunciation information correctly in complete. But for optimal usability some criteria for audio files should be considered.

2.3 Criteria for audio files in pronunciation databases

In audio files the audible pronunciation should reproduce unambiguously all features of the phonetic transcription such as aspiration of plosives, the in German very common devocalization of consonants in positions after voiceless fortis consonants or elisions. This is what many text-to-speech systems (TTS) are not able to achieve, although they have been improved significantly in recent years. When using TTS for generating audio files it is required that the API can be individually adapted to the transcriptions particularities. So it would be possible to train the TTS system to implement regularly occurring phonetic features in generating audio files.

An orthoepic database should be free of dialects or individual pronunciations. Websites like *forvo.com*, that produce their audio files with the help of the user community cannot meet this criterion. Often you can find more than one search result there and all of them give different pronunciations. For example if you search for the word *Bonbon* (what's a kind of candy): This word is used in German but has its origin in French. User may want to know, which syllable is stressed in German pronunciation - Does it show the typical stress on the first syllable, like many German words, or does it have its stress on the last syllable, what's typical for French word accent? In January 2021 *forvo.com* provides three audio files of the word *Bonbon* that were spoken in by three different persons. In one version the first syllable is stressed, in another version the second syllable is stressed and in the third version both syllables are equally stressed. Although there is even given an IPA transcription (which *forvo.com* doesn't provide for every word) users stay left alone with their question because in this transcription there is no primary stress marked: [bɔ̃.bɔ̃].

One last criterion for audio files that should be mentioned is the voices uniformity. If you use different voices (real or synthesized) it could happen that users are not able to distinguish which features are only individual and have no orthoepic basis (e. g. in aspiration degree or the specific realization of word accents). This illustrates another argument for TTS because with synthesized voices you are able to add or revise single words after years without differences in

voice quality, while a real speaker might then no longer be available or his/her voice has changed.

3 German Pronunciation Database (DAD)

The German Pronunciation Database (Deutsche Aussprachedatenbank DAD) has been developed by researchers at the Department for Speech Science and Phonetics at Martin-Luther-University Halle-Wittenberg in cooperation with the Professorship for business informatics at Technical University of Applied Sciences Wildau for some time. This database can be seen as a further development of the German Pronunciation Dictionary (DAWB; [2]). Already since 2011 the data pool with 130,000 entries exists. It was generated out of the DAWB dictionary. This data pool was subjected to a quality assurance control with regard to correctness and completeness of the entries as well as usability and systematics [4]. After that some systems were tested and the first version of the Pronunciation Database was set as a DokuWiki system. The first entries in this database can be dated to 2017. Based on the criteria mentioned in chapter 2.3 it was determined that audio files should be generated using a TTS system. For this purpose, the available data first had to be prepared and supplemented with additional metadata [5]. Since the middle of 2020, is being reengineered to overcome weaknesses that arose from old DokuWiki system.

3.1 Insight into the DAD

If you are looking for a certain word, you can use the search box. Depending on the entry one or more search results are displayed. For example the query for “Forsythie” (forsythia) only yields one hit, but the query for “Kaktus” (cactus) results in several hits including “Kaktusfeige” (prickly pear) and “Echinokaktus” (echinocactus). In figure 1 you can see the orthographic search for the word “Aussprache”, which returns four hits.

Wörterverzeichnis

Durchsuchen: [alle Sprachen](#)

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

Filterung mit %title% *~ Aussprache

[Zeige alle](#) ([Filter/Sortierung löschen](#))

#	↓Stichwort	Herkunft	Aussprache	Silbenzahl
	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
1	Aussprache		[ˈaʊ̯spʁɑːxə]	3
2	Aussprachefehler		[ˈaʊ̯spʁɑːxəfɛːlɐ]	5
3	Aussprachekodifizierung		[ˈaʊ̯spʁɑːxəkodifits̩iːʁʊŋ]	8
4	Aussprachewörterbuch		[ˈaʊ̯spʁɑːxəv̥ɛˈtɛbuːx]	6

 CSV-Export

Figure 1 - Screenshot of the DAD - search result for the word “Aussprache”

You can not only search for words in their graphematic form, but for combinations of IPA segments, too. The phonetic search is a central component of DAD and is supported by a virtual IPA keyboard that maps all vowels, consonants, diacritics and suprasegmentals according to the International Phonetic Alphabet (Figure 2).

