

Johns Hopkins University
Homewood Institutional Review Board (HIRB)

Application for Exemption

Limited types of human participant research are exempt from Federal regulations for the protection of human research participants [45 CFR (Code of Federal Regulations) 46]. Studies in which the **ONLY** involvement of human participants is in one or more of the six exempt categories of research listed in the Federal Regulations and in this form **MAY** qualify for exemption. HIRB is ultimately responsible for verifying whether studies meet the required criteria for exemption. Studies that qualify for exempt status may not be undertaken until HIRB verifies the exemption.

Exemption from federal regulations does not absolve investigators from their ethical obligation to protect the welfare and rights of human participants and, when appropriate, to obtain informed consent.

To determine whether your study qualifies for exempt status and under which category(s), please complete the Exemption Checklist in this document. If all of the research activities in your study fall into one or more of the exemption categories, you may submit an Application for Exemption, which is below. If all the research activities in your study do not fall into one or more of the exemption categories, you must complete and submit an Application for Expedited/Full Board Review.

Please ensure that you have fully completed the application, including providing the required signatures and, when applicable, additional documentation. Complete applications may be mailed, delivered, or e-mailed to HIRB. Only signature pages may be faxed. Faxing complete applications should be avoided except in unusual circumstances. If you have any questions, please contact us.

Homewood Institutional Review Board (HIRB)
151 Krieger Hall
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218
Phone: 410-516-6580
Fax: 410-516-0150
E-mail: hirb@jhu.edu

Johns Hopkins University
Homewood Institutional Review Board (HIRB)

Application for Exemption

PRINCIPAL INVESTIGATOR

Name: Chris Callison-Burch	Degree /Title: PhD
Position: JHU Faculty If Other, Please Specify:	
Department: Computer Science	Mailing Address: 3400 N Charles St CSEB room 226-B Baltimore, MD 21218
Phone: 410 516 3427	
Fax:	
E-mail: ccb@cs.jhu.edu	
Signature: <i>Chris Callison-Burch</i>	Date: Jan 6 2010

STUDENT INVESTIGATOR/CO-INVESTIGATORS (IF APPLICABLE)

<i>Student Investigators must have a faculty member serve as Principal Investigator for all research projects with human participants. For any additional investigators on this study (e.g., students, consultants) please complete a Research Team Members Form.</i>	
Name: Michael Bloodgood	Degree /Title: PhD
Position: Senior Staff If Other, Please Specify:	
Department: Human Language Technology Center of Excellence	Mailing Address: 810 Wyman Park Drive Baltimore, MD 21211
Phone: 410 516 7889	
Fax:	
E-mail: bloodgood@jhu.edu	
Signature: <i>Michael Bloodgood</i>	Date: 1/6/10

ADDITIONAL RESEARCH TEAM MEMBERS

<p>Research team members include anyone — whether affiliated with the Johns Hopkins University or not — who (a) interacts with research participants (e.g., in recruiting, obtaining informed consent, collecting data, or debriefing); (b) prepares materials for use with participants (e.g., letters to potential participants, interview protocols); or (c) has access to data collected from participants.</p> <p><i>Additional research team members on this study?</i> <input type="checkbox"/> No <input checked="" type="checkbox"/> Yes <i>!!attach Research Team Member Form for each</i></p>
--

Johns Hopkins University
Homewood Institutional Review Board (HIRB)

Research Team Member Form

This form is for a single research team member. Use a separate form for each additional team member.

STUDY ID

Study Title: Collection of language data for computational linguistics models with Amazon's Mechanical Turk	
HIRB Application # (leave blank if new application):	
Principal Investigator: Chris Callison-Burch, Michael Bloodgood	
Signature of PI: <i>Chris Callison-Burch</i>	Date: <i>Jan 6 2010</i>

RESEARCH TEAM MEMBER INFORMATION

Name:	Christine Piatko	
Position:	<input type="checkbox"/> JHU Faculty <input checked="" type="checkbox"/> JHU Staff – please specify: <u>JHU/APL on assignment at JHU HLT/COE</u> <input type="checkbox"/> Other – please specify: _____	
Department:	JHU/APL MERC	
Contact Info::	Phone: 410-516-4635	Fax:
	E-mail: christine.piatko@jhuapl.edu	
Mailing Address:	Stieff Building / 810 Wyman Park Drive Baltimore, MD 21211-2840	

TRAINING REQUIREMENT

Training in the Protection & Ethical Treatment of Human Research Participants – Must renew every 3 years	<input checked="" type="checkbox"/> JHU/JHSPH Human Subjects Research Education (Meets requirement for team members from Homewood or JHSPH) <input type="checkbox"/> JHM Human Subjects Research Course (Meets requirement for team members from JHM.) <input type="checkbox"/> Other Institution (Meets requirement only for non-JHU team members and new faculty.) Specify:
Date Completed:	<u>1/18/2007</u> (must be completed within last 3 years !)
Certification:	<input checked="" type="checkbox"/> Attached <input type="checkbox"/> Copy is on file with Homewood IRB

CONFLICTS OF INTEREST (*)

Does the team member's involvement in this study create any foreseeable conflicts of interest? (*) See <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No http://jhuresearch.jhu.edu/Policy_onConflict_of_Interest.pdf for information on JHU Conflict of Interest policies and procedures. *If Yes, please explain any conflict of interest:

PROJECT ROLES AND RESPONSIBILITIES

Please describe the team member's roles and responsibilities on the project: Help design & set up the modality language understanding experiments. Help retrieve the results and monitor for quality control. Help analyze the results and design algorithms to learn from the labeled examples. Assist with evaluating the effectiveness of the computational models and writing about findings.

