Priming and Actions: An Analysis in Conversational Search Systems

Xiao Fu
University College London
London, UK
xiao.fu.20@ucl.ac.uk

Aldo Lipani
University College London
London, UK
aldo.lipani@ucl.ac.uk

ABSTRACT

In order to accurately simulate users in conversational systems, it is essential to comprehend the factors that influence their behaviour. This is a critical challenge for the Information Retrieval (IR) field, as conventional methods are not well-suited for the interactive and unique sequential structure of conversational contexts. In this study, we employed the concept of Priming effects from the Psychology literature to identify core stimuli for each abstracted effect. We then examined these stimuli on various datasets to investigate their correlations with users' actions. Finally, we trained Logistic Regression (LR) models based on these stimuli to anticipate users' actions. Our findings offer a basis for creating more realistic user models and simulators, as we identified the subset of stimuli with strong relationships with users' actions. Additionally, we built a model that can predict users' actions.

CCS CONCEPTS

• **Information systems → Users and interactive retrieval**; Evaluation of retrieval results.

KEYWORDS

Dialogue Systems, User Modeling, Priming Effect

ACM Reference Format:

Xiao Fu and Aldo Lipani. 2023. Priming and Actions: An Analysis in Conversational Search Systems. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23), July 23–27, 2023, Taipei, Taiwan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3539618.3592041

1 INTRODUCTION

Conversational search systems (CSSs) are widely discussed and recognized as an established area within the Information Retrieval (IR) community. Although many studies and industry efforts are directed towards enhancing user experience, the method of assessment remains restricted [2, 6, 23]. Due to the interactional nature of CSSs, traditional metrics cannot capture user satisfaction [6, 19, 21, 22, 24]. Consequently, the industry primarily depends on online evaluation techniques for assessing CSSs [14].

SIGIR '23, July 23–27, 2023, Taipei, Taiwan

ACM ISBN 978-1-4503-9408-6/23/07.
https://doi.org/10.1145/3539618.3592041

Most traditional search systems treat queries as separate entities, but in conversational scenarios, context becomes crucial [8, 17, 24]. This implies that the current user's inquiry also relies on preceding inquiries and responses. An irrelevant response might be considered invalid in a search system, while in the case of CSSs, it can impact subsequent inquiries, leading the user to pose additional questions to refine their search.

One potential approach could be utilizing user simulation (US) for CSSs [11]. USs aim to generate data close to the data we would collect from users. The benefits of USs include: 1) The ability to operate without actual data. 2) The capability to predict the behaviour of users. 3) It is suitable for both developing and assessing CSSs.

One question that remains unanswered is what motivates users to take certain actions. In this paper, we focused on priming effects. Priming effect refers to an unconscious influence of past experience on current performance or behaviour [4, 30]. Tulving et al. [34] conducted an experiment to demonstrate the effect of priming where participants saw a list of 96 words. Then the participants were asked to finish several tasks, including completing graphemic word fragments 1 hour later and seven days later than having seen the list. This study shows that previous tasks help accelerating later tasks. The priming effect is widely used in social marketing. Fukawa [16] summarized that affects consumers' behaviour and judgments with an example: When an individual is exposed to primes, e.g., wholesome and nourishing, it can activate related concepts, e.g., being healthy, which makes them more susceptible to purchasing corresponding products, e.g., vegetable juice.

Although there is a long history of exploring priming effects, it is still a developing topic in CSSs. Church [7] proposed an adaptive language model for lexical adaptation to depict priming effects, where each document is divided into prime and target, where the prime will influence the target. In many studies, correlations between primes and targets are discovered, such as active/passive, verbal particle placement, etc. [9, 18, 20, 28, 33]. Reitter and Moore [28] demonstrated that lexical and syntactic repetition can predict task-oriented dialogues' success. While it is fruitful in modelling the conversation, the correlation between users' actions and the influence of priming effects is not yet clearly understood.

