

iMnem: Interactive Mnemonic Word Suggestion Using Phonetic Algorithms

Orapin Anonthanasap and Teerapong Leelanupab

Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang (KMUTL), Bangkok, Thailand. 10520

o.anonthanasap@gmail.com, teerapong@it.kmutl.ac.th

Abstract: To support language learning by using the principle of a Mnemonic technique, this paper proposes to automatically generate suggested mnemonic words by using “phonetic algorithms”, i.e., Soundex and Metaphone Levenshtein edit distance is employed to compare the phonetic similarity of foreign words and that of words in a known language using the sound transcriptions transformed by the proposed algorithms. Our new interactive cross-lingual system, called iMnem, is also introduced to support the task of searching for mnemonic words with images for better imagination of word association. To the best of our knowledge, this work is the first to apply phonetic algorithms for mnemonic word generation. In this study, we focus on suggesting mnemonic keywords in English for supporting learning of words in Japanese.

Keywords: Mnemonics, Phonetic Similarity, Foreign-Language Learning

1 INTRODUCTION

Learning a new language is often a matter of applying various memory techniques to memorize language structures such as grammar, vocabulary, and contextual usage of appropriate words and sentences. Mnemonics for *vocabulary* learning concerns learning techniques that assist in the retention of complex or unfamiliar foreign words that are, at the beginning, difficult to memorize, especially for new learners. One of these techniques transforms foreign words into the so-called mnemonic¹ keywords, which are immediately understandable and thus memorable to the learners. Mnemonic keywords are those that have phonetic similarities (or similar word pronunciation) to words in target foreign languages. For example, an English-speaking learner desires to learn Japanese language. To remember the word “Ki (Tree)” in Japanese, the learner can associate it with a phonetically similar or mnemonic keyword, e.g., “Key”, in his known language (i.e., English). Then, a traditional context is provided as a memorable form of a sentence or an image that links the mnemonic word with the foreign one. For example, “Imagine I put the Key under the Tree”. To this end, learners can use mnemonic words to facilitate learning and recall of foreign vocabulary rather than attempting to remember a foreign vocabulary individually.

In this paper, we propose to generate suggested mnemonic words by using Soundex and Metaphone “phonetic algorithms”. We also present a new interactive

cross-lingual system, called iMnem, which supports users (e.g. teachers or learners) to generate mnemonic materials for learning Japanese. Automatic generation of mnemonic words is achieved by transforming foreign words (i.e. Japanese) and candidate mnemonic words (i.e. English) into phonetic transcriptions for computing phonetic similarities between pronunciations of these words using Levenshtein distance. The selected mnemonic words can then be visualized on the dynamic and interactive interface of our iMnem system. The iMnem interface allows vocabulary learners to seamlessly browse a collection of Japanese words with suggesting mnemonic words of a known English language. However, our approaches of this paper is to demonstrate our proposed approach that can computationally generate effective mnemonic words for learning Japanese vocabulary. To the best of our knowledge, this work is the first to apply phonetic algorithms for mnemonic word generation. Given a Japanese word, the focus of our paper is on how our system effectively suggests phonetically relevant/mnemonic words in an English language.

A comprehensive study is conducted using the evaluation set of 18 Japanese words to learn, and 5,000 most frequent English words as candidates for mnemonic words. In this study, our main hypothesis is that phonetic algorithms can be used to effectively generate mnemonic vocabulary for increasing new foreign word retention. Our proposed approach using two phonetic algorithms are compared against a baseline approach for typical transcription of word pronunciation, i.e., International Phonetic Alphabet (IPA). It can be argued that although IPA can represent phonetic transcription of speech sounds for all

¹ A mnemonic keyword is also often called a link-word.

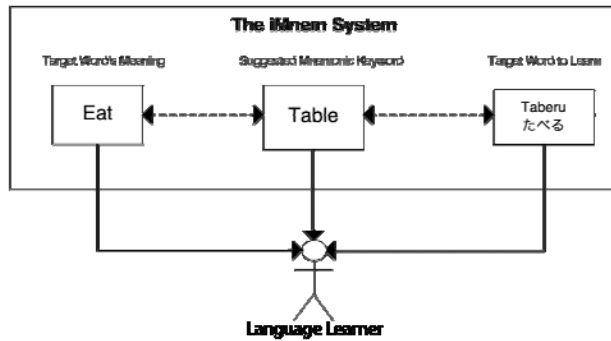


Fig. 1. The link system method of mnemonic learning strategy.

languages, IPA makes too fine distinctions between similar sounds. As a result, it is difficult to find exact match and nearly exact match for phonetic comparison of cross-language words. In addition, the aim of our research does not require such distinctions. The experimental results show that our proposed approach successfully suggests mnemonic words, leading to improved recall of newly learned words in Japanese.