CITI Collaborative Institutional Training Initiative

Social & Behavioral Research - Basic/Refresher Curriculum Completion Report

Printed on 1/5/2010

Learner: Chris Callison-Burch (username: ccb@cs.jhu.edu)

Institution: Johns Hopkins University

Contact Information Department: Computer Science

Phone: 410 516 3427

Email: ccb@cs.jhu.edu

Social & Behavioral Research - Basic/Refresher:

Stage 1. Basic Course Passed on 12/10/09 (Ref # 3484359)

Required Modules	Date Completed	Score
Belmont Report and CITI Course Introduction	09/07/09	3/3 (100%)
Students in Research - SBR	09/07/09	8/10 (80%)
History and Ethical Principles - SBR	09/07/09	3/4 (75%)
Defining Research with Human Subjects - SBR	09/07/09	4/5 (80%)
The Regulations and The Social and Behavioral Sciences - SBR	12/10/09	5/5 (100%)
Assessing Risk in Social and Behavioral Sciences - SBR	12/10/09	5/5 (100%)
Informed Consent - SBR	12/10/09	4/4 (100%)
Privacy and Confidentiality - SBR	12/10/09	3/3 (100%)
Research with Prisoners - SBR	12/10/09	4/4 (100%)
Research with Children - SBR	12/10/09	2/4 (50%)
Research in Public Elementary and Secondary Schools - SBR	12/10/09	4/4 (100%)
International Research - SBR	12/10/09	3/3 (100%)
Internet Research - SBR	12/10/09	5/5 (100%)
Johns Hopkins University	09/07/09	no quiz

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.
Professor, University of Miami
Director Office of Research Education
CITI Course Coordinator

Return

CITI Collaborative Institutional Training Initiative

Social & Behavioral Research - Basic/Refresher Curriculum Completion Report Printed on 10/31/2009

Learner: Michael Bloodgood (username: mbloodgood)

Institution: Johns Hopkins University

Contact Information Phone: 410-516-7889

Email: bloodgood@jhu.edu

Social & Behavioral Research - Basic/Refresher:

Stage 1. Basic Course Passed on 10/31/09 (Ref # 3650170)

Required Modules	Date Completed	Score
Belmont Report and CITI Course Introduction	10/26/09	3/3 (100%)
Students in Research - SBR	10/26/09	8/10 (80%)
History and Ethical Principles - SBR	10/27/09	4/4 (100%)
Defining Research with Human Subjects - SBR	10/27/09	5/5 (100%)
The Regulations and The Social and Behavioral Sciences - SBR	10/28/09	5/5 (100%)
Assessing Risk in Social and Behavioral Sciences - SBR	10/28/09	5/5 (100%)
Informed Consent - SBR	10/29/09	4/4 (100%)
Privacy and Confidentiality - SBR	10/29/09	3/3 (100%)
Research with Prisoners - SBR	10/29/09	4/4 (100%)
Research with Children - SBR	10/29/09	4/4 (100%)
Research in Public Elementary and Secondary Schools - SBR	10/30/09	4/4 (100%)
International Research - SBR	10/30/09	3/3 (100%)
Internet Research - SBR	10/30/09	5/5 (100%)
Johns Hopkins University	10/31/09	no quiz

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.
Professor, University of Miami
Director Office of Research Education
CITI Course Coordinator

Return

CITI Collaborative Institutional Training Initiative

Social & Behavioral Research - Basic/Refresher Curriculum Completion Report

Printed on 12/22/2009

Learner: Christine Piatko (username: cpiatko)

Institution: Johns Hopkins University

Contact Information Human Language Technology Center of Excellence

Stieff Building/810 Wyman Park Drive

Baltimore, MD 21211-2840 USA

Department: JHU HLT/COE and JHU/APL and WSE

Computer Science

Phone: 4105164635

Email: Christine.Piatko@jhuapl.edu

Social & Behavioral Research - Basic/Refresher:

Stage 1. Basic Course Passed on 12/22/09 (Ref # 3857485)

Required Modules	Date Completed	Score
Belmont Report and CITI Course Introduction	12/22/09	3/3 (100%)
Students in Research - SBR	12/21/09	9/10 (90%)
History and Ethical Principles - SBR	12/21/09	4/4 (100%)
Defining Research with Human Subjects - SBR	12/21/09	5/5 (100%)
The Regulations and The Social and Behavioral Sciences - SBR	12/21/09	5/5 (100%)
Assessing Risk in Social and Behavioral Sciences - SBR	12/22/09	5/5 (100%)
Informed Consent - SBR	12/22/09	4/4 (100%)
Privacy and Confidentiality - SBR	12/22/09	3/3 (100%)
Research with Prisoners - SBR	12/22/09	4/4 (100%)
Research with Children - SBR	12/22/09	4/4 (100%)
Research in Public Elementary and Secondary Schools - SBR	12/22/09	4/4 (100%)
International Research - SBR	12/22/09	3/3 (100%)
Internet Research - SBR	12/22/09	5/5 (100%)
Johns Hopkins University	12/22/09	no quiz

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.

RESEARCH PROTOCOL

STUDY TITLE

Collection of language data for computational linguistics models with Amazon's Mechanical Turk

STUDY SITE(S)

Where will the study be conducted? Check all locations that apply.