The objective of this paper is to find a relationship between context-based stimuli and the actions performed by users. We first modelled the interaction between users and systems. We then adapted priming effects from the psychology community, which have the potential to depict users' actions, to analyze our datasets. After that, we analyze the correlations between priming effects and users' actions. Finally, we proposed a model to predict users' actions.

Our contributions can be summarized as follows:

- An analysis of correlations between stimuli of priming effects and users' actions.
- A logistic regression (LR) model to predict users' actions.

Our findings establish a basis for creating reliable user simulations and better evaluation metrics for CSSs.

2 RESEARCH QUESTIONS

In this paper, we aim to answer the following research questions:

RQ1. *Are there relationships between stimuli and users' actions?*

To answer this question, we will calculate the correlation coefficients between various stimuli adapted from priming effects and users' actions. This is crucial for understanding the priming effects in CSSs.

RQ2. *What is the performance of models built on these stimuli?*

To answer this question, we will train LR models based on stimuli in **RQ1**. This will demonstrate if these stimuli include enough information to predict users' actions.

3 PRIMING EFFECTS

Researchers have identified various types of priming effects that have distinct effects on user behaviour. For example, some stimuli will speed up the processing [25], which is called positive priming effects, as well as negative priming effects will slow down the processing. This study focuses on different priming effects and corresponding stimuli, with consideration given to various dimensions such as depth of history, scope in history, and scope in the current turn.

In general, three dimensions are depth of history, scope in history, and scope in the current turn. Depth of history refers to the number of turns before the current turn when the stimuli are calculated. In this study, we chose 1, 2 and 5 turns. Scope in history and last turn refers on which part (i.e., queries, replies or both) we calculate the stimuli. The scope has r for replies, q for queries and rq for both.

Repetition priming effects. Repetition priming effects refer to the response when the stimuli are repeatedly presented [30]. Low-frequency words tend to have stronger repetition priming effects than high-frequency words [12]. They also discovered two components of normal repetition priming effects: a short-term effect independent of frequency and a long-term effect dependent on frequency. In this study, we consider two types of repetition: repetition of each part of the tokenized text and repetition of each noun phrase. According to different scopes, the number of repetitions will be counted as the stimuli.

Semantic priming effects. Foss [13] concluded semantic priming effect is "awakening" from the context, where stimuli will trigger the same semantic category. In their study, Blank and Foss [5] provide an example for semantic priming effects: *nurse* is a semantic prime of *doctor* rather than of *butter*. When a stimulus related to one word is present, not only the word but also related words are "awakened". For stimuli of these priming effects, Word2Vec [26] is used to calculate the semantic similarity between words. We use three methods to summarize the top 5 similarities in each scope. They are: Taking the average, Summing up and Taking the max.

Table 1: Properties of each dataset. Size refers to the total number of conversations, $\bar{l}$ refers to their average length, σ refers to its standard deviation. The last two columns indicate if the follow-up and switch-topic actions are labelled.

Dataset	Size	$\bar{l}$	σ	Answer Style	Follow-up	Switch-topic
TopiOCQA (Topi)	3920	13	3.3	Free-form	✗	✓
FAITHDIAL (FD)	5649	4.5	0.5	Free-form	✗	✗
QReCC	13598	5.9	2.5	Free-form	✗	✗
ORConvQA (ORC)	5644	7.2	2.1	Free-form	✓	✗
TREC CAsT 2021 (TREC)	26	9.2	1.7	Passage	✗	✗
EvalCran (Cran)	131	5.4	2.6	Passage	✗	✗
AllMixed (AllMix)	156	7.5	3.7	Mixed	✗	✗

Affective priming effects. Affective priming effects involve assessing people, ideas, objects, and goods not solely based on their physical characteristics but also based on their emotional context. Studies on affective priming effects typically involve presenting positive, neutral, or negative cues before a stimulus to influence how it is evaluated or responded to. Affective priming effects may be more powerful and widespread when the cue is barely noticed by the individual [29]. In this study, we will measure polarity and subjectivity via TextBlob[1] and similarly summarize them as we do for the semantic priming effects.

4 USERS' ACTIONS

In this study, we categorize users' actions when interacting with CSSs into three types, namely **Stopping**, **Following up** and **Switching topic**.