2 RELATED WORKS

In the following subsections, we review the literature related to this work.

2.1 Mnemonics for Vocabulary Learning

Mnemonics is a memorization technique and considered a powerful learning strategy for vocabulary learning [1]. Previous studies have reported that mnemonic keywords can improve learner's memorization of foreign vocabulary by creating a word association between foreign language and native language of learners [4]. Atkinson *et al.* [4] evaluated the effectiveness of a mnemonic procedure, called the *keyword* method, for learning a foreign language vocabularies.

In their study, it was showed that the mnemonic strategy could enhance recognition and recall of foreign vocabularies. This result was confirmed by the later study of Fritz *et al.* [8], in which the mnemonic keyword strategy could improve the recall and recognition in different conditions. In psycholinguistics, Ellis and Beaton [7] demonstrated that keyword with acoustic similarity is an essential factors that affect the ease of learning foreign language vocabulary.

In Figure 1, the word *Taberu* in Japanese has similar sound with the word *Table* in English. A language learner can form a sentence by using the word, *Table*, as a mnemonic keyword and its definition in English, *Eat*, for assisting the recognition and recall of the Japanese word

Taberu. For instance, imagine that “you *Eat* your lunch on a *Table*.”

2.2 Mnemonics in HCI

The use of computer technology with language learning has a large volume of prior works [3, 11]. Savva *et al.* [16] presented TransPhoner a cross-lingual system for generating mnemonic keywords by using International Phonetic Alphabet (IPA). The keywords were selected by considering phonetic, orthographic and semantic word similarity, and word concept imageability. However, their work mainly focused on evaluating the user performance of learning vocabulary as supported by TransPhoner, but not on the algorithms for generating the keywords like this work of us.

2.3 Linguistics Terminology: International Phonetic Alphabet (IPA)

The International Phonetic Alphabet, written in Latin and Greek alphabets, is a phonetic notation for representing speech sounds for all languages [9]. The IPA organizes and classifies all of consonants and vowels presented in human speech: phones, phonemes, intonation, and syllables. IPA can be used as a cross-language transcription of phones: the smallest sound unit of speech. Each language has a subset of phones, also known as language specific phonemes. These are basic units of the phonology to distinguish meanings in a language. Phonemes can be combined to form the pronunciation of words and considered equivalent to a set of phones.

3 PROPOSED APPROACH TO MNEMONIC KEYWORD SUGGESTION

3.1 Phonetic Algorithm

A phonetic algorithm indexes words by converting their pronunciations into different codes based upon applied algorithms. Homophones of two words are encoded to the same representation. Thus, the code of similar sound word can be matched although their actual spellings are different, such as “Sun” in English and “San” in Japanese. The algorithms are useful for applying in our mnemonic keyword suggestion for searching similar sound words in cross-languages. Two state-of-the-art phonetic algorithms that we proposed and investigated in this study are: i) Soundex and ii) Metaphone.

Two pragmatic issues must be addressed when phonetic encoding and matching are performed [18,19]. The first is the responses time to queries. Answers should be found and returned quickly. The second is the issue of effectiveness as for Information Retrieval (IR). That is,

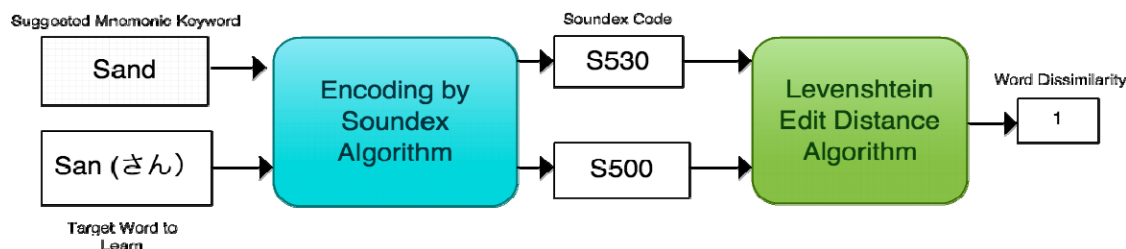


Fig. 2. The process of measuring phonetic similarity of words.

some approaches are better than others at identifying matches and ranking them at a top list. It is the issue of effectiveness that we explore in this paper.

