

Phon: Free Software for Phonological Transcription and Analysis

From <http://phon.ling.mun.ca/phontrac>

Reviewed by HEATHER BUCHAN,¹ *University of Wollongong*

1. OVERVIEW. Phon is an open-source program for the transcription and analysis of phonological and phonetic data. It was designed to help systematize research in children's phonological development, but many functions in Phon, particularly the powerful search function, can be used for a wide range of investigations in phonetics and phonology. Phon is compatible with other language processing programs and is not just limited to English, making it a useful tool for documenting and analyzing the phonological system of any spoken language.

The program has a user-friendly interface that makes transcription and analysis a relatively straightforward process. The layout is easily customized, as each major feature opens in a new view panel and so can be displayed as needed. Phon supports blind transcription, allowing multiple transcribers to work on the same file independently, and a validation feature allows the multiple transcriptions to be displayed together for consensus discussions and transcription validation. The media player supports both audio and video files, and media files can be segmented to link to sections of the transcription. Each transcript section is called a *record*, and, once linked to the media, the waveform of each record can be displayed.

Phon uses Unicode fonts, and there are built-in dictionaries for some languages that facilitate transcription of the target phonological form. Although the built-in dictionaries are only of major languages, there is also a dictionary utility which allows user-defined dictionary input that can show the IPA transcription of target phonological forms. There is also an automatic syllabification feature that can be adapted for other languages. Perhaps most impressive are the program's search capabilities. The search function permits systematic searches of phones and meta-phones, and the program developers have created Phon Expressions (PhonEx), a language that allows for complex searches using common linguistic terms. PhonEx is logical and easy to learn from the manual, especially for researchers already familiar with linguistic terminology. The ability to search phonetic transcriptions for feature classes in specific phonological environments makes Phon a very useful program for describing and analyzing phonetic features and phonological systems of spoken language.

Another very useful feature of Phon is that transcripts can be imported and exported in CSV and XML formats. This makes it easy to archive transcripts and to use Phon in conjunction with other language documentation software. For example, users can download

¹ I would like to thank my supervisors Caroline Jones of University of Wollongong and Felicity Meakins of The University of Queensland for their valuable suggestions. I am also very grateful to Yvan Rose of Memorial University for his helpful feedback.

the program Chatter, which is freely available online, to convert CHAT (the CHILDES transcription format) transcripts into XML. Users can then use another freely available script, xml2phon, to import these into Phon for phonetic and phonological analysis.²

Phon is a good primary transcription tool, especially for phonetic transcription as the IPA dictionary function can substantially speed up transcription time. However, transcripts can also be converted for compatibility with programs like Transcriber and ELAN. This means that, if a corpus has previously been transcribed in one of these programs, it can then be imported into Phon for phonetic transcription and analysis.³

Similarly, search results can be exported into a variety of formats including Microsoft Excel and Word, OpenOffice Calc and Writer, as well as more general PDF, HTML, and CSV formats. The latter is especially suited for transporting Phon data into other applications and statistical analysis programs. In this way, Phon's unique search function can be used in conjunction with functions offered by other programs.

Phon was developed by a team led by Yvan Rose and Greg Hedlund as part of the PhonBank project, a recent expansion of the CHILDES project (Child Language Data Exchange System), an international database that supports research in language acquisition (Rose et al. 2007). Coordinated by Brian MacWhinney of Carnegie Mellon University, the system facilitates data sharing across the field of child language research (MacWhinney 2000). CHILDES is a component of TalkBank, a multilingual corpus containing shared databases from several subfields of human and animal communication research. Within this larger database, PhonBank offers corpora documenting phonological development in a number of languages. PhonBank data are compatible with the Phon and CLAN software programs.

Good support is available for Phon users. A PDF version of the manual is available online and in the program's help menu. Users can ask for help and post feature requests in an online forum⁴ that is checked regularly by the developers. There is also a Google discussion group for PhonBank.⁵

Phon has very recently been released in a new version (v1.4). This review is of the newest version, Phon 1.4b975. This version is a substantial improvement over Phon 1.3, particularly in layout and navigation. User feature requests have been incorporated into v1.4, and existing functions have been updated. The program includes a simple tool for converting project files from v1.3 to v1.4.