- | | |
|--|--|
| <input type="checkbox"/> Applied Physics Laboratory (APL) | <input type="checkbox"/> Peabody Institute |
| <input type="checkbox"/> Carey Business School | <input type="checkbox"/> School of Advanced International Studies (SAIS) |
| <input type="checkbox"/> Krieger School of Arts and Sciences | <input type="checkbox"/> School of Education |
| <input type="checkbox"/> Kennedy-Krieger Institute (KKI) | <input checked="" type="checkbox"/> Whiting School of Engineering |

Other U.S. Site(s). Specify: JHU Human Language Technology Center of Excellence

International Site(s). Specify:

FUNDING STATUS

Indicate the project's funding status below - *Submit the complete grant proposal or application*

- Project is currently funded.
Funding Agency (e.g., NSF): DARPA
Grant/Contract Number: 105397
- Project has been submitted for funding.
Funding Agency (e.g., NSF):
- No Funding Obtained or Applied For.

PURPOSE OF THE STUDY

Describe the goals of the study clearly below.

The primary goals of the study are to (1) collect language data that will be used to train and/or evaluate statistical models of language and translation, (2) investigate Wikipedia-style collaborative editing and how it may be used to improve natural language processing software such as machine translation systems, and (3) evaluate the quality of data gathered by non-expert annotators who do such work on Amazon's Mechanical Turk online labor market.

DESIGN & METHODOLOGY

Describe the design and methodology of the study clearly below.

Our study investigates computational models of human languages. In this field of study, training data is required in order to estimate the parameters of these computational models. Our study's design will focus on the collection of such training data from human participants. The methodology will be to collect language data, particular to each of our models, using online voluntary participation. For example, for language understanding data we will ask participants whether sentences express various modalities such as belief or desire, and potentially who holds the belief or desire and what is believed or desired. As an example, if the sentence is "John wants to go to NY", then the participants would annotate that there is a desire, "John" is the holder of the desire, and "go to NY" is what is desired. For a sentence translation task, a participant would translate a sentence from one language into another. Following collection, the data will be used to train and evaluation computational models that produce predictions about language (for instance, who wants what, or produce an automatic translation of a foreign sentence into English, or which automatically transliterate words from Arabic script into the Latin alphabet, or which guess which words in a sentence are verbs). The research subsequent to data collection will involve using statistical methods combined with linguistic representations to train computational models. This research on how to improve these methods and representations that is performed subsequent to data collection will not involve human participants.

PROCEDURES & ACTIVITIES

Describe the study's procedures and activities that participants will be asked to perform. *Attach all surveys, interview questions and guidelines, and other relevant materials you will use in this study.*

Participants will be asked to provide annotations for language data through online surveys. For example, these surveys will ask whether one automatically produced translation is better than another, or whether a sentences expresses a belief or desire (and potentially who in the sentence has the belief or desire and what in the sentence is believed or desired), or what a translation of a foreign sentence should be. A set of example surveys is attached.

PARTICIPANTS

Describe the participant population(s) and the total number of participants to be enrolled.

Participants will be individuals who use Amazon's Mechanical Turk web site. The Mechanical Turk web site is a forum that Amazon has established to let companies and researchers pay people small sums of money in order to do ``human intelligence tasks" (things that are easy for people but difficult for computers). Our entire participant pool will be drawn from this site. The number of participants is difficult to know in advance, but will likely range from the tens to the low hundreds.

SPECIAL CATEGORIES OF PARTICIPANTS

Check all categories that apply to this study. Additional regulations and review requirements apply to some participant categories.

JHU Students

JHU Employees

Children (< 18 years old)

If checked, give age range:

Cognitively-Impaired Individuals

Individuals with Limited or No Reading Skills

Non-English Speakers

Pregnant Women/Fetuses

Prisoners

PARTICIPANT RECRUITMENT

In the area below, describe how you will identify and recruit participants. *Please attach copies of all recruitment materials (e.g., ads, bulletin board notices, e-mails, phone scripts).*

Participants will be recruited using the standard practice on Amazon's Mechanical Turk which is to display a short description of the task on along with a link to the task. Individuals can choose to participate in our task or not.

COMPENSATION

Describe how participants will be compensated for their participation, or say 'None'.

We will pay participants small sums of money to complete our task, ranging from \$0.01 to \$1. All participants can choose for themselves whether the compensation is fair, and opt not to do it if they deem the compensation to be too low. Amazon's Mechanical Turk has many other researchers and companies offering tasks, so we will offer compensation that is similar to what others offer.

RISKS

Please consider the risks your study may pose to participants, including physical, psychological, social, economic, and other risks or harms. In the area below, explain any risks posed by your study or state that your study poses no risks to participants.

Our study poses no risks to participants.

MINIMIZATION OF RISKS

Explain the procedures for minimizing any risks to participants, including procedures for protecting the anonymity of participants and confidentiality of data.

The anonymity of the participants is protected through the design of Amazon's Mechanical Turk web site. The only information that is provided to link a participant to a response is a cryptic number (like A2B9RL38SJXB2Q). All identifiable information is kept from us as the researchers, and we have no way to link the number to a participant's true identity. Amazon handles all financial transactions in a way that keeps participants' identities private.

BENEFITS

Describe any potential benefits to participants and/or society. Be realistic in your assessment of potential benefits. (*'Compensation' is not a benefit.*)

The data collected through this research has the potential to enable us to train computational models that can perform useful tasks such as automated language understanding and machine translation. Machine translation is a task that would be of obvious benefit to society, because it would allow greater communication and understanding across language and cultural boundaries. Automatic language understanding is a foundational technology that would support improved information retrieval, information extraction, and question answering by computers, which could benefit to society by improving Internet search, for example.

CONSENT

Exemption from Federal regulations does not negate the ethical obligation to inform participants about their involvement in a study, **when appropriate**. Please indicate whether consent will be obtained for this study by checking one of the boxes below.

