Analysis of the Accuracy of User-provided Information in Natural Language Queries for Music Information Retrieval*

음악정보 검색에서 이용자 자연어 질의의 정확성 연구

이진하(Jin-Ha Lee)**

ABSTRACT

Our limited understanding of real-life music information queries is an impediment to developing music information retrieval (MIR) systems that meet the needs of real users. This study aims to contribute to developing a theorized understanding of how people seek music information by an empirical investigation of real-life queries, in particular, focusing on the accuracy of user-provided information and users′ uncertainty expressions. This study found that much of users′ information is inaccurate; users made various syntactic and semantic errors in providing this information. Despite these inaccuracies and uncertainties, many queries were successful in eliciting correct answers. A theory from pragmatics is suggested as a partial explanation for the unexpected success of inaccurate queries.

초 록

실제 이용자들의 필요성을 충족하는 음악정보 검색 시스템을 개발하는데 있어서 실생활의 음악 정보 질의에 대한 부족한 이해가 장애가 되고 있다. 이 연구는 실생활 질의의 경험적 분석을 통해 이용자들이 어떻게 음악 정보를 찾는지에 대한 이론적인 이해를 돕고자 한다. 그 중에서도 미래의 음악정보 검색 시스템의 디자인, 특히 잠재적인 접근점을 선택하는데 있어서 결정적인 정보를 제공하는 역할을 하게 될 실생활의 음악정보 질의 내에서 이용자들이 제공한 정보의 정확성을 검토하고 있다. 이 연구는 이용자의 정보 중 상당 부분이 부정확한 정보임을 보여주며, 이런 부정확성과 불확실성에도 불구하고 다수의 질의가 성공적임을 알려준다. 또한 어용론의 이론으로써 부정확한 질의의 예기치 않은 성공에 대한 부분적 설명을 하였다.

키워드: music information retrieval, natural language queries, accuracy
음악정보 검색, 자연어 질의, 정확성

---

* 이 논문은 University of Illinois at Urbana-Champaign의 박사학위논문의 일부입니다.

** University of Illinois at Urbana-Champaign, Ph.D. (jinha.lee@gmail.com)

■ 논문접수일자 : 2008년 11월 15일 ■ 최초심사일자 : 2008년 11월 17일 ■ 게재확정일자 : 2008년 11월 26일
■ 情報管理學會誌, 25(4): 149-164, 2008. [DOI:10.3743/KOSIM.2008.25.4.149]

1. Introduction

The lack of empirical studies on users and real-life information seeking processes of MIR systems is a major issue in current Music Information Retrieval (MIR) research (Byrd & Crawford, 2002; Downie & Cunningham, 2002; Futrelle & Downie, 2003). This has led MIR researchers to rely on their assumptions of typical usage scenarios for designing MIR systems (Downie & Cunningham, 2002; Cunningham et al., 2003) rather than empirical user data. Futrelle and Downie (2003) criticized that this is why the current MIR research is weak on application to real users.

This study aims to further our understanding of the information features provided by users in real-life music information queries, particularly focusing on the accuracy of user-provided information. Music information queries here mean the natural language query statements that users expressed while searching for music objects or information about those objects (i.e., metadata), not bounded by the limited set of features provided in currently available catalogs or MIR systems. Previous studies of music information queries contributed in identifying the basic types of real- life music information needs and the types of information features that people use when seeking music information. Nevertheless, our current understanding of the real-life music information queries is still astonishingly poor; we have yet to improve our understanding of the use patterns of those features beyond their frequency of occurrence. For instance, from the previous studies, we know how often each feature appeared in the queries analyzed; however, we do not know how accurately those features were described, and when inaccurate, what kinds of errors were made by users. This clearly demonstrates our limited understanding of the real-life music information queries and calls for further investigation.

The central question of this study is as follows: How accurate are the information features that users provide in the successfully answered queries? Understanding the patterns of the level of accuracy across different information features may help improve the selection of (i) access points, and (ii) the appropriate kinds of search strategies (e.g., generalization of concept, term selection) to be applied to different types of information features.