3.2. Soundex

Soundex, the best-known algorithm for encoding word pronunciation to Latin alphabet, followed by three numerical digits [5,12,14]. The words that may have similar pronunciation, but their actual spellings are different, will be converted into the same code. For example, the Soundex code for the word “fish” is F200. The calculation is based on a group of characters with different numbers. The Soundex code can be computed as follows:

1. Keep the first letter of the word (in upper case)
2. Drop all of these letters (a,e,i,o,u,y,h,w) and replace with hyphens ‘-’
3. Replace the other consonants by numbers as follows (after the first letter):
 - (a) b,f,p,v -> 1
 - (b) c,g,j,k,q,s,x,z -> 2
 - (c) d,t -> 3
 - (d) l->4 (e) m,n -> 5 (f) r -> 6
4. Delete adjacent repeats of number.
5. Delete the hyphens.
6. Keep the first three numbers or pad out with zeros until there are three numbers.

3.3. Metaphone

Metaphone was improved from the Soundex algorithm by reducing English words to their basic sounds fundamentally. This algorithm more sensitive for changes in the sequence of the letters and also concerned information about variations and inconsistencies in English spelling and pronunciation to produce a more accurate encoding. The codes are set of the 16 characters 0BFHJKLMNPRSTWXY. The vowels AEIOU are also used, but only at the beginning of the code. We refer interested readers in a detailed algorithm of Meta- phone encoding to [15] or to ².

3.3. Phonetic Similarity Measurement

For the process of measuring phonetic similarity of words, we take each word from two different languages and measure differences in their pronunciations. We firstly convert each word into code based upon either Soundex or Metaphone phonetic algorithms. We then utilize Levenshtein edit distance to compute phonetic distance between two words of different languages. A result of the computed distance indicates the degree of phonetic dissimilarity between the words. In other words, it measures similarity of phones derived from the sound codes. The shorter distance, the more phonetic similarity between the words. Fig. 2 illustrates the process of measuring phonetic similarity of words.

3.4 Levenshtein Edit Distance

Levenshtein edit distance is an algorithm primarily used for measuring string similarity. This algorithm counts the least number of edit operations; insertions, deletions and substitutions of single characters required. The counting result of each string is used for similarity comparison with other strings [5, 13]. This approach does only edit operations between two strings. It has nothing related to phonetic matching between words. Nevertheless, if the matching strings are directly encoded based on the basis of pronunciation, then phonetic matching using this approach can be acceptable. Another research work has also been conducted using Levenshtein edit distance to compute phonetic similarities between IPA transcriptions of different languages [16]. However, it can be argued that the quality of similarity comparison using this approach is based on how the words are encoded phonetically.

3.4 iMnem System

We present a new mnemonic word suggestion system, iMnem, which employs our proposed phonetic algorithms for assisting second-language vocabulary learning. The system provides a collection of learning words in English and Japanese. Its dynamic and flexible interactive interfaces assist users (teachers/learners) to visually browse learning vocabulary (Japanese) by images. A learning word can also be entered as a query to search for mnemonic keywords in a

² <http://www.it.kmitl.ac.th/~teerapong/metaphone.pdf>

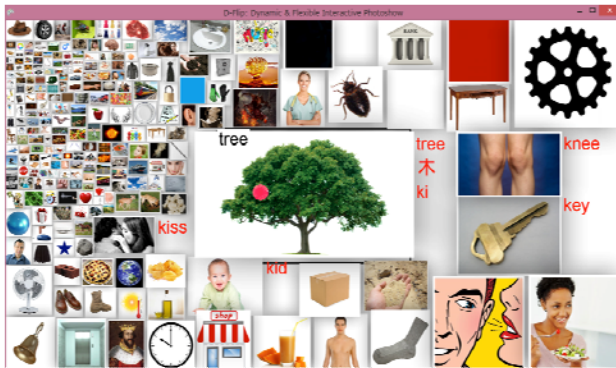


Fig. 3. Learners can get the mnemonics words of “Ki” including key, kiss, kid and knee.

system database. Suggested mnemonic keywords of a known (English) language are generated by phonetic algorithms based on selected vocabulary. They are then visualized as images dynamically on the interface of our iMnem system. A detailed explanation of our iMnem can be found in [2]. Fig. 3 illustrates the interface of the iMnem.