Yes - *describe the consent process and attach any disclosures, consent forms, and oral consent scripts.*

The participants will be shown the complete protocol and stimulus before participating, and will be explicitly asked to click on a button labeled "Accept Task" to indicate their consent to participate. Participants may quit the task at any time without repercussion, either by closing their web browser or by clicking on a button that says "Return Task".

No - *explain why consent is not required or is not appropriate for the proposed study.*

CONFLICTS OF INTEREST

Please indicate whether the principal investigator or any other key research personnel for this study have any potential conflicts of interest. Potential conflicts of interest arise when, for example, investigators have an interest in or serve as an officer for an outside entity whose financial interests may be affected by the research. See http://jhuresearch.jhu.edu/Policy_onConflict_of_Interest.pdf for information on JHU policies and procedures concerning conflict of interest.

- No
 Yes - *Please describe the nature of the potential conflict.*

INVESTIGATOR TRAINING

University policy requires that **ALL** individuals engaged in human participant research complete training in human research protections. Certification should be renewed every three years. If proof of training has been submitted to the HIRB for other research, it will remain on file until due for renewal. *If proof of training is not already on file with HIRB, please submit a copy for every team member with this application.* HIRB's training policy can be found on the HIRB website.

- Training completed within last three years Attached on file with HIRB

APPLICATION CHECKLIST

To ensure that your application is complete, please review the checklist below. All items may not apply to your study. All signatures must be obtained and applicable additional forms and documentation must be fully completed before the review process begins.

- | | |
|--|--|
| <input checked="" type="checkbox"/> Signature of Principal Investigator | <input checked="" type="checkbox"/> Surveys, Questionnaires, Interview Guides, and Other Relevant Measures Used in the Study |
| <input checked="" type="checkbox"/> Signature of Student Investigator/Co-Investigator | <input checked="" type="checkbox"/> Recruitment Materials (e.g., Flyers, Ads, Invitations) |
| <input checked="" type="checkbox"/> Research Team Member Forms | <input checked="" type="checkbox"/> Consent Forms/Scripts |
| <input checked="" type="checkbox"/> Certificates of Training in Human Participant Research | <input checked="" type="checkbox"/> Exemption Checklist |
| <input type="checkbox"/> Grant Proposal and Application | |

Surveys and Questionnaires

The following pages show a representative sample of the surveys and questionnaires that will be presented to the human subjects in our study. All of the questionnaires gather information about the English language or about translation from a foreign language into English.

Compare machine translation outputs

Instructions:

- You will be asked to judge two **automatic translations** of an **Urdu sentence**
- If you do not know Urdu, you can use the provided reference translation that was translated by a person
- Your task: Judge which of the two English outputs is a better translation.

NOTE: You need JavaScript to participate in this task. It is probably already enabled on your browser. If you're not sure, see [here](#) (Explorer users) or [here](#) (Firefox users).

Source: فرانس کی تجویز کی حمایت		
Reference: Support of France's Recommendation		
Translation 1	Translation 2	
Supported Proposal of France	France Supported the Proposal	
Translation 1 is better <input type="radio"/>	No difference <input type="radio"/>	Translation 2 is better <input type="radio"/>

Source: کمیٹی کا اجلاس جمعہ کے روز پارلیمنٹ ہاؤس میں کمیٹی کے چیئرمین مشاہد حسین سید کی سربراہی میں ہوا جس میں حکومت کو یہ تجویز دی گئی کہ وہ امدادی اشیاء کے علاوہ تمام سیاسی جماعتوں کے نمائندوں کا ایک پارلیمانی وفد فلسطین بھیجے۔		
Reference: The Committee meeting was held in Parliament House on Friday under the direction of Committee Chairman Mushahid Hussain Sayed, wherein it was recommended to the government that, in addition to sending relief supplies, it should also send a parliamentary delegation of the representatives of all political parties to Palestine.		
Translation 1	Translation 2	
Meeting of the committee in the parliament house on Friday, chairman of the committee under the chairmanship of Mushahid Hussain Sayed, in which the sense of relief goods to this proposal that he sent a parliamentary delegation to Palestine in addition to representatives of all political parties.	The meeting of the committee in the parliament house on Friday in the leadership of the Committee Chairman Mushahid Hussain Syed, a parliamentary delegation of representatives of all political parties in the country in which it was suggested that the relief goods sent to Palestine.	
Translation 1 is better <input type="radio"/>	No difference <input type="radio"/>	Translation 2 is better <input type="radio"/>

Transliterate Urdu words into English

Given an Urdu word, write an English transliteration of the word (an English version of the word that sounds the same). For example, given the Urdu word "ملٹی میڈیا", you may write "multimedia" as an English transliteration. Please transliterate these yourself, and **do not** use online machine translation systems.

Are you an Urdu and English speaker?

- Yes
 No

Give English transliterations of these Urdu words:

مانیکل بسی	<input type="text"/>
دلاور	<input type="text"/>
رچرڈ آرمیٹیج	<input type="text"/>
ذوالفقار مرزا	<input type="text"/>
صدیق خان	<input type="text"/>
ورسک	<input type="text"/>
تہمینہ دولتانہ	<input type="text"/>
زبیر احمد	<input type="text"/>
امان اللہ کنرانی	<input type="text"/>
قیوم	<input type="text"/>

If you have any comments, please enter them below:

Submit

Translate these sentences

Your task is to translate the Urdu sentences into English. Please make sure that your English translation:

- Is faithful to the original in both meaning and style
- Is grammatical, fluent, and natural-sounding English
- Does not add or delete information from the original text
- Does not contain any spelling errors

When creating your translation, please follow these guidelines:

- **Do not use any machine translation systems (like translation.babylon.com)**
- **You may** look up a word on [an online dictionary](#) if you do not know its translation

This task is for people who are fluent in both Urdu and English.