2. Related Works

In the past few years, a limited number of user studies were conducted in an attempt to address the lack of empirical studies on users and their information seeking processes. These studies employed various research methods including user surveys (Downie, 1994; Lee & Downie, 2004; Taheri-Panah & MacFarlane, 2004), interviews (Cunningham et al., 2003; Taheri-Panah & MacFarlane, 2004; Laplante & Downie, 2006), transaction log analysis (Itoh, 2000; McPherson & Bainbridge, 2001), natural language query analysis (Downie & Cunningham, 2002; Bainbridge et

al., 2003; Lee et al., 2005a), ethnographic observations (Cunningham et al., 2003), and user experiments (Kim & Belkin, 2002; Lee & Moon, 2006). These studies provided invaluable preliminary information on the needs, uses, and information behaviors of the MIR system users, and also helped raise general awareness of the importance of understanding users among MIR researchers.

The accuracy of user-provided information in music information seeking process, however, has not been studied much in the past. In fact, the concept has not been fully explored in the general field of information seeking. Baker and Lancaster (1991) point out that few evaluators have explored how factors like "accuracy (e.g., is the answer to the question recorded somewhere?), the scope (e.g., is the question's subject within the scope of the library?) and stability (e.g., is the question stable or has its answer changed recently?)" of the reference question affect the performance of reference librarians.

Whereas the accuracy of search results has always been regarded as a major factor in evaluating the search performance of IR systems or reference librarians, the accuracy of information in queries/questions (with regards to the content of information beyond syntactical errors such as spelling mistakes) has not been given much attention. The queries used for evaluating IR systems are often artificially derived from the documents in the test collection (e.g., as in TREC) and typically, researchers do not intentionally include incorrect information in the test queries.

Part of the reason why the accuracy of queries has not been studied much is because obtaining a data set suitable for this type of study is difficult. Studies of transaction log analysis typically do not provide information on the users' motivation or the success of the search (Cunningham et al., 2003), thus it is difficult (in some cases, impossible) to determine the accuracy of user-provided information and study the impact of this accuracy in search tasks. Even in the case of real-life queries, the accuracy of user-provided information can only be assessed when the correct answer is known, in other words, there exists a means for obtaining the ground truth information.

That users often bring information to the search that is ambiguous, vague, underspecified, false, and sometimes even inconsistent is already a well-known phenomenon in the library and information science (LIS) field (Jackson & American Library Association, 1958; Swanson, 1972; Lewis, 1987; Dwyer et al., 1991). Similar findings were reported within the MIR domain as well. Bainbridge et al. (2003) analyzed a collection of music information queries posted on an online reference website (Google Answers) and found that users often expressed difficulty in precisely describing various attributes of the sought music object and also expressed uncertainty as to the accuracy of their descriptions. Lee et al. (2005a) also found that users searching for music information in a cross-cultural/multilingual setting were experiencing difficulties with common bibliographic access

points. In the analyzed queries, users mentioned that they could not remember the exact artist names and/or song titles, understand and provide accurate lyric information, or specify the "correct" genre label. Based on what we know from the studies of reference and catalog uses, it is likely that some of the features provided by users in these queries contain errors, but we do not know which features they are, and also the nature of these errors (e.g., frequency of inaccurate information for each feature, types of errors).

This study attempts to provide some empirical evidence to help us advance our understanding of one of the less studied variables that affect the search success. The accuracy of user-provided information is especially important for queries searching for a known- item because in a Boolean search statement, even one non-matching term combined using "and" will result in a failure of information retrieval (Allen, 1989). It may be true that as long as some features are correct, the inaccuracies of other features do not matter. In this case, it would be valuable to know which features were accurate or not.

3. Research Design

The particular website selected for collecting real-life queries was the Google Answers (GA) website, an online reference service provided by Google. The rationale for selecting Google Answers is that the amount of information the users provide in their queries and the quality of the replies are impressive because it is a fee-based service (Katz, 2002/ 2003). Upon receiving the approval from the Institutional Review Board (IRB), all the queries posted under the music category on the GA website (2208 queries) were collected on April 27, 2005.

Among the various types of queries, the queries that are of interest to this study were those in which the user is trying to identify particular music object(s) and/or artist(s). These queries were identified by the author manually examining the query data. The answers and comments were also collected since they are necessary to identify successfully answered queries among the whole pool of query documents. "Successfully answered" queries refer to the cases where the inquirers verified the answer to be correct and explicitly stated so in their feedback/rating.