4 EXPERIMENTAL DESIGN

In this paper, we focus on evaluating the effectiveness of phonetic algorithms, i.e., Soundex and Metaphone, for mnemonic keyword generation. This work is different from our previous study [2] that focused on the acquisition and supportiveness of our iMnem system in preparing mnemonics materials to language learners or teachers. This section describes the experimental setup used to conduct our experiments. We hypothesized that using phonetic algorithm for generating Mnemonic keywords provides a better (or comparable) result than using IPA in terms of retrieval effectiveness. In order to validate this hypothesis, we perform a system-oriented evaluation, which means that the experiments require a test collection, queries (a set of foreign words in Japanese), and relevance judgments made for each mnemonic keyword suggested in English whether it is relevant or non-relevant to a foreign word.

Table 1. The statistics of a test collection

#Candidate mnemonic vocabularies	#Query	Avg. word per topic	Rel. word per topic
5000	18	24.15	7.5

4.1. Test Collection

Table 1 shows the statistics of a test collection.

- **Candidate mnemonic vocabularies:** five thousand most frequently used noun words. This set of vocabularies is obtained from a free available

subset of the word frequency data corpus of contemporary American English (COCA) [6].

- **Queries:** eighteen Japanese words/queries randomly sampled from the above set and in a recommended list of words to learn in a level of beginner or Japanese-Language Proficiency Test (JLPT) N5.
- **Ground truth data:** a pooling technique [5,17] was used to create the assessment list from a subset of suggested mnemonic keywords. This approach performs relevance assessments on a pool of English vocabularies suggested as mnemonic keywords generated by three studying approaches/systems, i.e., Soundex, Metaphone and IPA. This pool was created by merging results from the top 10 keywords, returned in response to each query/Japanese word from the suggestion systems. Three Japanese language teachers volunteered to be the experimental assessors to provide relevance judgments only for those keywords contained in the pool. Graded relevance with a 5-point scale³, i.e., {0,1,2,3,4}, was employed according to their preferences on the basis of the suitability of words to be mnemonic keywords for each query. Unjudged words are considered non-relevant.

4.2 Evaluation Measures

– Set-based Measures

Precision and recall are common measures used for evaluating the effectiveness of retrieval tasks. Both measures can also be evaluated at a given cut-off rank, considering only the topmost results, i.e., candidate mnemonic keywords, returned by the system.

Precision: Precision refers to the fraction of retrieved mnemonic keywords that are relevant.

$$\text{Precision} = \frac{\#(\text{relevant mnemonic keywords retrieved})}{\#(\text{retrieved mnemonic keywords})} \quad (1)$$

Recall: Recall refers to the fraction of relevant mnemonic keywords that are retrieved.

$$\text{Recall} = \frac{\#(\text{relevant mnemonic keywords retrieved})}{\#(\text{relevant mnemonic keywords})} \quad (2)$$

³ 0 is for non-relevant, 1 for marginally relevant, 2 for fairly relevant, 3 for relevant, and 4 for highly relevant

– Ranked Retrieval Measures

Mean Average Precision (MAP): Average precision (AP) refers to the average of the precision obtained for the top k mnemonic keywords, each relevant mnemonic keywords is retrieved. This value of AP is then averaged over the set of queries, or learning English words in this context, to obtain the mean average precision (MAP), i.e.

$$\text{MAP} = \frac{1}{|Q|} \sum_{i:q_i \in Q} \frac{1}{|R_{i,k}|} \sum_{j:d_j \in R_{i,k}} \text{Precision}@j \quad (3)$$

where q_i refers to the query in the set of queries Q , $R_{i,k}$ refers to the set of relevant documents for the i -th query from the top result until the document at rank k , and d_j is the relevant mnemonic keywords placed at rank j .

Normalised Discounted Cumulative Gain (nDCG): nDCG is an effectiveness measure that assesses document utility from cumulative gains [10]. The discounted cumulative gain (DCG) at rank r is defined as follows:

$$\text{DCG}@r = \sum_{i=1}^r \frac{J(d_i, q)}{\log_2(1 + i)} \quad (4)$$

where $J(d_i, q)$ is the relevance judgement of the i -th document given a query q in the ranked list, and the logarithmic denominator is the discount factor based on the rank positions of documents. nDCG is obtained by normalising this score by the DCG score of the ideal ranked list.

