

Query reformulation by phrases in interactive CLIR

Fernando López-Ostenero

Julio Gonzalo

Anselmo Peñas

Felisa Verdejo

UNED (Madrid)

UNED hypothesis

- Assisted query term translation is a hard task for the user.
- Substitute assisted query translation by monolingual phrase selection + automatic phrase translation.

UNED experiment

Reference system (WORDS)

- Assisted query translation (by inverse dictionary).
- Assisted query refinement (id.)

Contrastive system (PHRASES)

- Assisted query formulation by phrases.
- Assisted query refinement by phrases.

• Document translation: CL summaries based in noun phrases (UNED@iCLEF 2001)

• CLEF English documents, 8 monolingual Spanish users.

WORDS: query translation

Mozilla: Consulta número 1 con el sistema 1

<input type="checkbox"/> enfermedad	<input type="checkbox"/> disease <input type="checkbox"/> ailment <input type="checkbox"/> dolencia achaque achuchón alifafe arrechucho enfermedad padecimiento <input type="checkbox"/> ill	<input type="checkbox"/> desarrollo	<input type="checkbox"/> buildup <input type="checkbox"/> growing <input type="checkbox"/> growth <input type="checkbox"/> maturation <input type="checkbox"/> ontogenesis <input type="checkbox"/> ontogeny <input type="checkbox"/> evolution
	<input type="checkbox"/> cause <input type="checkbox"/> bring_about <input type="checkbox"/> give <input type="checkbox"/> get <input type="checkbox"/> have		<input type="checkbox"/> human <input type="checkbox"/> individual

Por favor, seleccione los términos de traducción que considere adecuados y pulse el botón *Realizar Consulta*

WORDS: document ranking

Mozilla: Consulta número 1 con el sistema 1

<input type="checkbox"/> gen	<input type="checkbox"/> i gene		<input type="checkbox"/> i build-up
	<input type="checkbox"/> i disease		<input type="checkbox"/> i buildup
<input type="checkbox"/> enfermedad	<input type="checkbox"/> i ailment	<input type="checkbox"/> desarrollo	<input type="checkbox"/> i growing
	<input type="checkbox"/> i complaint		<input type="checkbox"/> i growth
	<input type="checkbox"/> i ill		<input type="checkbox"/> i maturation
	<input type="checkbox"/> i cause		<input type="checkbox"/> i ontogenesis
	<input type="checkbox"/> i bring_about		<input type="checkbox"/> i ontogeny
	<input type="checkbox"/> i give		<input type="checkbox"/> i evolution
	<input type="checkbox"/> i get		<input type="checkbox"/> i human

Leyenda de Símbolos

- No Juzgado
- Relevante
- No Relevante
- ? Dudoso

- [CONDADO NEWSWATCH DE VENTURA](#)
- [FOCO ANARANJADO DEL CONDADO: ANARANJADO; ALCALDE ENFERMO SAYS ÉL BUSCARÁ LA REELECCIÓN](#)
- [INNOVACIÓN / MICHAEL SCHRAGE: Y EN SU IZQUIERDA, USTED PUEDE APENAS HACER HACIA FUERA GENE PROFUNDO DEL COLUMNISTA](#)
- [LOS CIENTÍFICOS ENCUENTRAN EL GENE DEFECTUOSO LIGADO PARA PELAR EL CÁNCER](#)
- [2DO EL GENE DEL CÁNCER DE LOS DOS PUNTOS ENCONTRÓ; LAS PRUEBAS ESPERARON PRONTO](#)
- [LOS INVESTIGADORES IDENTIFICAN EL GENE DOMINANTE DEL CÁNCER](#)
- [EL DESCUBRIMIENTO DEL GENE DEL CÁNCER DE PECHO LLAMÓ AVANCE IMPORTANTE; MEDICINA: INVESTIGADORES LOCALICE EL DEFECTO QUE CAUSA LA MITAD DE CASOS HEREDADOS. LA DETECCIÓN TEMPRANA PUEDE](#)

WORDS: assisted refinement

Mozilla: Consulta número 1 con el sistema 1

<input type="checkbox"/> gen	<input type="checkbox"/> i gene ←		<input type="checkbox"/> i build-up
	<input type="checkbox"/> i disease		<input type="checkbox"/> i buildup
<input type="checkbox"/> enfermedad	<input type="checkbox"/> i ailment	<input type="checkbox"/> desarrollo	<input type="checkbox"/> i growing
	<input type="checkbox"/> i complaint		<input type="checkbox"/> i growth
	<input type="checkbox"/> i ill		<input type="checkbox"/> i maturation
	<input type="checkbox"/> i cause		<input type="checkbox"/> i ontogenesis
	<input type="checkbox"/> i bring_about		<input type="checkbox"/> i ontogeny
	<input type="checkbox"/> i give		<input type="checkbox"/> i evolution
	<input type="checkbox"/> i get		<input type="checkbox"/> i human

Añadir término a la consulta

Evalúe el **gen **i** de la obesidad**
Documento

Relevante
 No Relevante
 ? Dudoso

notable descubrimiento
gordo **gene** **i**
aplastante mayoría
malo alimento
genetico disfuncion_fisica