Do you speak both Urdu and English? Yes No

اس سلسلے میں سب سے بڑا جلسہ صوبائی دارالحکومت میں ہوگا جس میں پشاور کے علاوہ نوشہرہ اور چارسدہ ضلع سے پارٹی کارکن حصہ لیں گے۔	Translation of the first sentence goes here.
'ہم جنس شادیاں غیر قانونی ہیں' کیلے فورنیا میں سپریم کورٹ نے تقریباً چار ہزار ہم جنسوں میں شادیوں کو یہ کہہ کر غلط قرار دیا ہے کہ شہر کے مئیر نے ہم جنسوں پرستوں کو شادی کا لائسنس جاری کر کے اپنے اختیارات سے تجاوز کیا ہے۔	Translation of the second sentence goes here.
ان کا کہنا تھا کہ بات چیت سے بہتر نتائج حاصل ہوسکتے ہیں۔	Translation of the third sentence goes here.
میرا نائر فلم کا اختتام نیا ہے جس میں مرکزی کردار بیگی شارپ ایک نئی زندگی کے آغاز کے لیے انڈیا روانہ ہوجاتی ہے۔	Translation of the fourth sentence goes here.
ہان کیمنون زلزلہ زدہ شہر کے دورے پر اقوام متحدہ کے سیکریٹری جنرل بان کیمنون زلزلے سے متاثرہ چین کے صوبے سچوان کا دورہ کر رہے ہیں جہاں ہلاکتوں کی تعداد ساٹھ ہزار سے زائد ہوگئی ہے۔	
ممبئی: بھارتی اخبارات ممبئی کی خیر تمام خبروں پر چھائی ہوئی ہے	
پرنٹ کریں ویاکرا نہیں تو 'لیوٹرا'	
تیرہ رکنی سلامتی کونسل جس کے سربراہ صدر ملکت ہیں، اس کے اراکین میں وزیراعظم، چیئرمین سینیٹ، قومی اسمبلی کے سپیکر اور قائد حزب اختلاف، چاروں صوبوں کے وزراء اعلیٰ، چیئرمین جوائنٹ چیفس آف سٹاف کمیٹی، اور تینوں مسلح افواج (بری، بحری اور فضائیہ) کے سربراہ شامل ہیں۔	
اپنے اعتراضات بتاتے ہوئے انہوں نے کہا ملکی قانون کے تحت کوئی بھی سزا یافتہ یا مقرر شخص الیکشن نہیں لڑسکتا ہے جبکہ بینظیر بھٹو قومی احتساب بیورو کے تین ریفرنسوں میں سزا یافتہ ہیں۔	
فوج معاوضہ ادا کرے: عدالت 'فوجیوں کو زمین نہیں دی جا سکتی' زمیندار فوج زرعی زمین کوڑیوں کے دام	

Did you use a machine translation system to produce these translations? Yes No

Please provide any comments that you have about this HIT. We appreciate your input.

Submit

Indicate whether sentences express a want/desire or not

Instructions:

- For each sentence, indicate whether any part of the sentence expresses a want/desire or not.
- If unsure for a sentence, make a best guess and write in the comments box at the bottom what made that sentence difficult.
- **How to handle Negative Wants:** If a part of a sentence expresses a negative want, then you should still answer yes. This is because in English it seems that to not want X is the same as to want not X. If you think we gave you a sentence where this is not true, make a comment about it in the comments box at the bottom.
- **How to handle different tenses:** If a part of a sentence expresses a want in any tense, please answer yes (even if another part of the sentence indicates that the want no longer exists or that the want was only speculative in the first place -- see examples below).

Examples:

These are examples where "Yes" should be selected:

- John wants to buy a house.
- John doesn't want to buy a house.
- John wanted to buy a house before he found out how expensive they were.
- If John wanted to buy a house, he would have done it by now.

These are examples where "No" should be selected:

- The moon is made of cheese.
- Yesterday it rained a lot.

Your Sentences:

We have measured the length of the line.

- Yes
- No

\$sent2

- Yes
- No

\$sent3

- Yes
- No

Recruitment Materials

The following two pages show an example of the recruitment mechanism available on Amazon's Mechanical Turk web site. Users of the web site, who are our hoped for human subjects, may search for interesting tasks. They are shown a title, a short description, and the amount that they will be paid if they complete the task. In the figure, we have shown a set of example tasks on Mechanical Turk relating to language. The first item shows an example of the recruitment material that we will use, and the others are real examples of what other researcher and companies use to recruit participants.

All HITS | HITS Available To You | HITS Assigned To You

(HITS: 5)

language

0.00

60

HITS containing 'language'

1-10 of 11 Results

Sort by: HIT Creation Date (newest first) 60

[Show all details](#)[Hide all details](#)[Show all details](#)[Hide all details](#)1 2 > [Next](#) >> [Last](#)**Transliterate Urdu words into English**[View a HIT in this group](#)

Requester: [Human Language Technology Center of Excellence](#) HIT Expiration Date: Jan 16, 2010 (1 week 2 days) Reward: \$0.10
Time Allotted: 60 minutes HITS Available: 100

Description: Given an Urdu word, write a corresponding English transliteration (an English version of the word that roughly sounds the same)

Keywords: [Urdu](#), [English](#), [Translation](#), [Transliteration](#), [Language](#), [Research](#)

Qualifications Required: Your Value

HIT approval rate (%) is not less than 95 95 You meet this qualification requirement

[Contact the Requester of this HIT](#)**Listening puzzle**[View a HIT in this group](#)

Requester: [Michael C. Frank](#) HIT Expiration Date: Jan 7, 2010 (23 hours 52 minutes) Reward: \$0.50
Time Allotted: 60 minutes HITS Available: 5

Description: Listen to a strange language for a few minutes -- after a while you may be able to recognize the words. PLEASE DO THIS HIT ONLY ONCE.