Stopping. Stopping occurs when users decide to terminate a conversation, and it typically marks the final turn of the exchange. This action holds significance in measuring various effects, such as the principle of least effort [35] and the recency effect, wherein individuals tend to start recalling information with the most recent items.

Following up. Conversational sessions frequently involve follow-up queries that rely on prior interactions, taking into account the absence of certain context and references to previously mentioned subjects [27]. Users ask follow-up queries to correct their search space and seek better answers.

Switching topic. According to Stede and Schlangen [32], the inquisitive user in an ongoing interaction may develop an interest in additional, yet related topics based on the information presented in the responses. This phenomenon is commonly referred to as topic-switching behaviour, which is frequently observed in information-seeking conversations, particularly when utilizing search systems for information gathering [31].

5 DATA

In this study, we use 6 datasets to analyze the relationship between priming effects and behaviours of users in CSSs. The datasets used are shown in Table 1: TopiOCQA (Topi) [1], FAITHDIAL (FD) [10],

[1]https://textblob.readthedocs.io/

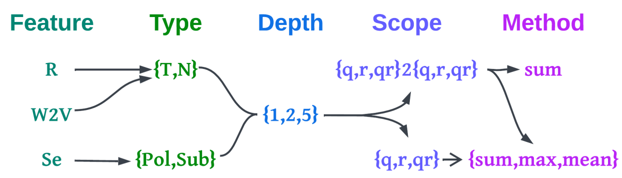

Figure 1: Structure of stimuli.

Table 2: Scores of LR models on datasets, while AUC refers to the area under P-R curves.

Action	Dataset	Precision	Recall	F1	AUC
Stopping	AllMix	0.844	0.574	0.684	0.835
	Topi	0.977	0.951	0.963	0.995
	FD	0.998	0.999	0.998	1.000
	TREC	1.000	0.625	0.769	0.800
	Cran	0.795	0.875	0.833	0.883
	QReCC	0.953	0.997	0.974	0.990
	ORC	0.965	0.979	0.972	0.997
Following up	ORC	0.656	0.779	0.712	0.698
Switching topic	Topi	0.569	0.162	0.252	0.476

QReCC [3], ORConvQA (ORC) [27], TREC CAsT 2021 (TREC) [2], and EvalCran (Cran) [15].

To broaden the range of scenarios, we selected the following datasets. Some datasets, like Topi and FD, feature agent responses in free-form, while others consist of passages from documents. Additionally, the datasets possess distinct characteristics. For instance, the average length of conversations in Topi is 13, with a standard deviation of 3.3, while FD has an average length of 4.5, with a standard deviation of 0.5. Consequently, Topi conversations tend to be lengthier with a broader distribution, while in FD, most conversations end after 4 or 5 turns. In datasets like Topi and FD, the agent's answers are composed of free-form responses, making the detection of repeated retrieved documents challenging since the same content can be expressed differently.

To analyze the following-up action, we chose ORC, where follow-up questions are labelled. In this dataset, 53% of questions are follow-up questions. To analyze the switching-topic action, we chose Topi, where switching-topic queries are also labelled. In Topi, 27% of questions are switching-topic queries.

We created a new dataset AllMix based on the above original datasets. As the minimal size of the seven datasets is 26 (TREC), in AllMix, we randomly picked 26 samples from each dataset to keep the balance of different datasets.

6 EXPERIMENTS AND FINDINGS

6.1 Stimuli vs Actions: Correlation

To answer **RQ1**, we calculated correlation coefficients between stimuli and users' actions on various datasets. As mentioned in Section 3, we took into account stimuli with diverse scopes and

