



The Leeds Arabic Discourse Treebank: Guidelines for Annotating Discourse Connectives and Relations

Amal Al-Saif, Katja Markert

School of Computing
University of Leeds



- Discourse coherence
- Theoretical and empirical studies & applications
- What about Arabic?
- Collection methodology of discourse connectives and relations
- Discourse annotation tool & annotation guidelines
- Discourse corpus distribution
- Usages of Leeds Arabic Discourse Treebank

- Discourse relations play a critical role in making a coherent text/dialogue
 - beside other cohesion devices such as reference, substitution and ellipsis (Halliday & Hasan 1976)
- Discourse relations are intentional or informational/semantic relations
- **Informational/semantic relations** such as CONTRAST, CAUSE, CONDITIONAL, TEMPORAL, etc. between abstract entities of facts, beliefs, eventualities, etc. called **Abstract Objects** (AOs) (Asher, 1993).

- **Discourse Structure Theories and relations taxonomies**
 - Intentional Discourse Model – G&S (Grosz and Sidner 1986).
 - Linguistic Discourse Model -LDM (Polanyi 1998)
 - Rhetorical Structure Theory- RST (Mann and Thompson 1988)
 - Segmented Discourse Representation Theory – SDRT (Asher and Lascarides 2003)
 - Discourse GraphBank by Wolf and Gibson (Wolf and Gibson 2005)
 - Discourse Lexicalized TAG (Webber, Joshi, Stone, Knott)
- **Empirical Studies (corpus-based study)**
 - RST Discourse Treebank (Carlson and Marcu. 2001)
 - PDTB - Penn Discourse Treebank (2008)



Discourse Relations:

- Contrast relation is signalled using the explicit discourse connective: **Instead**
- Cause relation is inferred from the context -> implicit discourse connective (**because**)

- Machine Translation

Condition Relations - PDTB	Eng Conn.	Arabic Connectives
- General	if	في حال - اذا - ما دام
- Unreal_Past		لو
- Factual_Past		في حال
- Unreal_Present		في حال - اذا - ما دام
- Factual_Present		في حال - اذا - ما دام

- Text Summarization

Exemplification

تتعرض دول الشرق الأوسط لأزمة في الإستقرار الأمني. فعلى سبيل المثال، تستمر الإحتجاجات في

The nations of the Middle East crisis in the stability of security. For example, protests continue in

Conjunction

Conjunction

اليمن و يتعرض الشعب الليبي للقصف. و امتدت المظاهرات ايضا الى البحرين احد دول الخليج العربي.

Yemen, and the Libyan people being bombed. Demonstrations are also extended to Bahrain one of Arabic Gulf countries.

- Machine Translation

Condition Relations - PDTB		Eng Conn.	Arabic Connectives
-	General	if	في حال - اذا - ما دام
-	Unreal_Past		لو
-	Factual_Past		في حال
-	Unreal_Present		في حال - اذا - ما دام
-	Factual_Present		في حال - اذا - ما دام

- Text Summarization

تتعرض دول الشرق الأوسط لأزمة في الإستقرار الأمني. فعلى سبيل المثال، تستمر الإحتجاجات في

The nations of the Middle East crisis in the stability of security.

اليمن و يتعرض الشعب الليبي للقصف. و امتدت المظاهرات ايضا الى البحرين احد دول الخليج العربي.

- Question- answering and information extraction systems
- Speech recognition
- Text Generation
- Essay scoring
-

Arabic

- ❑ **No** comprehensive list of discourse connectives
- ❑ **No** standard definitions of discourse relations
- ❑ **No** large scale studies of discourse structure
- ❑ **No** discourse annotation tools
- ❑ **No** discourse annotated corpora
- ❑ **NO** machine learning algorithms for recognizing discourse relations



Our Study..