Keywords: [language](#), [sound](#), [listening](#), [game](#)

Qualifications Required: Your Value

HIT approval rate (%) is not less than 95 95 You meet this qualification requirement

[Contact the Requester of this HIT](#)**Write a 50 Word Advertisement for a Language Exchange**[View a HIT in this group](#)

Requester: [Andrew H Playford](#) HIT Expiration Date: Jan 8, 2010 (1 day 22 hours) Reward: \$0.30
Time Allotted: 60 minutes HITS Available: 136

Description: Pretend you are the language student below and write a descriptive Ad looking for a language exchange partner.

Keywords: [write](#), [an](#), [ad](#), [advertisement](#), [writing](#), [creative](#), [writing](#), [language](#), [language](#), [student](#), [learn](#), [language](#)

Qualifications Required: Your Value

HIT approval rate (%) is greater than 90 95 You meet this qualification requirement

[Contact the Requester of this HIT](#)**Write a short answer to a question about language**[View a HIT in this group](#)

Requester: [conjecture corporation](#) HIT Expiration Date: Jan 19, 2010 (1 week 5 days) Reward: \$0.03
Time Allotted: 60 minutes HITS Available: 3

Description: Review all of our current questions related to language, and answer as many as you like

Keywords: [opinion](#), [advice](#), [question](#), [answer](#), [short](#), [answer](#), [writing](#)

Qualifications Required: None

[Contact the Requester of this HIT](#)**Language guessing game**[View a HIT in this group](#)

Requester: [tedlab](#) HIT Expiration Date: Jan 12, 2010 (5 days 20 hours) Reward: \$0.05
Time Allotted: 15 hours HITS Available: 26

Description: **A general description of HITS in this group.** en cut off, and guess what comes next. NOT COMPATIBLE WITH INTERNET EXPLORERKeywords: [survey](#), [answers](#), [language](#), [quick](#), [multiple](#), [choice](#), [understanding](#)

Qualifications Required: Your Value

Location is US US You meet this qualification requirement

HIT approval rate (%) is not less than 90 95 You meet this qualification requirement

[Contact the Requester of this HIT](#)**Опрос о предложениях на русском языке [Survey about Russian sentences]**[View a HIT in this group](#)

Requester: [Alya Asarina](#) HIT Expiration Date: Jan 11, 2010 (4 days 14 hours) Reward: \$0.08
Time Allotted: 60 minutes HITS Available: 5

Description: Вы прослушаете 16 предложений, и ответите про каждое, можно ли так сказать.

Keywords: [Russian](#), [language](#), [linguistics](#)

Qualifications Required: Your Value

HIT approval rate (%) is not less than 95 95 You meet this qualification requirement

[Contact the Requester of this HIT](#)**German Language Translation**[Take Qualification test \(Why?\)](#)[View a HIT in this group](#)

Requester: [Music Service Bureau](#) HIT Expiration Date: Jan 10, 2010 (4 days 10 hours) Reward: \$1.25
Time Allotted: 60 minutes HITS Available: 203

Description: This HIT is for you if you know German well. Enter the local movie title and local synopsis in the fields provided. You will enjoy working on these HITs if you love movies. Good luck!

Keywords: [German](#), [translate](#), [movies](#), [translation](#), [Music](#), [Service](#), [Bureau](#), [films](#), [synopsis](#), [synopses](#), [local](#), [description](#), [movie](#)

Qualifications Required:

Your Value

German Language Qualification(#66127) is not less than 60

none [Take Qualification test](#)

[Contact the Requester of this HIT](#)

[Korean Language Translation](#)

[Take Qualification test \(Why?\)](#) | [View a HIT in this group](#)

Requester: [Music Service Bureau](#)

HIT Expiration Date: Jan 6, 2010 (12 hours 30 minutes) Reward: \$1.00

Time Allotted: 60 minutes

HITs Available: 253

Description: This HIT is for you if you know Korean well. Enter the local movie title and local synopsis in the fields provided. You will enjoy working on these HITs if you love movies. Good luck!

Keywords: [Korean](#), [translate](#), [movies](#), [translation](#), [Music](#), [Service](#), [Bureau](#), [films](#), [synopsis](#), [synopses](#), [local](#), [description](#), [movie](#)

Qualifications Required:

Your Value

Korean Language Qualification(#31042) is not less than 60

none [Take Qualification test](#)

[Contact the Requester of this HIT](#)

[German Language Translation](#)

[Take Qualification test \(Why?\)](#) | [View a HIT in this group](#)

Requester: [Music Service Bureau](#)

HIT Expiration Date: Jan 10, 2010 (4 days 10 hours) Reward: \$1.00

Time Allotted: 60 minutes

HITs Available: 575

Description: This HIT is for you if you know German well. Enter the local movie title and local synopsis in the fields provided. You will enjoy working on these HITs if you love movies. Good luck!

Keywords: [German](#), [translate](#), [movies](#), [translation](#), [Music](#), [Service](#), [Bureau](#), [films](#), [synopsis](#), [synopses](#), [local](#), [description](#), [movie](#)

Qualifications Required:

Your Value

German Language Qualification(#66127) is not less than 60

none [Take Qualification test](#)

[Contact the Requester of this HIT](#)

[Provide a Seven Syllable English Language Statement](#)

[View a HIT in this group](#)

Requester: [cyrus durain](#)

HIT Expiration Date: Apr 6, 2010 (12 weeks 5 days) Reward: \$0.05

Time Allotted: 24 hours

HITs Available: 1

Description: This will be used as the middle statement in a haiku. It should be a complete thought or statement, but can be (should be?) dramatically out of context.