[2]https://www.treccast.ai/

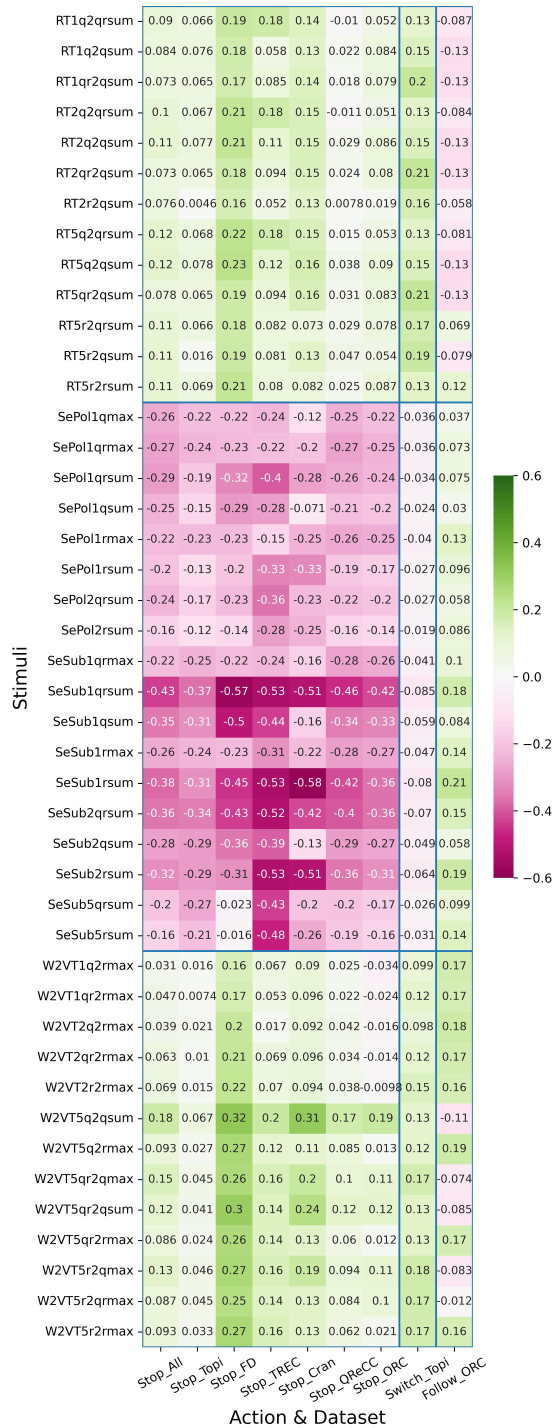

Figure 2: Spearman's correlation coefficient between stimuli and actions in different datasets.

conditions. In this study, there count a total of 270 stimuli. The labels on the y-axis in Figure 1 exemplify the names of such stimuli.

Stimuli with different features begin with different names. For instance, *W2V* denotes the semantic similarity assessed by Word2Vec, *R* represents repetition, and *Se* refers to sentiment. Then comes the type of stimuli, where *T* refers to stimuli computed from tokenized lists, *N* refers to noun lists, *Sub* refers to subjectivity, and *Pol* refers to polarity. The following number denotes the depth of the history. After that, the scope is indicated. For W2V and R, the scope consists of two parts and is written as *h2c* where *h* refers to the scope of history, and *c* refers to the scope of the current turn. For Se, there is only one part that refers to the scope of history. The scopes have three types, as mentioned in Section 3. Finally, the method used to compute the stimuli is indicated, which includes *sum*, *max*, and *mean*.

In this section, we utilized Spearman's rho to gauge the correlation coefficients between stimuli and users' actions across various datasets. We processed every turn of each conversation for each stimulus and concatenated them into one stimulus list for each corresponding dataset. We then calculated the correlation coefficients with corresponding actions at each turn. Figure 2 shows the result. In this figure, we selected the top 10 stimuli for each dataset and action based on the absolute value of their Spearman's rho without any duplication. There are a total of 44 stimuli in the figure. According to the figure, we have the following findings:

First, sentiment-based stimuli have a reasonable association with stopping points. For example, SeSub1qrsum has a negative correlation with stopping in most datasets. It indicates that this stimulus has the potential to prolong the conversation.