Although there are numerous features that can be analyzed, only a limited number of features were examined due to the nature of the study which is that only certain features can be determined as accurate or not (at least, arguably). The accuracy of six different features was analyzed: *Lyric, Title, Person name, Corporate Name, Date,* and *Place reference.*

4. Results

4.1 Accuracy of Six Features

The total number of queries that received an

answer from the whole sample was 1,062 (48%). However, the number of queries known to have been successfully answered was only 266 as many answers were never verified and/or rated by the inquirer. All the instances of the six features in successfully answered queries were marked up and qualitatively examined.

The tables providing a full comparison of the user-provided information and the correct information features identified in the answers can be found in Appendix I of Lee (2008). Due to the limited space, only a few selected examples are presented in this work. The following sub-sections provide discussions on each feature.

4.1.1 Lyric

Of all the successfully answered queries that had lyric information, the accuracy of 129 cases could be verified, but only 25 of them were accurate. All the instances of lyric information were manually analyzed in order to understand the kinds of mistakes people make in providing lyric information. By analyzing all the instances, the following 11 categories of user errors emerged from the data (Table 1). The Other category included errors made in using articles (e.g., "a" vs. "the"), relative pronouns (e.g., "who" vs. "that"), singular/plural, phrases in a reverse order, British/American spelling, slang (e.g., "your" vs. "ya") and so on. The

〈Table 1〉 Types of errors in lyric information

Error Type	Example(s)
Missing one or more word(s)	user: alone without you correct: while I'm alone without you
Additional word(s)	user: I said I'm cool correct: I'm cool
Misspelling word(s)	user: surley there is a mine correct: surely there is a mine
Errors in using contractions	user: scarecrow's waving correct: scarecrow is waving
Errors in using pronouns	user: you've won the league correct: they've won the league
Errors in using prepositions	user: sun shines up the mountain correct: sun shines on the mountain
Errors in tense	user: can't even remember correct: couldn't even remember
Use of similar sounding word(s)	user: your conscious correct: you're gorgeous
Use of different words with the similar sense and/or in the same semantic category	user: feel my hand correct: take my hand
Errors in using different words (that do not sound similar or have similar meaning)	user: South africa correct: love africa
Other	user: come on come on correct: c'mon c'mon

errors were categorized relying on the author's best judgment; however it is acknowledged that some of these may still be contestable and therefore all the raw data is provided in Lee (2008) for cross-examination.

Figure 1 shows the distribution of the counts of queries containing each type of error found in the analyzed queries. The most surprising observation here is that there were more queries where users provided words and phrases that had similar senses (meanings) and/or in the same semantic category as the correct lyrics than the ones where users provided words and phrases that sounded similar. This is especially surprising since we typically think confusion between similar sounding words mainly account for misheard lyrics.

More examples of errors using words with similar senses (or in the same semantic category) are provided below:

User (U): landing in LA
Correct (C): coming into LA

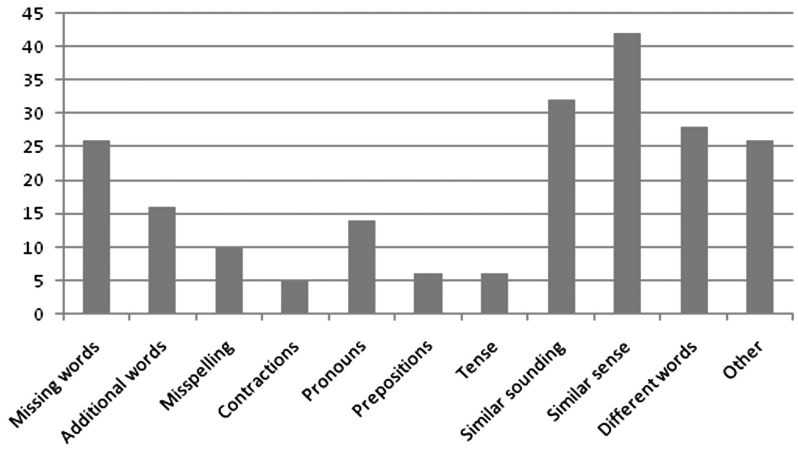
U: nothing's gonna stop me now
C: nothing's gonna hold me back

U: floating up in a cardboard sky
C: sailing over a cardboard sea

U: bullet in my back
C: bullet in my shoulder

U: be true with her eyes of blue
C: be true to her lips of blue

The meaning of the original lyrics are preserved in the first three examples although the users expressed them in different phrases. In some cases, different body parts were expressed as shown in the last two examples.