Note that in this study, graded relevance is used in nDCG, while binary relevance⁴ is used for Precision, Recall and MAP⁴.

4.3 Experimental Systems and their Settings

To evaluate the mnemonic keywords generated by three studying systems, we use “TREC eval”, the standard evaluation tool used by the IR community. We employed three systems including our two phonetic algorithms, i.e. Soundex, Metaphone and a baseline approach for typical transcription of word pronunciation.

We started by obtaining candidate mnemonic vocabularies from five thousand most frequently used noun words from the word frequency data corpus of contemporary American English (COCA) [6]. We then asked Japanese volunteers to translate them into Japanese

and used as queries. For a baseline approach International Phonetic Alphabet (IPA), the English IPA transcriptions were retrieved from pronunciation dictionaries and using rule-based for Japanese language [16]. The word similarity between input queries and candidate mnemonic vocabularies were computed by Levenshtein Distance using IPA transcriptions of them as well as the phonetic code in the cases of Soundex and Metaphone. We took only a top 10 result of each query from each investigating systems to evaluate its performance and construct ground truth data as described in Subsection 4.1.

5 RESULTS AND DISCUSSION

Table 2. Retrieval performances on Soundex, Metaphone and IPA

	Soundex	Metaphone	IPA
P@5, @10	0.59* , 0.42*	0.49*, 0.45*	0.31, 0.27
R@5, @10	0.44* , 0.56*	0.33*, 0.59*	0.19, 0.35
nDCG@5, @10	0.60* , 0.61*	0.45, 0.54*	0.32, 0.36
MAP@5, @10	0.38* , 0.46*	0.25, 0.39*	0.17, 0.23

Table 2 reports the results of Soundex, Metaphone and IPA, evaluated by Precision (denoted by P), Recall (indicated by R), MAP and nDCG at ranking positions 5 and 10. The results highlighted in bold show the best performance of the runs regarding the given measures. When statistical significant difference (according to a two-tailed t-test, with $p < 0.05$) against a baseline approach IPA is individuated, we report it with *. We compute statistical significance against IPA because it is considered as a state-of-the-art approach in the context of phonetic transcription.

As shown in the table, the baseline approach IPA had the lowest performance in all the values of evaluation metrics, followed by Metaphone and Soundex. Soundex obtained the highest score with significant difference to IPA in Precision@5 (0.59) and Recall@5 (0.44) - a user can find the most number of relevant mnemonic keywords at rank 5 in the Soundex system. While at rank position 10 (i.e., when ten mnemonic keywords are retrieved), Metaphone performed the best in terms of Precision and Recall, 0.45 and 0.59 respectively.

Furthermore, both Soundex and Metaphone obtained high scores in rank-based measures, MAP and nDCG. These results indicated that our proposed phonetic algorithms attempted to return relevant mnemonic keywords as early as possible. That is, language learners are more likely to encounter relevant mnemonic keywords at the top of the ranking.

⁴ We obtained binary relevance by combining four partial relevance grades {1,2,3,4} into 1 whereas 0 still stands for non-relevant.

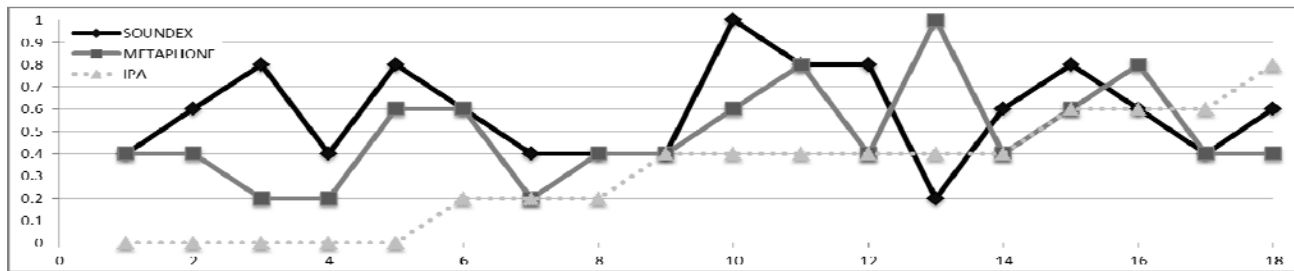


Fig. 4. Results of P@5 for all queries

From these results, it is suggested that our two phonetic algorithms can be employed to effectively generate mnemonic keywords as measured by all evaluation metrics. Fig. 4 shows the results of Precision at 5 for each query. As we can see, Soundex outperformed in almost all queries.