Second, stimuli based on the same feature tend to have a similar relationship with the same action. In Figure 2, there is a notable trend that almost all the similar stimuli play a similar role for the same action. For example, most of the stimuli based on sentiment play a negative role in stopping under most conditions. In contrast, in switching topics, this group of stimuli always have a smaller strength compared with other stimuli. We can also observe this trend in stimuli based on W2V in stopping and switching topics, as well as in stimuli based on repetition in stopping, switching topics and following up.

Finally, only a few stimuli have a weak correlation with switching topics and following up. RT1q2qsum, RT2qr2qsum, and RT5qr2qsum are the only three stimuli that have a correlation coefficient larger than 0.2 for switching topics, while SeSub1rsum is the only stimulus with a correlation coefficient larger than 0.2 for following up. This suggests that predicting these two actions based on stimuli is challenging, unlike predicting stopping points.

6.2 Stimuli vs Actions: LR models

To address **RQ2**, LR models were trained for three actions using the 270 stimuli outlined in Section 3. Each dataset was randomly split into 70% training and 30% test sets for LR model training. At each turn of the conversation, the LR models received 270 features of stimuli as input and the action to be taken as the target. The LR models consisted of a Min-Max scaler for normalization and an LR layer for classification.

Table 2 displays the scores of LR models in the corresponding test sets. The LR models exhibit good performance in all datasets for stopping, poor scores in Topi for switching topics, and fair scores in ORC for following up. This indicates that the 270 stimuli used in this study can represent stopping behaviour and capture some information about asking follow-up queries. However, predicting switching topics solely based on repetition, semantic similarity, and sentiment is difficult.

The performance of LR models also aligns with the correlation coefficients distribution presented in Section 6.1, where the stopping action exhibited stronger relationships than the other two actions, as LR models have the best performance in predicting it.

Different from correlation coefficients, the weights of the LR models focus on the stimuli aggregated using the mean. The top 3 stimuli, sorted by the accumulated reciprocal of the rank for the stopping action, are SeSub2qrmean, SeSub2rmean and SeSub2qmean; for the switching topic action are SeSub5qmean, SeSub5qrmean and W2VT1q2qmean, and; for the following up action are: W2VT5q2rmean, W2VT5qr2qrmean and SeSub1qmean.

7 DISCUSSION AND CONCLUSION

This paper examined different stimuli based on several priming effects. Initially, we combined six original datasets to create a diverse dataset. Next, we adjusted priming effects to match the CSS setting. Then, we generated 270 stimuli based on varying feature, type, depth, scope, and method. Following that, we computed and scrutinized the correlation between stimuli and users' actions. Finally, we employed logistic regressions on seven datasets with 270 stimuli and evaluated their effectiveness in predicting users' actions.

Our results show:

- Stimuli based on sentiment have reasonable relationships with the stopping action.
- There is a limited correlation between certain stimuli and the act of switching topics and following up.
- Stimuli based on the same feature share a similar relationship with the same action.
- LR models based on repetition, semantic similarity and sentiment stimuli are capable of predicting the stopping and follow-up actions but cannot predict when the topic of the conversation changes.
- Different from correlation coefficients, LR models prefer stimuli aggregated using the mean operator rather than sum or max.

Our study analyzed how different stimuli, such as repetition, semantic similarity, and sentiment influenced the three user's actions in various datasets. According to our results, LR models that utilize these stimuli are capable of predicting the stopping of the conversation and the asking follow-up questions, but they cannot anticipate a change of topic. This would be beneficial in upcoming research involving simulating user behaviour in CSSs.

This study has identified two limitations that warrant attention in future research. Firstly, incorporating a more diverse set of priming effects beyond just repetition, semantic similarity, and sentiment may be beneficial, as these factors only represent a subset of potential priming effects. If labelled data is available, introducing additional priming effects and stimuli could improve future studies. Secondly, since this study was restricted to using only one dataset for the topic-switching and follow-up actions, using more datasets would give us more confidence in the results of this paper.