〈Figure 1〉 Frequency of occurrences of each type of error in lyrics

4.1.2 Title

Of all the successfully answered queries that contained title information, 106 cases could be verified as either accurate or not, and 36 of them contained some errors. Misspelling, omitting, adding and substituting the right words with similar sounding words were common mistakes made by users. The cases where the user- provided titles were significantly different than the correct titles (e.g., "I Am Canada" vs. "Gotta hear you say it too", "Bullet in my back" vs. "Wanted man") were the ones in which users mistook a part of the repeated chorus as the title. There were also some questionable cases with regards to judging the accuracy of the title information such as the titles of TV commercials. These are different in nature from other titles as sort of an ad-hoc title, not something that was officially labeled as such. Table 2 presents a few examples showing the discrepancies between the user-provided titles and correct titles.

4.1.3 Person Name

The accuracy of 88 cases could be verified and 25 of them turned out to be incorrect. There were quite a few cases that were difficult to judge if the user's information was correct or not, since users provided names of artists who were similar or artists that were popular around the same time as the one that they were actually looking for. Another type of difficult case was where users list several options (e.g., "is either Carolyn or Caroline, Hillyer or Hilyer") and one of them turned out to be correct. In such cases, the user is technically not "incorrect" but it is also somewhat questionable if we can say the user was "accurate." Most errors with person names were spelling mistakes and occasionally a user mistook a similar sounding artist to be the correct one. An interesting use of title and person name was when users provided title and artist of a song that was played on the radio or other media before/after the song they were actually seeking. Table 3 presents examples of the

〈Table 2〉 Examples of discrepancies between user-provided titles and correct titles

User-provided titles	Correct titles
Toccatarentella, or something like that.	Tarantella
Raise Bild	Reisebilder
Michaud est monte dans une peuplier (or possibly "Michaud est tombee")	Michaud Est Tombe
I'm cleaning out my closet	Cleaning out my closet
Miss You Blue	Misty Blue
crone woman	Night Woman
The man from Iwo Jima	The Girl From Ipanema
'afro' something	Wings of Dawn
Skinnamarink the Sergeant	Skin-a-ma-link the Sergeant
The Clap Clap Song	The Clapping Song

〈Table 3〉 Examples of discrepancies between user-provided person names and correct person names

User-provided person names	Correct person names
I thought it was Lisa Lobe, but now not sure - anyway it sounded like her.	Juliana Hatfield
…either Little Anthony & the Imperials or Smokey Robinson & the Miracles …	The Royalettes
Boston	Kansas
A Thousand Suns	A Million Sons
Ambrosia	Deliverance
The Groovy Grubworms	Harlow Wilcox & the Oakies
McNally??	Peter McNeeley
Laura? Lora? Laurie?	Laurie
Boss Hog	Crunt
…middle names "Hagerich" and "Viliers"…	Edvard Hagerup Grieg, Charles Villiers Stanford

discrepancies between the user-provided person names and correct person names.

4.1.4 Corporate Name

There were 12 instances of corporate names found in the successfully answered queries. Although the corporate names were not used as often as other features, almost all of them (except one case) were correct names. The only incorrect case was the one in which the user expressed that he/she was uncertain about the information. Most of them had to do with the corporate names for commercial advertisements that users saw on TV or the names of record label companies. In the only case in which the user expressed uncertainty about his/her information, the correct label was Warner Bros, not Sony as specified by the user.