6 CONCLUSION AND FUTURE WORK

In this paper, we proposed an automatic mnemonic word suggestion based on phonetic association using phonetic algorithms. A system-oriented evaluation showed that our two approaches outperforms a baseline approach to generate mnemonic keywords, in particular Soundex algorithm. Future work will focus on a user study of iMnem to evaluate its effectiveness in assisting learning vocabulary and improving recall and retention of foreign target words.

REFERENCES

- [1] M. Amiryousefi and S. Ketabi. Mnemonic Instruction: A Way to Boost Vocabulary Learning and Recall. *Journal of Language Teaching and Research*, 2(1):178–182, Jan 2011.
- [2] O. Anonathanasap, C. He, K. Takashima, T. Leelanupab, and Y. Kitamura. Mnemonic-Based Interactive Interface for Second-Language Vocabulary Learning. In *Proceedings of the Human Interface Society, HIS' 14*.
- [3] N. Arnold and L. Ducate. Present and Future Promises of CALL: From Theory and Research to New Directions in Language Teaching. *Computer Assisted Language Instruction Consortium*, 2011.
- [4] R. C. Atkinson and M. R. Raugh. An Application of the Mnemonic Keyword Method to the Acquisition of a Russian Vocabulary. *Journal of Experimental Psychology: Human Learning and Memory* 1, 2:126–133, Mar 1975.
- [5] B. Croft, D. Metzler, and T. Strohman. *Search Engines: Information Retrieval in Practice*. Addison-Wesley, 1st edition, 2009.
- [6] M. Davies. The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, Architecture, and Linguistic Insights. *International Journal of Corpus Linguistics*, 14(2):159–190, 2009.
- [7] N. C. Ellis and A. Beaton. Psycholinguistic Determinants of Foreign Language Vocabulary Learning. *Language Learning*, 43(4):559–617, 1993.

- [8] C. O. Fritz, P. E. Morris, M. Acton, A. R. Voelkel, and R. Etkind. Comparing and Combining Retrieval Practice and the Keyword Mnemonic for Foreign Vocabulary Learning. *Applied Cognitive Psychology*, 21(4):499–526, 2007.
- [9] D. Horga. Handbook of the International Phonetic Association. A Guide to the Use of the International Phonetic Alphabet. *Govor*, 16(2):181–188, 1999.
- [10] K. Ja'rvälin and J. Keka'läinen. Cumulated Gain-Based Evaluation of IR Techniques. *ACM Transactions on Information Systems*, 20(4):422–446, 2002.
- [11] R. Kern. Perspectives on Technology in Learning and Teaching Languages. *TESOL Quarterly*, 40(1):183–210, 2006.
- [12] A. Lait and B. Randell. An Assessment of Name Matching Algorithms. Technical Report Series-University of Newcastle upon Tyne Computing Science, 1996.
- [13] V. Levenshtein. Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady*, 10:707, 1966.
- [14] K. Odell and R. Russell. Soundex Phonetic Comparison System. vol. 1261167 edition. US Patent, 1261167, 1918.
- [15] L. Philips. Hanging on the Metaphone. *Computer Language*, 7(12 (December)), 1990.
- [16] M. Savva, A. X. Chang, C. D. Manning, and P. Hanrahan. Transphoner: Automated mnemonic keyword generation. In *Proceedings of CHI '14*, pages 3725–3734.
- [17] K. Spa'ček-Jones and C. J. van Rijsbergen. Report on the Need for and Provision of an "Ideal" Information Retrieval Test Collection. Technical Report 5266, Computer Laboratory, University of Cambridge, 1975.
- [18] J. Zobel and P. Dart. Finding Approximate Matches in Large Lexicons. *Softw. Pract. Exper.*, 25(3):331–345, Mar. 1995.
- [19] J. Zobel and P. Dart. Phonetic String Matching: Lessons from information retrieval. In *Proceedings of SIGIR '96*, pages 166–172.