REFERENCES

[1] Vaibhav Adlakha, Shehzaad Dhuliawala, Kaheer Suleman, Harm de Vries, and Siva Reddy. 2022. TopiOCQA: Open-domain Conversational Question Answering with Topic Switching. *Transactions of the Association for Computational Linguistics* 10, 468–483. https://doi.org/10.1162/tacl_a_00471 arXiv:https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00471/2008126/tacl_a_00471.pdf

[2] Avishek Anand, Lawrence Cavedon, Hideo Joho, Mark Sanderson, and Benno Stein. 2020. Conversational Search (Dagstuhl Seminar 19461). *Dagstuhl Reports* 9, 11 (2020). https://doi.org/10.4230/DagRep.9.11.34

[3] Raviteja Anantha, Svitlana Vakulenko, Zhucheng Tu, Shayne Longpre, Stephen Pulman, and Srinivas Chappidi. 2021. Open-Domain Question Answering Goes Conversational via Question Rewriting. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (2021).

[4] John A Bargh and Tanya L Chartrand. 2014. The mind in the middle: A practical guide to priming and automaticity research. (2014).

[5] Michelle A Blank and Donald J Foss. 1978. Semantic facilitation and lexical access during sentence processing. *Memory & Cognition* 6, 6 (1978), 644–652.

[6] Jason Ingyu Choi, Ali Ahmadvand, and Eugene Agichtein. 2019. Offline and Online Satisfaction Prediction in Open-Domain Conversational Systems. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (Beijing, China) *(CIKM '19)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3357384.3358047

[7] Kenneth Church. 2000. Empirical estimates of adaptation: The chance of two noriegas is closer to p/2 than p2. In *COLING 2000 Volume 1: The 18th International Conference on Computational Linguistics*.

[8] J Shane Culpepper, Fernando Diaz, and Mark D Smucker. 2018. Research frontiers in information retrieval: Report from the third strategic workshop on information retrieval in lorne (swirl 2018), In ACM SIGIR Forum. *SIGIR Forum* 52, 1, 34–90. https://doi.org/10.1145/3274784.3274788

[9] Amit Dubey, Patrick Sturt, and Frank Keller. 2005. Parallelism in coordination as an instance of syntactic priming: Evidence from corpus-based modeling. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*. 827–834.

[10] Nouha Dziri, Ehsan Kamalloo, Sivan Milton, Osmar Zaiane, Mo Yu, Edoardo M Ponti, and Siva Reddy. 2022. FaithDial: A Faithful Benchmark for Information-Seeking Dialogue. *Transactions of the Association for Computational Linguistics* 10 (12 2022), 1473–1490. https://doi.org/10.1162/tacl_a_00529

[11] Pierre Erbacher, Laure Soulier, and Ludovic Denoyer. 2021. State of the Art of User Simulation approaches for conversational information retrieval. In *CSR-Sim4IR 2021 Causality in Search and Recommendation and Simulation of Information Retrieval Evaluation 2021*.

[12] Kenneth I Forster and Chris Davis. 1984. Repetition priming and frequency attenuation in lexical access. *Journal of experimental psychology: Learning, Memory, and Cognition* 10, 4 (1984), 680.

[13] Donald J Foss. 1982. A discourse on semantic priming. *Cognitive Psychology* 14, 4 (1982), 590–607. https://doi.org/10.1016/0010-0285(82)90020-2

[14] Steve Fox, Kuldeep Karnawat, Mark Mydland, Susan Dumais, and Thomas White. 2005. Evaluating Implicit Measures to Improve Web Search. *ACM Trans. Inf. Syst.* 23, 2 (2005), 147–168. https://doi.org/10.1145/1059981.1059982

[15] Xiao Fu, Emine Yilmaz, and Aldo Lipani. 2022. Evaluating the Cranfield Paradigm for Conversational Search Systems. In *Proceedings of the 2022 ACM SIGIR International Conference on Theory of Information Retrieval*. 275–280.

[16] Nobuyuki Fukawa. 2016. Priming effects on affective preference for healthy products over unhealthy products upon brand exposure. *Social Marketing Quarterly* 22, 1 (2016), 34–53.