UNIVERSITY OF LEEDS

Arabic



- ❑ Comprehensive list of discourse connectives



- ❑ Standard definitions of discourse relations for Arabic



- ❑ Discourse annotation guidelines



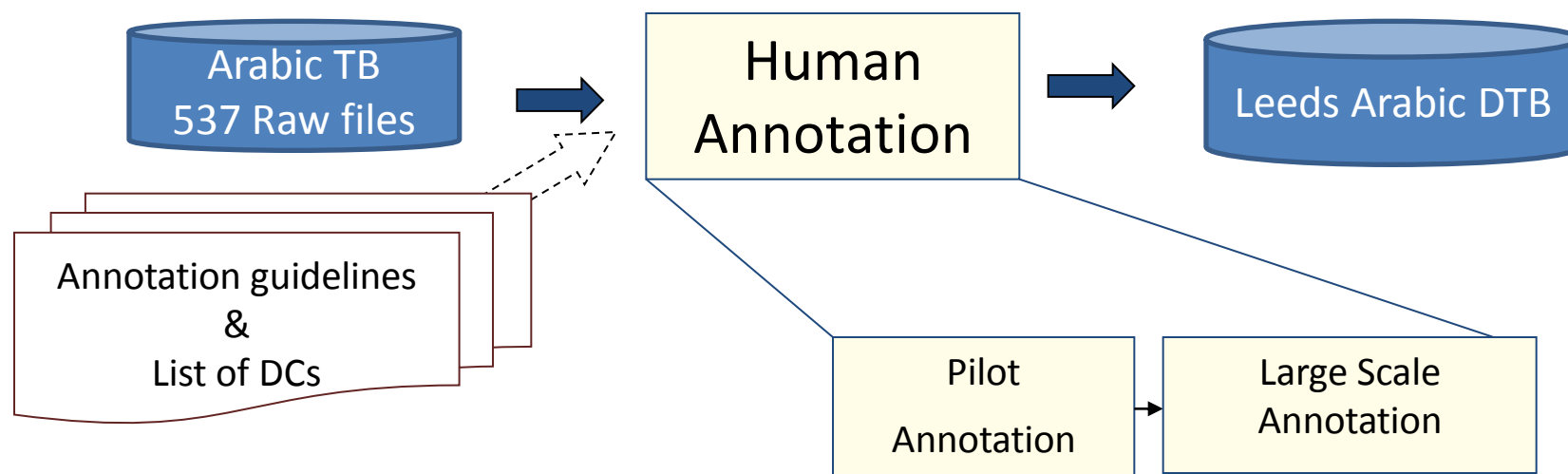
- ❑ Discourse annotation tools for Arabic/English



- ❑ LADTB – discourse annotated Arabic Treebank



- ❑ First machine learning algorithms for recognizing discourse relations



Only Explicitly signalled discourse relations

- 76% of sentences starting with potential connective (537 files of ATB Part1)

Based on similar principles as PDTB scheme (Prasad, Dinesh et al. 2008)

- Theory-neutral
- More flexibility & reliability with a fine-grained hierarchical structure of discourse relations
- Tested practically for English and adapted recently for Chinese, Hindi and Turkish

سنظل نواجه نوعًا من التحدي في جعل أفكارنا واضحة ومتألقة، وهذا يعود إلى أن كل اللغات مصابة بالقصور الذاتي على مستوى نظم التعبير والصياغة وعلى مستوى نظم الفهم والتفسير والتأويل، وقد قال أحد المفكرين: لو شرحت فكرتك للناس عشرين مرة، ووجدت أنهم قد فهموا عنك ما تريد على نحو تام، فأنت محظوظ! وليس هذا وحده هو مصدر غموض الأفكار بل هناك شيء يتعلق ببنية الأفكار ذاتها، وهو أن الحقيقة ذات طبقات عدّة، وكلما غصنا في طبقاتها وجدنا أنفسنا أقل قدرة على الفهم والإدراك، وقليلون جدًا أولئك الذين يتقنون (الحفر المعرفي) وأولئك الذين يشعرون بالحاجة إليه. نحن دائمًا نظن أن أفكارنا واضحة، ولهذا فإننا لا نهتم بتوضيحها كما ينبغي، ولا نكتشف مدى ما فيها من إبهام إلا حين يُطلب منا شرحها، أو حين نحاول إثراءها بالأمثلة.