4.1.5 Date (Time)

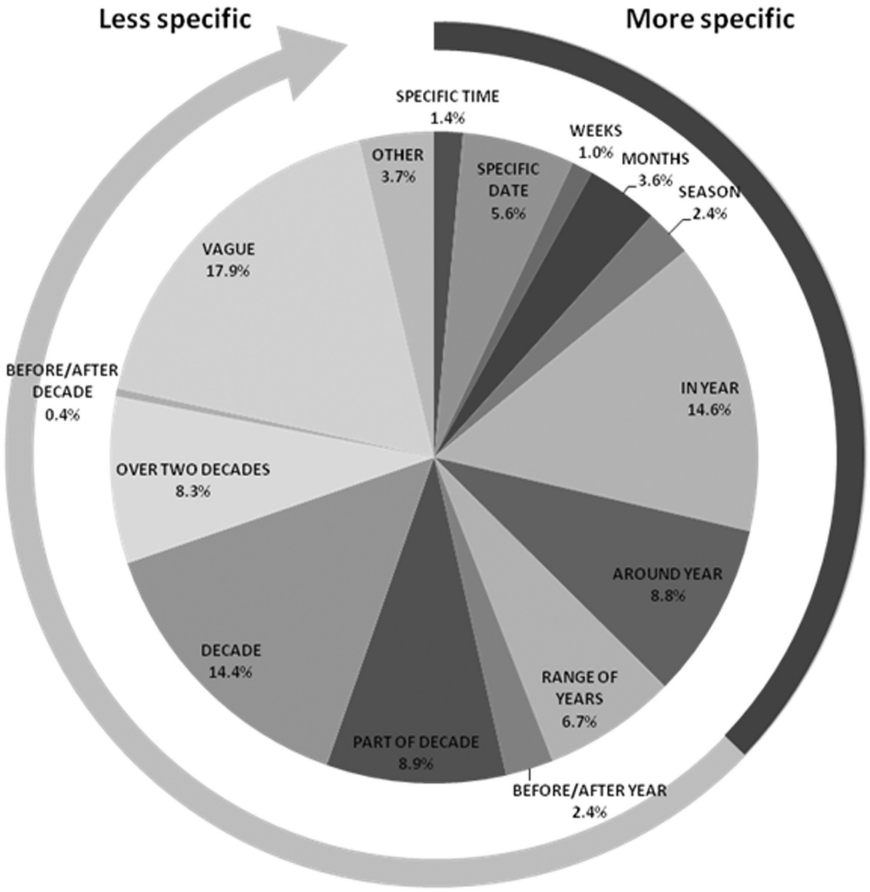
Information related to the time dimension was prevalent throughout the analyzed queries. The specificity of this information ranged from vague (e.g., "old", "recent") through a specific decade, to a particular year, and sometimes even to a specific date and time. 188 of the 1051 unique instances of the date feature were vague temporal references. When users did specify date information more specifically than "old" or "recent," most users specified a particular year (153 out of 1051) or a particular decade (151 out of 1051) [See Figure 2].

Determining the accuracy of the date information presented some difficulties. Many of the date instances had to do with the date when the user had encountered the music which cannot be verified, except for a few cases where users set a specific range of years. Also in the cases where the users provided vague terms such as "old" or "recent", it is tricky to set the boundary of the time periods that will count as old or recent music. Even when the user provides specific time period in their descriptions, often it is of a very wide timeframe (e.g., a decade, over several decades). Therefore

a few of them were technically incorrect, but many of them were rather vague. A total of 87 cases could be verified and 11 were determined to be incorrect.

4.1.6 Place reference

Similarly to the date information, users' references of places where they encountered the music could not be verified except when the performance was a one-time-only event and the information about the event could be tracked down. When the place reference could actually be verified, however, the users' information was correct most of the time. Of the 36 cases that could be verified, only 4 were incorrect. In three cases, the users mistook a British singer for Hungarian, a Swedish for British, and a Romanian boy for Dutch. The other user thought a group from L.A. were Nicaraguan.



〈Figure 2〉 Distribution of date with regards to the level of specificity

5. Discussion

5.1 Potential Use of Inaccurate Information in Searches

Table 4 presents the summary of the accuracy of the six selected information features. The second column shows the total number of queries which contained the particular feature and were successfully answered. The third column shows the number of queries in which the particular information feature could be verified to be correct or not based on the information given in the answer. The fourth column shows the number of queries that had the particular feature but the accuracy of the information could not be verified either due to the nature of the information itself (e.g., date of user's contact with the music, place where the user heard the music) or a lack of information in the answer. The next two columns show the number of queries that contained some sort of errors, and queries that contained correct information without any errors. The proportions of queries with and without errors over the total number of verifiable queries are given inside the parentheses.

Among the features analyzed, the lyric feature had significantly more errors (80.6%) than the rest. Not only did people often mishear lyrics and have trouble remembering the exact words and phrases, but also since the lyric feature typically contains a lot more text than other features, the possibility for error tends to be larger, thus this is not very surprising. One interesting insight from this analysis is that the inaccuracy of users' information might not always be such a negative thing, as demonstrated in the example below:

> *Subject: I need to find the title and artist of a song I heard on the radio.*
> I only caught some of the lyrics. I only heard it once. I am not sure that the wording is absolutely correct. IT was a female singer, very plain, beautiful, slow. "there will be no black flag on my door" "I am in love" "I will go down with vengence (sic)"