[17] Jianfeng Gao, Chenyan Xiong, and Paul Bennett. 2020. Recent Advances in Conversational Information Retrieval. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Virtual Event, China) *(SIGIR '20)*. Association for Computing Machinery, New York, NY, USA, 2421–2424. https://doi.org/10.1145/3397271.3401418

[18] Stefan Gries. 2005. Syntactic Priming: A Corpus-based Approach. *Journal of psycholinguistic research* 34 (08 2005), 365–99. https://doi.org/10.1007/s10936-005-6139-3

[19] Ahmed Hassan, Rosie Jones, and Kristina Lisa Klinkner. 2010. Beyond DCG: User Behavior as a Predictor of a Successful Search. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining* (New York, New York, USA) *(WSDM '10)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/1718487.1718515

[20] Tim Florian Jaeger. 2006. *Redundancy and syntactic reduction in spontaneous speech*. Ph. D. Dissertation. Stanford University Stanford, CA.

[21] Kalervo Järvelin, Susan L. Price, Lois M. L. Delcambre, and Marianne Lykke Nielsen. 2008. Discounted Cumulated Gain Based Evaluation of Multiple-Query IR Sessions. In *Advances in Information Retrieval*, Craig Macdonald, Iadh Ounis, Vassilis Plachouras, Ian Ruthven, and Ryen W. White (Eds.). Springer, Springer Berlin Heidelberg, Berlin, Heidelberg, 4–15.

[22] Jiepu Jiang, Ahmed Hassan Awadallah, Rosie Jones, Umut Ozertem, Imed Zitouni, Ranjitha Gurunath Kulkarni, and Omar Zia Khan. 2015. Automatic Online Evaluation of Intelligent Assistants. In *Proceedings of the 24th International Conference on World Wide Web* (Florence, Italy) *(WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE. https://doi.org/10.1145/2736277.2741669

[23] Julia Kiseleva, Kyle Williams, Jiepu Jiang, Ahmed Hassan Awadallah, Aidan C. Crook, Imed Zitouni, and Tasos Anastasakos. 2016. Understanding User Satisfaction with Intelligent Assistants. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval* (Carrboro, North Carolina, USA) *(CHIIR '16)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2854946.2854961

[24] Aldo Lipani, Ben Carterette, and Emine Yilmaz. 2021. How Am I Doing?: Evaluating Conversational Search Systems Offline. *ACM Trans. Inf. Syst.* 39, 4, Article 51 (2021). https://doi.org/10.1145/3451160

[25] Susanne Mayr and Axel Buchner. 2007. Negative priming as a memory phenomenon. *Zeitschrift für Psychologie/Journal of Psychology* 215, 1 (2007), 35–51.

[26] Tomás Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. In *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1301.3781

[27] Chen Qu, Liu Yang, Cen Chen, Minghui Qiu, W Bruce Croft, and Mohit Iyyer. 2020. Open-retrieval conversational question answering. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*. 539–548.

[28] David Reitter and Johanna D. Moore. 2007. Predicting Success in Dialogue. In *Proceedings of 45th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 808–815.

[29] Mark Rotteveel, Paulien De Groot, Arjen Geutskens, and R Hans Phaf. 2001. Stronger suboptimal than optimal affective priming? *Emotion* 1, 4 (2001), 348.

[30] Daniel L Schacter and Randy L Buckner. 1998. Priming and the brain. *Neuron* 20, 2 (1998), 185–195.

[31] Amanda Spink, Hüseyin Özmutlu, and Seda Özmutlu. 2002. Multitasking information seeking and searching processes. *JASIST* 53 (06 2002), 639–652. https://doi.org/10.1002/asi.10124

[32] Manfred Stede and David Schlangen. 2004. Information-Seeking Chat : Dialogue Management by Topic Structure.

[33] Benedikt Szmrecsanyi. 2005. Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. (2005).

[34] Endel Tulving, Daniel L Schacter, and Heather A Stark. 1982. Priming effects in word-fragment completion are independent of recognition memory. *Journal of experimental psychology: learning, memory, and cognition* 8, 4 (1982), 336.

[35] George Kingsley Zipf. 2016. *Human behavior and the principle of least effort: An introduction to human ecology*. Ravenio Books.