2. GETTING STARTED. Projects are organized hierarchically, with each project made up of one or more corpora, and each corpus containing one or more sessions. At this stage, the user needs to decide whether to work in the *Editor's* mode or the *Blind* mode, the former

² Chatter and xml2phon can both be downloaded for free from <http://talkbank.org/>

³ Trevor Johnston at Macquarie University identified the potential use of Phon in this way.

⁴ <http://phon.ling.mun.ca/phontrac/discussion>

⁵ phon-subscribe@googlegroups.com

being the default. Transcription done in *Editor's* mode will not appear in the *Blind* mode, so it is important to make a decision about which mode to use at the beginning of the project. The *Blind* mode enables blind transcription from multiple transcribers and should be checked at the start of the project if comparison of phonetic transcription by two or more transcribers is desired. This feature is useful when discussing transcript discrepancies as part of a validation procedure, and also for documentation as a way to keep track of changes made through consultation with native speakers.

Transcription is done in the *Session Editor*. The layout here can be organized according to the user's needs at the time. The displays (view panels) for different functions (e.g., media segmentation, automatic IPA lookup, search, and search results) can be opened, moved around, and docked or undocked as needed. Customization of the layout is one of the strong points of Phon, allowing users to display only the functions relevant to the specific task at hand (e.g., there can be different configurations of view panels for transcribing, validating transcriptions, or searching). Figure 1 is an example setup for transcribing in Phon. The *Record List* on the left makes it easy to navigate between *records*, which are portions of the transcript that are useful for research, typically consisting of single utterances. The *Session Information* and *Record Data* in the middle display the participants' details and the record transcript respectively. The *IPA Lookup* on the right displays the dictionary IPA transcription of the *Record Data*.

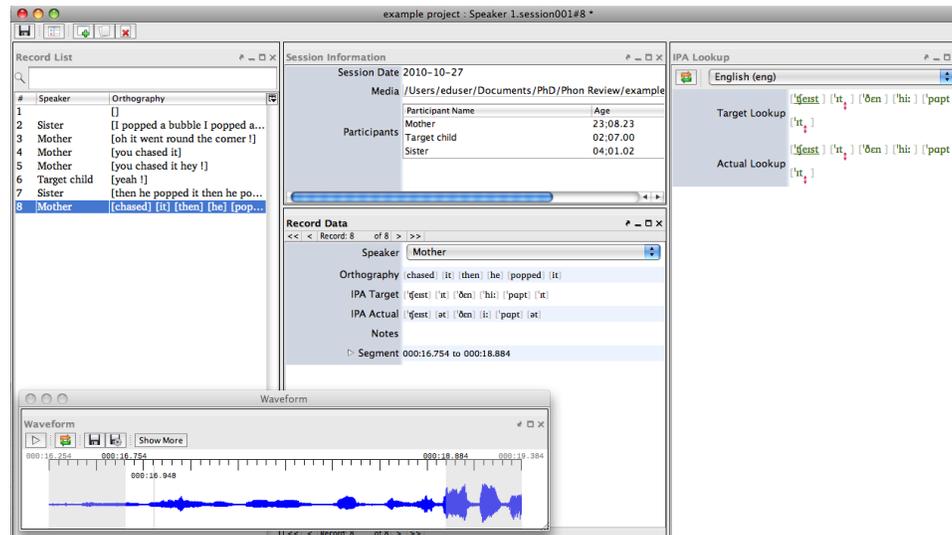


FIGURE 1. Example of a customized session editor display in Phon.

3. MEDIA. The media player in Phon supports a range of audio and video formats and can handle high quality recordings. (I use 24 bit 48 kHz audio files, which work well in this version.) Media files can be added by locating the file in the *Session Information* view panel and played by opening the player from the *Media* menu. Media can be segmented to create *records* (though data records can also be created in the absence of recorded media).