@de 'aɔ̯s

Die Suche kann mit @de auf das Aussprachewörterbuch eingeschränkt werden. [Hilfe zur Suche](#)

IPA Keyboard

Konsonanten											
	bilabial	labiodental	dental	alveolar	postalveolar	retroflex	palatal	velar	uvular	pharyngeal	glottal
Plosiv	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Vibrant	ɸ			r					ʀ		
Tap oder Flap		ɸ		ɾ			ɽ				
Frikativ	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateralfrikativ				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateralapproximant				l		ɭ	ʎ	ʟ			

Vokale			
	vorn	zentral	hinten
hoch, gespannt	i y	ɨ ɯ	ɯ u
hoch, ungespannt	ɪ ʏ		ʊ
mittelhoch, gespannt	e ø	ɘ ɵ	ɤ ɔ
mittelhoch, ungespannt	ɛ œ æ	ə ɜ ɞ ɐ	ʌ ɔ̄
flach	a ɶ		ɑ ɒ

Diakritika und Suprasegmentalia	
Haupt-/Nebenakzent	ˈ ˌ
lang	ː
Silbengrenze	ˑ
verbunden	˗ˊˆ ˗ˊˆ˗
stimmlos	̥ ̦
silbisch/nichtsilbisch	̩ ̪ ̫ ̬
nasaliert	̃

Weitere Symbole	
[ɸ] reduziert	ɸ̥
[x] reduziert	x̥
[e] reduziert	e̥
stimmhaft labial-velarer Approximant	w

Figure 2 - Screenshot of the DAD - virtual IPA keyboard

For example, the phonetic search in the editorial portal for the syllable ['aɔ̯s] finds three hits, which have to be revised by editors (Figure 3).

Redaktionssuche

Filterung mit de.ipa *~ 'aɔ̯s

[Zeige alle \(Filter/Sortierung löschen\)](#)

#	↑Status	Stichwort	Herkunft	Aussprache	Silbenzahl
	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value="'aɔ̯s"/>	<input type="text" value=""/>
1	<input type="checkbox"/> Revision	Auster		['aɔ̯ste], ['ɔ̯ste]	2
2	<input type="checkbox"/> Revision	Austria		['aɔ̯stʝia:], ['ɔ̯stʝia:]	3
3	<input type="checkbox"/> Revision	Firdausi (pers.)	persisch	[fɪʔdˈaɔ̯si:]	3

[CSV-Export](#)

Figure 3 - Screenshot of the DAD - search results for the phone combination ['aɔ̯s]

An entry in the DAD consists of five or six lines. A sixth line comes into play when entries have a foreign language origin. As you can see in figure 4, the first line is the word in graphematic form. Below is the transcribed pronunciation with IPA characters. In lines three

and four you can find information on the number of syllables and the accent pattern. At last the transcription in XSAMPA characters is given.