PHRASES: query formulation

Mozilla: Consulta número 4 con el sistema 2

Seleccione las expresiones que desea incluir en la consulta

La consulta contiene 5 expresiones

<input type="checkbox"/> atención al paciente	<input type="checkbox"/> consecuencia de la huelga
<input type="checkbox"/> caso de ataque	<input type="checkbox"/> motivo de la huelga
<input checked="" type="checkbox"/> huelga de hambre	

Si lo desea puede añadir nuevos términos individuales a la consulta

Leyenda de Símbolos

- No Juzgado
- Relevante
- No Relevante
- ? Dudoso

1	<input checked="" type="radio"/> <u>NOTICIAS DE LA COMUNIDAD: CENTRAL: PARQUE DEL ECO; PLAN EXCESIVO EFECTUADO DEL ALZA DEL AUTOBÚS DE LA HUELGA DE HAMBRE</u>
2	<input checked="" type="radio"/> <u>MUJER PARA TERMINAR HUELGA DE HAMBRE DE GUATEMALA; DERECHOS HUMANOS: EL ABOGADO DE ESTADOS UNIDOS BUSCARÁ CARGAS DEL CRIMINAL CONTRA OFICIALES DEL EJÉRCITO. ELLA SE HABÍA MUERTO DE HAMBRE POR 32 DÍAS ADENTRO ESFUERZO DE APRENDER DEL SINO DE SU MARIDO.</u>
3	<input checked="" type="radio"/> <u>EL ACUERDO TERMINA HUELGA DE HAMBRE DE SANTA DE LOS ESTUDIANTES UC DE BARBARA</u>
4	<input checked="" type="radio"/> <u>UCLA ESCOGE LA CABEZA DEL INTERINO PARA LOS ESTUDIOS DE CHICANO; EDUCACIÓN: ALGUNOS ACTIVISTAS CRITIQUE LA CITA DEL WHO DEL ADMINISTRADOR AYUDADA PARA NEGOCIAR EL EXTREMO PARA TENER HAMBRE HUELGA.</u>
5	<input checked="" type="radio"/> <u>ABOGADOS EGIPCIOS EN LOS COLEGAS EXCESIVOS DE LA HUELGA DE HAMBRE QUE ENCARCELAN</u>

PHRASES: query refinement

Mozilla: Consulta número 4 con el sistema 2

Seleccione las expresiones que desea incluir en la consulta

La consulta contiene 5 expresiones

<input type="checkbox"/> atención al paciente	<input type="checkbox"/> consecuencia de la huelga
<input type="checkbox"/> caso de ataque	<input type="checkbox"/> motivo de la huelga
<input checked="" type="checkbox"/> huelga de hambre	

Si lo desea puede añadir nuevos términos individuales a la consulta

Evalúe el Documento

Relevante

No Relevante

? Dudoso

huelga de hambre en guatemala ←

derechos humanos

criminal acusaciones contra oficiales del ejército

jornada de trabajo

abogados estadounidenses

PHRASES: query refinement

Mozilla: Consulta número 4 con el sistema 2

Seleccione las expresiones que desea incluir en la consulta

La consulta contiene 6 expresiones

<input type="checkbox"/> atención al paciente	<input type="checkbox"/> consecuencia de la huelga
<input type="checkbox"/> caso de ataque	<input type="checkbox"/> motivo de la huelga
<input type="checkbox"/> huelga de hambre	<input type="checkbox"/> huelga de hambre en guatemala