The segmentation feature is simple to use. After entering participant information, the user can decide to segment media from the beginning of the recorded media file or from the end of the last record segmented. To segment, simply use either the keyboard shortcuts or click on a participant in the media segmentation view panel, as shown in Figure 2. This will automatically create a record for that participant that links to the media segment. The default segment window (timeframe of an automatically-created segment) is 3000ms and can be changed just by typing in another timeframe. I found the default 3000ms window generally to work well for naturalistic interactions between, for example, mothers and children. Segment times can also be easily adjusted in the segment tier of each record, or by moving the boundaries in the waveform. The segmentation tool is also useful in that it allows the user to export individual media clips as separate files. This is good for storing soundbites and is also very helpful when audio in the segment is unclear, because the user can then quickly open the individual segment in a speech analysis program such as PRAAT for closer inspection.



FIGURE 2. Using the segmentation tool to create a record that is linked to the media.

4. TRANSCRIPTION. Once a record has been created, the associated utterance(s) can be transcribed. The *Record Data* window in the example in Figure 2 shows the default transcript tiers. The *Tier Management* function allows the user to re-order and show or hide tiers and to create new tiers. There is an IPA chart, available from the Tools menu, for insertion of phonetic characters, suprasegmentals, and diacritics.

At this stage the user should decide how to group words. *Word groups* in Phon refer to user-defined boundaries and are useful for defining boundaries of a lexical word such as *should*, or a prosodic word like *oughta*. Word groups are represented by square brackets and can be created in the *Orthography* tier, which then automatically creates groups in the *IPA Target* tier (transcription of the model or standard utterance form) and the *IPA Actual* tier (transcription of the produced utterance form), as well as any user-defined tiers that have been created to align with word groups. This feature is a way in which the transcript structure can be adapted to specific requirements of the research since it allows utterances to be segmented at any level. So, depending on the level of analysis, groups can be created for prosodic phrases, morphemes, or other segments.

The IPA target refers to the model phonological form, and the search function in Phon

can make systematic comparisons between target and actual phonological forms. This is a very useful feature that can be applied, for example, to compare phonemic and phonetic transcriptions, standard and colloquial pronunciations, adult forms and actual child pronunciations, or standard and baby-talk words.

The target phonological forms can be found in a built-in dictionary and inserted into the transcript. Dictionaries are already built in for some major languages (American English, Catalan, German, French, Icelandic, Dutch, Italian, and Spanish), and can be imported for other languages. A fieldworker could therefore create a word list with IPA transcriptions of the most frequent words in a corpus and build it into Phon as a basic dictionary. The phonological word forms can then be displayed in the dictionary look-up view panel and inserted into the transcript simply by clicking on each word form. If a word in the dictionary look-up has more than one IPA entry, the arrow keys can be used to scroll through the possible alternatives. (For example you can choose between [ət] and [ɪt] for the English word *it*.) Combined with the waveform feature, which allows the user to isolate segments in an utterance, this feature is an efficient way for a fieldworker to check transcriptions with native speakers; the user can play a segment and browse potential phonological forms for that segment simultaneously.

Phon also has tiers for target and actual *Syllabification* and *Alignment*, which offer syllable-level data annotation as well as systematic phone alignment between the *IPA Target* and *IPA Actual* tiers (see Figure 3 for an example). Syllable information is automatically labelled and displayed graphically in the transcript, with boxes around each syllable and different colours representing syllable constituents. The automatic syllabification can be edited manually by simply clicking on a syllable constituent and selecting the correct label from a drop-down list. The syllabification feature does not require the use of stress symbols in transcription and provides a quick and easy way of seeing how utterances are segmented into syllables and the alignment between the *IPA Target* and *IPA Actual* tiers.