The screenshot shows the homepage of the Deutsche Aussprachedatenbank (DAD) at Martin-Luther-Universität Halle-Wittenberg. The header features the logo [de:.a:d'e:] and the site name. A search bar is present with the text 'Suche'. Below the search bar, there is a note: 'Die Suche kann mit @de auf das Aussprachewörterbuch eingeschränkt werden. Hilfe zur Suche'. An 'IPA Keyboard' button is also visible. A navigation bar shows 'Zuletzt angesehen: Aktuelle Informationen Sachsen-Anhalt'. On the left, a sidebar contains a list of links: Hauptseite, Erweiterte Suche im Aussprachewörterbuch, Über die DAD, Hinweise zur Eindeutschung, Hinweise zur Transkription, Redaktionsportal, Kontakt, and Rechtliches. The main content area displays the entry for 'Sachsen-Anhalt' with the following details:

Aussprache:	[z,aks̩n̩ ¹ anhalt]
Silbenzahl:	4
Akzentmuster:	3142
SAMPA:	z%ak.sn = ." ?an.halt

Figure 4 - Screenshot of the DAD with the opened entry “Sachsen-Anhalt”

The accent pattern reflects the accent level assignment which was carried out automatically based on complex rules. This approach was particularly important with regard to speech synthesis. In German standard pronunciation we find a typical rhythm, that is created by a change from stressed and non-stressed syllables. To depict this rhythmic structure all entries were provided with a four-level pattern:

Level 4 represents the primary stressed syllable. Every word has only one primary stress. Level 3 marks all syllables that are secondary stressed. Not every word contains a secondary stress, but there are also words with more than one, like multi-part compounds. Level 2 indicates all syllables that could potentially be stressed in German standard pronunciation but aren't stressed in the concrete word. Eventually level 1 marks syllables that cannot be stressed in neutral realization (not emotional, not contrasting) of German standard pronunciation. That concerns syllables which include German reduction vowels [ə] or [ɐ] or elided vowels.

3.2 Problems of the application used

The beta version of the DAD currently in use is the web-based Open Source wiki software DokuWiki. The application enables user administration and rights assignment, editing and research as well as versioning of changes.

The DAD currently has 27 users of scientific departments. Of these, 11 are editors with full writing rights and 8 are lecturers who verify changes made by editors. Another 8 users are guests from scientific institutions who accompany the development of the beta version.

Since November 2018, 20,295 changes have been made by editors to the database, which currently contains 133,908 keyword entries (as of January 17th 2021). With the intention of increasing the number of verified entries and adding many new entries to the database, the number of editors and users should also increase.

However, with the use of DokuWiki comes the need to integrate many third-party dependencies. The further development of these community plugins, which are necessary for editing and searching entries, cannot be controlled and no future support is guaranteed. This results in the demand for an optimised front-end back-end structure according to the data format, which allows searching the pronunciation database, editing entries and versioning changes.

Even if basic functionalities of queries and filters of the database already exist with the search of the Dokuwiki application, long loading times of the search queries of the SQLite database as well as missing functions such as pagination, proper alphabetical sorting or an easy-to-use IPA keyboard severely limit the user experience.

4 Further developments of the German Pronunciation Database (DAD)

4.1 Re-implementation with a universal web framework

Due to the aforementioned limitations a new web interface for the DAD is being developed in a student project at TH Wildau. The new concept envisages to use the web framework Django which gives the flexibility to implement complex features that cannot be realised with DokuWiki. Django is a very versatile framework for web applications. It is implemented in Python, which is among the most popular programming languages [6]. Since the initial release in 2005 Django is widely used by many companies like Instagram, Dropbox and Uber, where it powers very large applications [7]. The most prominent features of Django are [8]:

- user authentication,
- database models,
- admin interface,
- support for multiple database back-ends.

The built-in user authentication and user management system is used to realise the role model for editors, lectors and users. The concept of the user management is similar to that of a web content management system. It consists of different permissions which are assigned to user roles. Most important is the copy-editing process which is distributed over multiple steps and persons.

In terms of the data base models, the improvements over the current version include the extension of the database schema. The DokuWiki solution cannot store with additional content and information like audio files, tags, meta data and semantic data, therefore this must be incorporated into the new data model. Furthermore, the revision system of changes in the DAD has to be adjusted in the database. Each editor must be able to display a history of the changes of a specific dictionary entry.