The correct answer to this query identified by the intermediary and verified by the inquirer was "White Flag" by Dido. Note the errors in the

〈Table 4〉 Overall accuracy of selected features

Feature	Total number of queries	Verified cases	Unverifiable cases	Queries with errors	Queries without errors
Lyrics	145	129	16	104 (80.6%)	25 (19.4%)
Title	123	106	17	36 (34.0%)	70 (66.0%)
Person name	108	88	20	25 (28.4%)	63 (71.6%)
Date	170	87	83	11 (12.6%)	76 (87.4%)
Place reference	66	36	30	4 (11.1%)	32 (88.9%)
Corporate name	12	12	0	1 (8.3%)	11 (91.6%)

user-provided lyric information by comparing it with the correct lyrics below:

I will go down with this ship
And I won't put my hands up and surrender
There will be no white flag above my door
I'm in love and always will be

Naturally one would think that the chance of finding the right song with incorrect or incomplete clues is very low, thus it is unlikely that those imperfect queries will be successfully answered. Yet they perform surprisingly well in human mediated searches. In the case of GA service, we can point to numerous individual examples demonstrating that a successful search is possible even with flawed queries.

This idea may be unorthodox as it is quite opposite from the general attitude toward inaccurate information in the field. However it seems like, at least in certain situations, other aspects such as consistency may be more important than accuracy. Finding music with misheard lyric information is one such case, but situations like this are not limited to just search and retrieval, and may also be found in other areas in information science. For instance, machine learning algorithms can sometimes perform better on bad data that contains numerous, consistent and systematic OCR errors than clean data with fewer, but more varied, errors. Future studies will be required to further explore the use of incorrect information in real-life information seeking contexts.

5.2 The Role of Pragmatics in Understanding Queries

The results of this study led to an interesting question: how is it that imperfect queries riddled with incorrect, incomplete and/or vague information still seem to perform surprisingly well in eliciting correct answers in human intermediated searches? A typical view in the field would be that such queries are simply semantically flawed by inaccurate, approximate, and ambiguous information, requiring various kinds of "correction" and this correction can be made by employing a number of well-known retrieval strategies designed to accommodate such problems (e.g., fuzzy searching, disambiguation, authority control, controlled vocabulary, spellchecking) (Lee & Renear, 2007). However, the analysis of GA queries suggests a different view, showing that human intermediaries responding to the queries are not simply employing analogues to correcting and approximating retrieval strategies. More satisfying explanations of this phenomenon may be found in pragmatics, an area of linguistics.

5.2.1 Two Uses of Definite Descriptions in Pragmatics

The potential for applying the philosophy of language for investigating information retrieval systems has previously been mentioned by researchers such as Blair(2003). There are several concepts and theories that seem particularly relevant such as speech act theory and conversational

implicature (Grice, 1975). In this work, in order to provide an explanation for the success of imperfect queries for eliciting correct answers, the attributive/referential distinction in pragmatics, initially developed by the philosopher Keith Donnellan (1966), is applied. There exist other studies (Birner, 1991; Kronfeld, 1986; Onishi & Murphy, 2002) which have applied Donnellan's distinction to computational models of reference and discourse; however, there has been little application of pragmatics to actual IR interactions with the exception of Ng's adaptation of Habermas (2002).

Donnellan (1966) argues that there are two uses of definite descriptions, providing a famous example of "the man drinking a martini":

> Suppose one is at a party and seeing an interesting-looking person holding a martini glass, one asks, "Who is the man drinking a martini?" If it should turn out that there is only water in the glass, one has nevertheless asked a question about a particular person, a question that it is possible for someone to answer.

This illustrates the referential use of the definite description. When the description is used referentially, it is merely a device for calling attention to some person or thing which may or may not fit the description in question. When the phrase is used attributively, however, the referent of the description is whoever or whatever fits that description. Donnellan offers the following example of a use that is ineluctably attributive:

> we are told that someone has laid a book on our prize antique table, where nothing should be put. The order, "Bring me the book on the table" cannot now be obeyed unless there is a book that has been placed on the table. There is no possibility of bringing back a book which was never on the table and having it be the one that was meant.