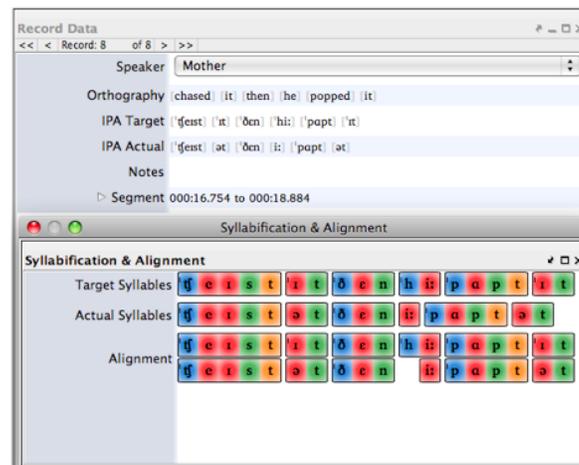


FIGURE 3. Example of Phon's syllabification and alignment feature

5. SEARCH. The search function in Phon is excellent and allows for advanced queries. While this is just an overview of the main search features, searching in Phon can be adapted to the most complex and detailed level necessary. Searches can be performed both within and across transcripts and results can be filtered by speaker. Tiers can be queried individually using the *Data Tiers* search (e.g., to search for all instances of a particular word in the *Orthography* tier), or multiple tiers can be queried at once (e.g., to compare segments in the *IPA Target* and *IPA Actual* tiers). Phon also allows searches for stress patterns and syllable types (i.e. consonant, vowel, and glide combinations). The results can be exported in the variety of formats listed above, making it easier for further analysis. For example, CSV report exports are useful for data post-processing in statistical analysis programs.

The search function is particularly useful because it was specifically designed for phonological analysis. There are tools for searching vowel and consonant harmony and consonant metathesis, in which the user can select the type of harmony and list the features and directionality of the process. As well as using plain text and regular expressions, complex queries can also be created using PhonEx, a language built for Phon using common linguistic terms. The manual describes PhonEx structure and terms in detail, and it does not take long to learn how to create queries. PhonEx allows search expressions to be as broad or specific as needed. For example, you can search for all instances of word-initial [k], all vowels in unstressed syllables, dorsal sonorants in primary stressed syllables in onset position, one or more phones not in coda position and that occur at the end of a word group, and so on. The output of searches gives summaries of the total number of instances in a transcript, as well as a detailed report with the record numbers for each instance. (Searching within a session even permits users to click in the results box to bring up the records in which each instance occurs.) The advanced search system is what really sets Phon apart from other phonetic transcription software. It is not as complicated as it may appear at first, results are easy to navigate, and it allows for a very wide range of complex searches depending on the researcher's needs.

6. CONCLUSION. Phon is a very effective program for transcription and analysis of phonological and phonetic data. Transcribing in Phon is straightforward, and the media segmentation and dictionary IPA look-up features help speed up the transcription process. The powerful search capabilities make it possible to describe the phonological system of a language variety in great detail. Phon would be a good asset for a wide variety of investigations in phonetics and phonology. The range of useful features and the customisable layout make Phon a very handy tool for fieldworkers interested in describing the phonology of an underdocumented language.

Primary function:	Transcription and analysis of phonological and phonetic data
Pros:	Useful interface for phonetic transcription (includes support for blind transcriptions) and annotation; support for media segmentation and transcript linkage; powerful search function for complex searches of many transcripts at once; compatibility with other transcription and analysis programs using importing and exporting functions; compliance with Unicode font encoding; good support from program developers
Cons:	None significant
Platforms:	Mac OS X 10.4 and higher, Windows XP and higher (Java 1.6 compatibility required on all platforms)
Open Source:	Yes, Phon is licensed under the GNU General Public License Version 2.
Proprietary:	Users can use Phon under the GNU-GPL license.
Available from:	Phon can be downloaded from http://phon.ling.mun.ca/phontrac/wiki/Downloads
Cost:	Free
Reviewed version:	Phon 1.4b975
Application size:	Mac OS X: 41MB; Windows XP/Vista/7: 95MB (includes VLC 1.10 for the media player)
Documentation:	http://phon.ling.mun.ca/phontrac/wiki/

REFERENCES

- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk. Volume 1: Transcription format and programs. Volume 2: The Database*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Rose, Yvan, Gregory J. Hedlund, Rod Byrne, Todd Wareham & Brian MacWhinney. 2007. Phon 1.2: A computational basis for phonological database elaboration and model testing. In Paula Buttery, Aline Villavicencio & Anna Korhonen (eds.), *Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition, 45th Annual Meeting of the Association for Computational Linguistics*, 17–24. Stroudsburg, PA: ACL. For recent developments see <http://phon.ling.mun.ca/phontrac/wiki/>.

Heather Buchan
hb307@uowmail.edu.au