Keywords: [Seven](#), [Syllable](#), [English](#), [Language](#), [Statement](#), [Question](#)

Qualifications Required:

Your Value

HIT approval rate (%) is not less than 95

95

You meet this qualification requirement

[Contact the Requester of this HIT](#)

Johns Hopkins University
Homewood Institutional Review Board (HIRB)

Recruitment Texts

Title: Collection of language data for computational linguistics models with Amazon's Mechanical Turk

Principal Investigator: **Chris Callison-Burch – Computer Science Department**

Date: **Jan 13, 2010**

Recruitment texts:

[The following are examples of the short recruitment texts that will be posted on the Amazon Mechanical Turk page for advertising tasks to workers – see <https://www.mturk.com/mturk/findbits?match=false>. These texts will be filled into the web page, along with the description of tasks from other Requesters. Workers/subjects are able to sort and search based on keyword or reward, and select only those tasks that they are interested in doing.]

Title: Urdu-English Word Alignment

Description: Look at an English translation of an Urdu sentence and show the word-for-word correspondences

Keywords: Urdu, English, translation, word, alignment, research, JHU

Reward: \$1.00

Title: Urdu-English Translation

Description: Translate 10 Urdu sentences into English

Keywords: Urdu, English, translation, writing, research, JHU

Reward: \$1.00

Title: Urdu to English Translation

Description: Translate 10 Urdu phrases into English

Keywords: Urdu, English, translation, research, JHU

Reward: \$0.10

Title: Change machine translation so that its meaning is the same as a human translation

Description: Read a human translation and a machine translated version, and make as few changes to the MT to make it correct

Keywords: machine translation, writing, editing, language, research, JHU

Reward: \$0.40

Title: Judge edited machine translation output

Description: Judge whether edited machine translation output has the same meaning as a human translation, and whether it is passable English

Keywords: machine translation, reading, grading, text, English
Reward: \$0.10

Title: Evaluate machine translation quality

Description: Compare different machine translation systems by ranking their translated output

Keywords: machine translation, translation, quality, evaluation, research, language, English, German, JHU

Reward: \$0.02

Title: Create reading comprehension test questions

Description: Read a short newspaper article and create three reading comprehension questions about it

Keywords: machine translation, translation, quality, evaluation, research, language, reading comprehension, writing, text, English

Reward: \$0.20

Title: Answer questions about machine translation

Description: Read a newspaper article that was automatically translated from another language, and answer questions about it

Keywords: machine translation, reading comprehension, research, English

Reward: \$0.15

Title: Transliterate Urdu words into English

Description: Given an Urdu word, write a corresponding English transliteration (an English version of the word that roughly sounds the same)

Keywords: Urdu, English, Translation, Transliteration, Language, Research

Reward: \$0.10

Title: Determine whether text expresses a want/desire or not

Description: Given a sentence, indicate whether the sentence expresses a want or not

Keywords: English, modality, research, want, classification, JHU, language, machine learning

Reward: \$0.10

Consent

The following page shows the consent mechanism that is available on Amazon's Mechanical Turk web site. A subject is allowed to preview the entire survey/human intelligence task (HIT) they will complete to gauge whether the compensation is fair and whether the task is interesting to them. They then give their consent by clicking on the "Accept HIT" button. They can also opt not to do the task by clicking on the "Skip HIT" button. They can withdraw their consent without penalty by clicking on the "Return HIT" button (which is shown once the task begins) or by closing their web browser window.

All HITS | HITS Available To You | HITS Assigned To You

(HITS) [dropdown]

0.00

[GO]

Timer: 00:00:00 of 60 minutes Want to work on this HIT? Want to see other HITS? Total Earned: \$5.57 Total HITS Submitted: 285

Urdu-English Word Alignment

Requester: Human Language Technology Center of Excellence Reward: \$0.05 per HIT HITS Available: 73 Duration: 60 minutes
Qualifications Required: HIT approval rate (%) is not less than 95

Word alignment

The initial alignment that you see was created by a computer, and it contains errors. Please indicate which words are in correspondence. If there is not a direct correspondence (as in the boxes). If any of the computer's initial suggestions are wrong please uncheck them.

This task should only be completed by people who speak both Urdu and English. Do you speak both?

Yes No

However	,	the	New	York	police	have	confirmed	that	explosion	was	not	a	result	of	terrorism	but	was	caused	by	the	bursting	of	a	steam	in	the	building	
تاہم																												
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
نیویارک																												
پولیس																												
نے																												
کھدق																												

Want to work on this HIT? Want to see other HITS?

Report this HIT: [violates the Amazon Mechanical Turk policies](#) or [broken](#) (Why?)

Johns Hopkins University
Homewood Institutional Review Board (HIRB)

Informed Consent Form

Title: Collection of language data for computational linguistics models with Amazon's Mechanical Turk

Principal Investigator: Chris Callison-Burch – Computer Science Department

Date: Jan 13, 2010

PURPOSE OF RESEARCH STUDY:

The purpose of this research study is to collect information about languages, with the aim of improving translation software, speech recognition, web search, and other technologies.

PROCEDURES:

[Procedures will vary depending on task. See attachments to HIRB application for other procedures.]