The advantages of the built-in admin interface mostly lie in the reduction of development time, since in all phases of the implementation access to data on a non-programmer level is granted. This helps in the transition process from the old system to the new one. The full flexibility of a web framework is also needed in the user interface. Due to new user demands, a new start page should show a clear, user-friendly front-end, which is also optimized for mobile devices. Django offers the possibility to implement modern user interfaces with state-of-the-art Javascript libraries [7].

The flexibility in changing the database back-end is an important factor in future development. The current prototype runs on MySQL/MariaDB. Nonetheless future requirements that originate from advanced search queries may demand a different database back-end.

4.2 Further developments in terms of content

The most urgent point is the completion of all entries with audio files. Various TTS systems are currently being tested to see which one best meets the requirements (cf. chapter 2.3).

Another idea is to group the vocabulary. So users could selectively display word from certain categories. Different classifications are conceivable. It would be possible to sort all words of the basic vocabulary according to learning level corresponding to the Common European Framework of Reference for Languages. Terms that are often used in certain situations could also be grouped together (e. g. in restaurant, at the doctor's, terms from the field of politics, chemical elements, toponyms, athlete names etc.)

An important target is to add inflected form to the DAD. Especially for learners of German as a foreign language not only the pronunciation of an infinitive is of interest, but also inflexion forms. To give an example: German verb "sprechen" shows the following stem forms: "spricht" - "sprach" - "hat gesprochen". These inflexion forms go with a vowel change. This in turn is associated with a change in the subsequent fricative, because in German standard pronunciation <ch> is realized as [ç] only after front tongue vowels, while it is realized as [x] after central and back vowels: [ʃpʁ'ɛçŋ] – [ʃpʁɪçt] – [ʃpʁa:x] – [hat ɡəʃpʁ'ɔxŋ]. Of course, these extensions should be supplemented by audio files, too.

5 Acknowledgements

Thanks to Seher Aksoy, Bénédicte Loesert-Albrecht, Hai Yen Nguyen, our student project team from TH Wildau, who helped us with gathering requirements, planning the project and implementing a prototype.

Literatur

- [1] KRECH, E.-M., STOCK, E., HIRSCHFELD, U and ANDERS, L. C.: *Deutsches Aussprachewörterbuch*. De Gruyter, Berlin/New York, 2010.
- [2] EBEL, A.: *Aussprache russischer Namen in der bundesdeutschen Standardsprache. Untersuchung zur Akzeptanz verschiedener Eindeutschungsgrade*. Frank & Timme, Berlin, 2015.
- [3] HOLLMACH, U.: *Untersuchung zur Kodifizierung der Standardaussprache im Deutschen*. Peter Lang, Frankfurt a. M., 2007.
- [4] FÖRSTER, J.: *Aufbau und Entwicklung der Deutschen Aussprachedatenbank (DAD). Softwarequalitätssicherung für ein digitales Aussprachewörterbuch*. In: EBEL, A. (Ed.): *Aussprache und Sprechen im interkulturellen, medienvermittelten und pädagogischen Kontext*, S. 105 – 117. <http://nbn-resolving.de/urn:nbn:de:gbv:3:2-24373>, 2014.
- [5] DRECHSEL, S.: *Aufbereitung des Halle-Korpus für die maschinelle Verarbeitung*. In: EBEL, A. (Ed.): *Anwendungsbeispiele neuer und etablierter Analyseverfahren in der sprechwissenschaftlichen Phonetik und Rhetorik*, S. 45 – 65. <http://nbn-resolving.de/urn:nbn:de:gbv:3:2-109356>, 2020.
- [6] TIOBE Index for January 2021: Online: <https://www.tiobe.com/tiobe-index/> (17 January 2021)
- [7] YUDIN, A.: *Building Versatile Mobile Apps with Python and REST*. apress, New York, 2020.
- [8] VINCENT, W. S.: *Django for Beginners: Build websites with Python and Django*. Leanpub, Victoria, 2020.