With this observation, it is possible to explain some of the limitations of IR systems that most IR systems treat all definite descriptions as attributive, even when they are intended as referential (Lee & Renear, 2007). As Donnellan (1966) says "when a definite description is used attributively in a command or question and nothing fits the description, the command cannot be obeyed and the question cannot be answered." On the other hand a human intermediary can easily identify and accommodate referential use, as in the example below:

> [Q.] Looking for a song. They lyrics go "My Mamma done told me, When I was in knee-socks" It's a jazzy number.
> [A.] ...Well, you were close. The lyrics refer to "knee pants" not "knee socks" and the genre is blues rather than jazz. In fact, the tune is called "BLUES IN THE NIGHT". And yes, it's a very cool song to be sure...

In this example, the human intermediary understood part of the user's description as referential

(i.e., "knee-socks" and "jazzy") and was still able to find the correct song and suggest it to the user as the correct answer. If the user's description was understood attributively, the intermediary would not have been able to suggest this answer as the correct one since part of the information about the song does not match some of the user's description. The intermediary's background knowledge that people often confuse Jazz with Blues probably helped the search in this particular case. Unlike a human intermediary who is capable of flexible reading of description, IR systems that treat all descriptions as attributive will return no results if genre information was used in conjunction with the lyric information in the search statement, since there is no "Jazz" song that satisfies the other condition.

One may still be tempted to see what is going on here as simply a failure of accuracy or precision that is "corrected" by the human intermediary. This temptation, however, must be resisted. For one thing, there may be no correction at all: the human intermediary may have no reason to believe that the description is incorrect even when it is, or they may believe that the description is incorrect, but have no grounds for improvement. In addition, correction may actually reduce the likelihood of a successful outcome, as it is sometimes easier to successfully identify the referent through the incorrect description. But most significantly, as Donnellan points out, the referential use of definite description is entirely independent of descriptive accuracy perse.

Another important distinction between human intermediaries and machines is that unlike retrieval systems, humans routinely exploit a vast background of social facts and conventions, in other words, context metadata (Lee et al., 2005b). Included in this context metadata are the conventions of language, not only the attributive/ referential distinction as described above, but other discourse conventions, common knowledge and expectations as well. Incorrect information such as commonly made mistakes is one of many interesting kinds of context metadata.

Donnellan's attributive/referential distinction in description seems to provide at least part of the explanation for the unexpected success of flawed queries. However, there are many more related questions that need to be further explored in future studies. To name a few, is it possible to help the system distinguish between the two types of descriptions in particular searches? How can we systematically collect, store, and organize the referential descriptions including common misinformation? And how do we incorporate the incorrect information about the objects without compromising information quality?

6. Conclusion

Users made various syntactic and semantic errors in providing the analyzed information features in Google Answers queries. Among the six features examined, the lyric feature was the one with most

user errors and the corporate name feature had the least number of errors. However, due to the limited number of cases that could be verified for analyzing the accuracy of user-provided information, additional studies with a larger number of verifiable cases will be necessary to further validate these results. Donnellan's theory of the two uses of definite description provides a theoretical lens for explaining how queries laden with incomplete and inaccurate information are still often successfully answered by human intermediaries. This shows the importance of pragmatics, in addition to syntax and semantics, in understanding music queries.

References

Allen, B. 1989. "Recall cues in known-item retrieval." *Journal of the American Society for Information Science,* 40(4): 246-252.

Baker, S. L. & F. W. Lancaster. 1991. *The measurement and evaluation of library services* (2nd ed.). Arlington, VA: Information Resources Press.

Bainbridge, D., S. J. Cunningham, and J. S. Downie. 2003. "How people describe their music information needs: a grounded theory analysis of music queries." *Proc. of ISMIR (International Society of Music Information Retrieval) 2003* (pp.221-222). Baltimore, USA.

Birner, B. J. 1991. Discourse entities and the referential/ attributive distinction. Paper presented at the 65th Meeting of the Linguistic Society of America, Chicago.

Blair, D. C. 2003. "Information retrieval and the philosophy of language." *Annual Review of Information Science and Technology,* 37: 3-50.

Byrd, D. and T. Crawford. 2002. "Problems of music information retrieval in the real world." *Information Processing and Management,* 38(2): 249-271.

Cunningham, S. J., N. Reeves, and M. Britland. 2003. An ethnographic study of music information seeking: implications for the design of a music digital library. *Proc. of JCDL 2003* (pp.5-16). Houston, Texas.

Donnellan, K. S. 1966. "Reference and Definite Descriptions." *Philosophical Review,* 75: 281-304.