Your task is to translate the Urdu sentences into English. Please make sure that your English translation:

- Is faithful to the original in both meaning and style
- Is grammatical, fluent, and natural-sounding English
- Does not add or delete information from the original text
- Does not contain any spelling errors
- When creating your translation, please follow these guidelines:

Do not use any machine translation systems (like translation.babylon.com, or translate.google.com). You may look up a word on an online dictionary if you do not know its translation

This task is only for people who are fluent in both Urdu and English.

RISKS/DISCOMFORTS:

There are no risks for participating in this study beyond those associated with normal computer use.

BENEFITS:

Although it will not directly benefit you, this study may benefit society by improving how computers process human languages. This could lead to better translation software, improved web searching, or new user interfaces for computers and mobile devices.

Title:

PI:

Date:

VOLUNTARY PARTICIPATION AND RIGHT TO WITHDRAW:

Participation in this study is voluntary, and you can stop at any time without any penalty. To stop, click on the “Return HIT” button, or close your browser window.

CIRCUMSTANCES THAT COULD LEAD US TO END YOUR PARTICIPATION:

We may decide to end your participation if we determine that you do not have adequate knowledge of a foreign language, or that you are not following the instructions closely enough, or that your answers significantly deviate from true, gold standard annotations.

CONFIDENTIALITY:

The only identifying information kept based on your participation in this study will be your Amazon Mechanical Turk WorkerID serial number. This information may be disclosed to other researchers.

COMPENSATION:

If you satisfactorily complete the study, you will receive \$1 to compensate you for your participation. \$0.50 of this amount is a bonus for completing all of the sessions. If you end your participation before completing the study, you will be paid for your participation up to that time, at a rate of \$0.10 per sentence. Payments are made via Amazon’s payment system.

IF YOU HAVE QUESTIONS OR CONCERNS:

You can ask questions about this research study now or at any time during the study by e-mailing Chris Callison-Burch at ccb@cs.jhu.

If you have questions about your rights as a research participant or feel that you have not been treated fairly, please call the Homewood Institutional Review Board at Johns Hopkins University at (410) 516-6580.

CLICKING ACCEPT

By clicking on the “Accept HIT” button, you indicated that you agree to participate in the study and that you understand the information in this consent form. You have not waived any legal rights you otherwise would have as a participant in a research study.

Johns Hopkins University Homewood Institutional Review Board (HIRB)

Checklist for Exempt Status

Below are six categories of research that may be exempt from Federal regulations for the protection of human research participants. Please review each of the categories below (including any additional tests where indicated) and check any that apply to your proposed study. If the activities in your study do **not** fit entirely into the exemption categories, please complete an Application for Expedited/Full Board Review.

Please note that the exemption categories **DO NOT** apply when research activities involve any of the following:

- Clinical studies of medical devices, procedures, treatments, or drugs.
- Pregnant women, fetuses, neonates, or human in vitro fertilization.
- Prisoners.
- Surveys or interviews that include minors (i.e., children) as participants.
- Collection or study of existing data if the information is recorded in such a way that participants can be identified, either directly or through identifiers linked to the participants and disclosure of the participants' responses outside the research could reasonably place them at risk of criminal or civil liability or be damaging to their financial standing, employability, or reputation.
- Research techniques or activities that expose participants to discomfort or risk beyond that encountered in daily life.
- Deception of research participants.

Studies involving any of these elements **DO NOT** qualify for exemption, and investigators should submit an Application for Expedited/Full Board Review.

=====

QUALIFICATION FOR EXEMPTION

Do **ALL** of the human participant research activities in your study fall into one or more of the following exemption categories and pass **ALL** additional required tests?

<input checked="" type="checkbox"/> YES	Select the category number(s) in which the research falls and specify: <u>2</u> Your study may qualify for exemption. Complete the Application for Exemption following the checklist, and submit both components to HIRB.
<input type="checkbox"/> NO	Your study does not qualify for exemption. Please complete and submit the Application for Expedited/Full Board Review.

EXEMPTION CATEGORIES

CATEGORY 1

Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (a) research on regular and special education instructional strategies and (b) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

Check if any part of your project falls within Category 1.

CATEGORY 2

Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior.

To be eligible for exemption under this category, the research must pass **TWO** additional tests.

Test I: Children

Children are persons who have not attained the legal age for consent to treatments or procedures involved in the research, under the applicable law of the jurisdiction in which the research will be conducted.

Does any part of your project involve research in which children are studied through (a) survey procedures, (b) interview procedures, or (c) observations of public behavior in which the investigators participate in the activities being observed?

No ➔ Your project passes Test I.

Yes ➔ Your project fails Test I, and does **NOT** qualify for exemption from review.

Test II: Identifiability & Risks

a. Will the information obtained in the research be recorded in such a manner that the human participants can be identified, directly or through identifiers linked to the participants?

Yes No

b. Could any disclosure of the human participants' responses outside the research reasonably place the participants at risk of criminal or civil liability or be damaging to the participants' financial standing, employability, or reputation? Yes No

**** IMPORTANT ****

For (a), what matters is **NOT** whether participants will be identified when the results of the research are reported. What matters is whether the data are recorded in such a way that the information obtained from a participant can be linked to the identity of that participant.

Example: Data for each participant are recorded on a page that identifies the participant only by a code number. The researcher keeps a key that associates a name with each code number. Even if no names or code numbers are listed in publications or other reports, the data are recorded in such a way that participants can be identified, and the answer to (b) is YES.

If you answered:

No to (a) **OR** No to (b) ➔ Your project passes Test II.

Yes to (a) **AND** Yes to (b) ➔ Your project fails Test II, and is **NOT** exempt.

Check if any part of your project falls within Category 2 and passes **BOTH** additional tests.