Si lo desea puede añadir nuevos términos individuales a la consulta

Leyenda de Símbolos

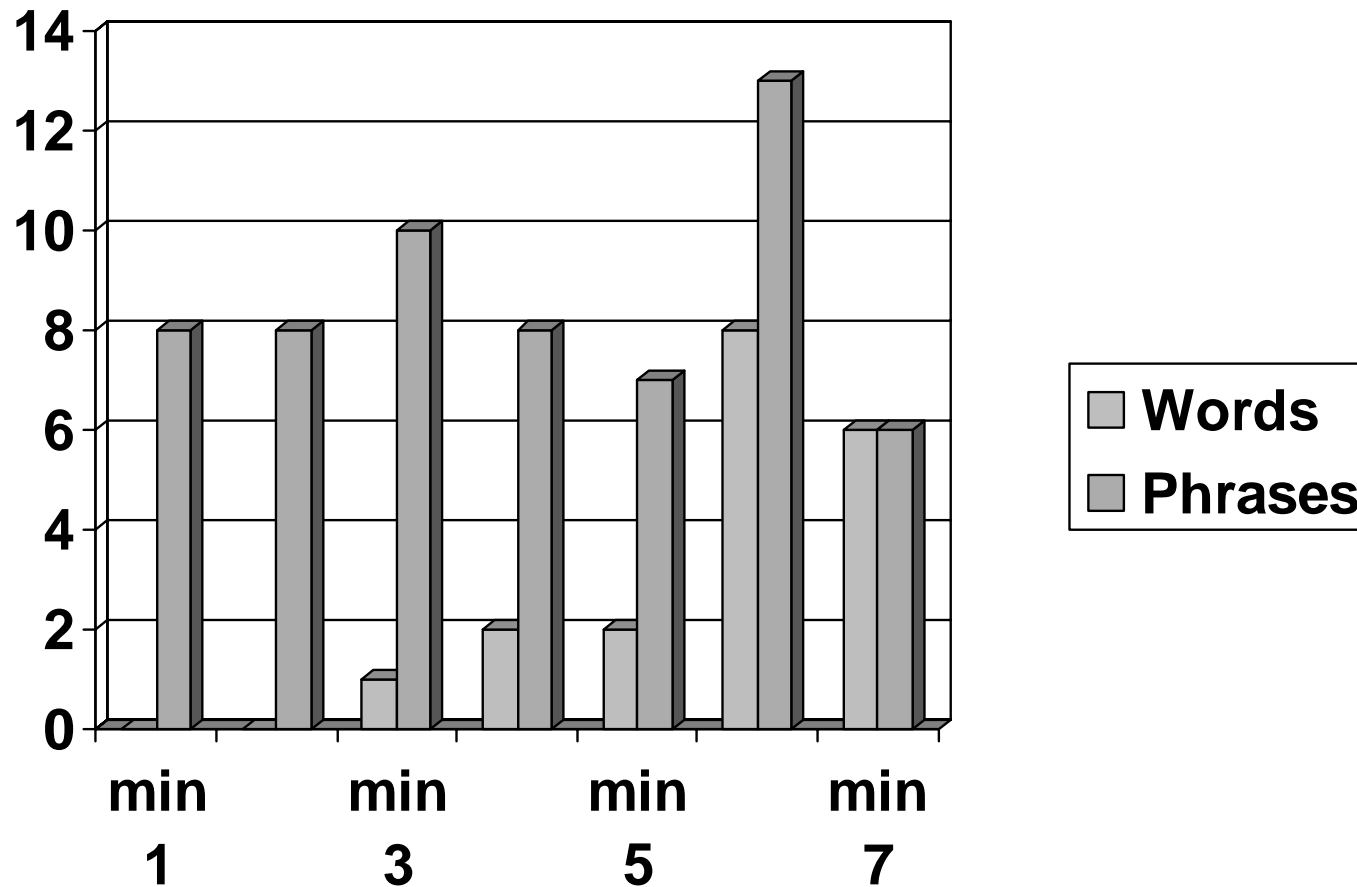
- No Juzgado
- Relevante
- No Relevante
- ? Dudoso

1	<input checked="" type="radio"/>	<u>MUJER PARA TERMINAR HUELGA DE HAMBRE DE GUATEMALA; DERECHOS HUMANOS: EL ABOGADO DE ESTADOS UNIDOS BUSCARÁ CARGAS DEL CRIMINAL CONTRA OFICIALES DEL EJÉRCITO. ELLA SE HABÍA MUERTO DE HAMBRE POR 32 DÍAS ADENTRO ESFUERZO DE APRENDER DEL SINO DE SU MARIDO.</u>
2	<input checked="" type="radio"/>	<u>ABOGADO DE ESTADOS UNIDOS EN LA HUELGA DE HAMBRE PARA EL ESPOSO PERDIDO DEL GUERRILLA; GUATEMALA: JENNIFER HARBURY DICE QUE ELLA SE MORIRÁ DE HAMBRE A LA MUERTE A MENOS QUE EL EJÉRCITO POSEA HASTA LO QUE ELLA VE COMO UN EXPEDIENTE 2 1/2-YEAR DEL ENGAN@O. ELLA INSISTE QUE ÉL ES UN PRISIONERO DE GUERRA Y SE HA TORTURADO SIN PIEDAD.</u>
3	<input checked="" type="radio"/>	<u>NOTICIAS DE LA COMUNIDAD: CENTRAL: PARQUE DEL ECO; PLAN EXCESIVO EFECTUADO DEL ALZA DEL AUTOBÚS DE LA HUELGA DE HAMBRE</u>
4	<input checked="" type="radio"/>	<u>ALBANIA FACILITA EL SOPORTE EN CARACTERÍSTICA CONFISCADA</u>
5	<input checked="" type="radio"/>	<u>EL ACUERDO TERMINA HUELGA DE HAMBRE DE SANTA DE LOS</u>

Official results

System	$F_{\alpha=0.8}$
WORDS	.23
PHRASES	.37 (+65%)

Document selections across time



Precision of initial queries

System	P @ 20
Assisted query translation	.19
Formulation by phrases	.29

Number of query refinements

System	# iterations
Assisted query translation	44
Formulation by phrases	93

Questionnaires

- All users stated that the phrases system was
 - Easier to learn.
 - Easier to use.
 - Better overall.
- Better not to see English terms at any moment.
- Better to select phrases rather than terms.
- Too many translations per term in the dict.

Conclusions

- All sources of evidence (quantitative data, questionnaires, observational studies) support our hypothesis.
- It is worth pursuing research in interactive query formulation at a monolingual level, keeping the translation component hidden to the user.