Downie, J. S. 1994. "The Musifind musical information retrieval project, phase ii: user assessment survey." *Proc. of the 22nd Annual Conf. of the CAIS,* 149-166.

Downie, J. S., and S. J. Cunningham. 2002. Toward a theory of music information retrieval queries: system design implications. *Proc. of ISMIR 2002* (pp.299-300). Paris, France.

Dwyer, C. M., E. A. Gossen, and L. M. Martin. 1991. "Known- item search failure in an OPAC." *RQ,* 31(2): 228-236.

Futrelle, J. and J. S. Downie. 2003. "Interdisciplinary research issues in music information retrieval: ISMIR 2000-2002." *Journal of New Music Research,* 32(2): 121-131.

Grice, H. P. 1975. Logic and conversation. In *Syntax and semantics: Speech acts* (Vol.3, pp.41-58). New York: Seminar Press.

Itoh, M. 2000. Subject search for music: quantitative analysis of access point selection. *Proc. of ISMIR 2000.* Amherst, MA.

Jackson, S. L., and ALA. 1958. *Catalog use study: director's report.* Chicago: ALA.

Katz, B. 2002/2003. "Digital reference: an overview." *The Reference Librarian,* (79/80), 1-17.

Kim, J.-Y., and N. J. Belkin. 2002. Categories of music description and search terms and phrases used by non-music experts. *Proc. of ISMIR 2002* (pp.209-214). Paris, France.

Kronfeld, A. 1986. "Donnellan's distinction and a computational model of reference." *Proc. of the 24th Meeting of the Association for Computational Linguistics,* 185-191.

Lancaster, F. W. 1998. *Indexing and abstracting in theory and practice* (2nd ed.). Champaign, IL: University of Illinois.

Laplante, A., and J. S. Downie. 2006. "Everyday life music information-seeking behaviour of young adults." *Proc. of ISMIR 2006* (pp.381-382). Victoria, Canada.

Lee, J. H. 2008. *Analysis of information features in natural language queries for music information retrieval: use patterns and accuracy.* Champaign, IL: University of Illinois at Urbana-Champaign.

Lee, J. H., and J. S. Downie. 2004. Survey of music information needs, uses, and seeking behaviours: preliminary findings. *Proc. of ISMIR 2004* (pp.441-446). Barcelona, Spain.

Lee, J. H., J. S. Downie, and S. J. Cunningham. 2005a. Challenges in cross-cultural/ multilingual music information seeking. *Proc. of ISMIR 2005* (pp.1-7). London, UK.

Lee, J. H., X. Hu, and J. S. Downie. 2005b. Q&A websites: rich research resources for contextualizing information retrieval behaviors. *Proc. of the ACM SIGIR 2005 Workshop on Info. Retrieval in Context* (pp.33-36). Salvador, Brazil.

Lee, J. H. and A. Renear. 2007. How incorrect information delivers correct search results: a pragmatic analysis of queries. *Proc. of the 70th ASIS&T Annual Meeting.* Milwaukee, Wisconson.

Lee, Y. J. and S. B. Moon. 2006. A user study on information searching behaviors for designing user-centered query interface of content-based music information retrieval system. *Journal of the Korean Society for information Management,* 23(2): 5-19.

Lewis, D. W. 1987. "Research on the use of online catalogs and its implications for library

practice." *Journal of Academic Librarianship,* 13(3): 152-156.

McPherson, J. R. and D. Bainbridge. 2001. Usage of the MELDEX digital music library. *Proc. of ISMIR 2001*(pp.19-20). Bloomington, IN.

Ng, K. B. 2002. "The applicability of universal pragmatics in IR interaction: a pilot study." *Information Processing and Management,* 38: 237-248.

Onishi, K. H., and G. L. Murphy. 2002. "Discourse model representation of referential and attributive descriptions." *Language and Cognitive Processes,* 17: 97-123.

Swanson, D. R. 1972. *Requirements study for future catalogs. Final report to the National Science Foundation.* Chicago: The University.

Taheri-Panah, S., and A. MacFarlane. 2004. "Music information retrieval systems: why do individuals use them and what are their needs?" *Proc. of ISMIR 2004* (pp.455-460). Barcelona, Spain.

University of Chicago. Graduate Library School, Swanson, D. R., Vaughan, D. K., & Cooper, W. S. (1968). *Requirements study for future catalogs. Progress report no. 2.* Chicago.