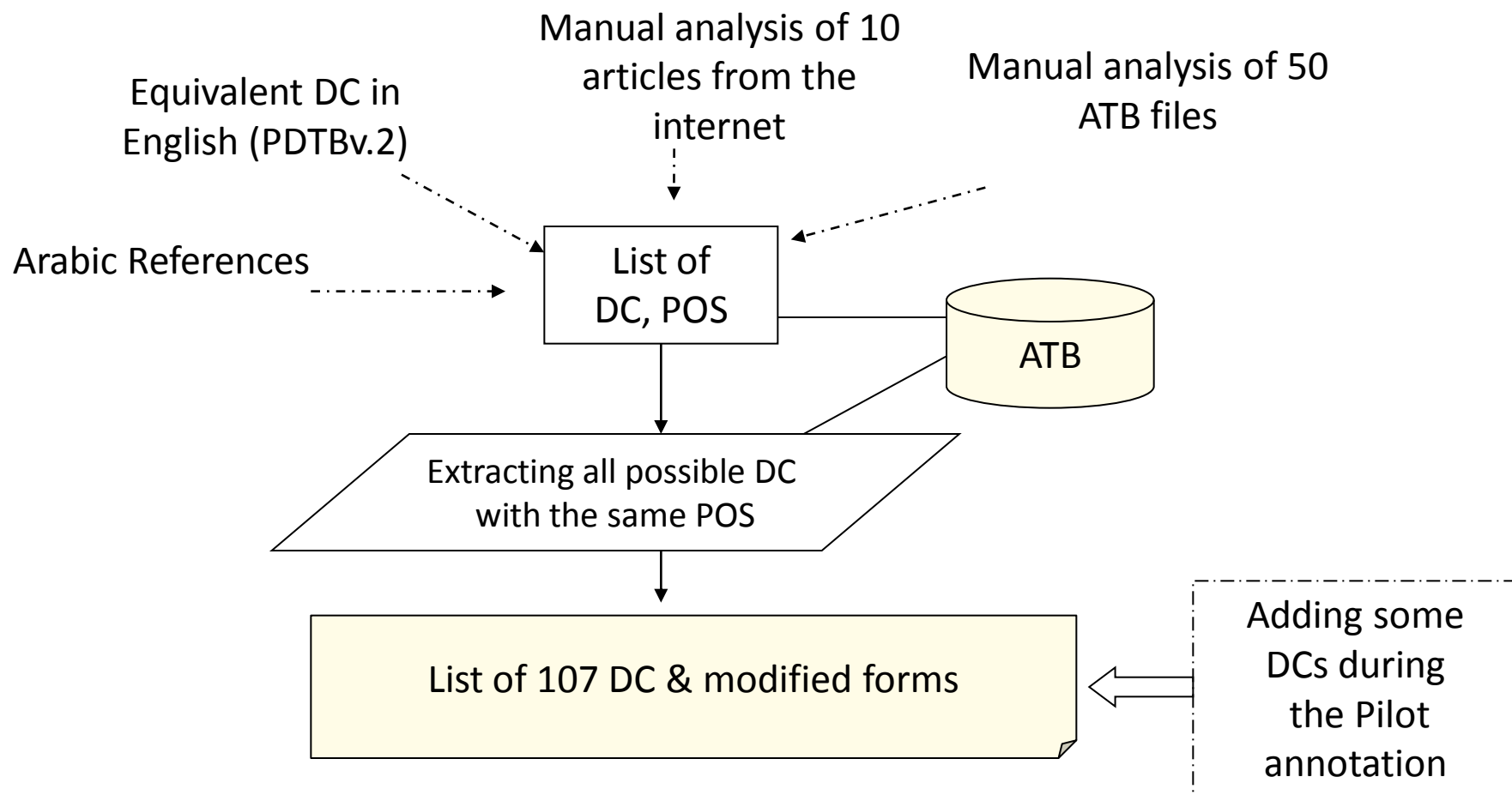
نحن اليوم في حاجة ماسّة لأن نكون أكثر وضوحًا، وذلك بسبب الضعف اللغوي الخطير لدى الأجيال الجديدة وبسبب ما تنشره (الإنترنت) من سطحية التعامل مع الرؤى والمفاهيم على مستوى العالم، ومع هذا وذاك فإن علينا أن نقول: إنه سيظل من الصعب توضيح الأفكار المعقدة والتعبير عنها بالفاظ سهلة. ففي ذلك نوع من الإساءة إليها، ولذا فإن علينا أن نسدّد ونقارب.

أ.د. عبدالكريم بكار، 2011

Collection of Discourse Connectives



UNIVERSITY OF LEEDS



Collection of Discourse Connectives



UNIVERSITY OF LEEDS

107 potential discourse connectives in a word cloud generated by www.wordle.net



The guidelines provide details about:

- **What are Discourse Connectives?**
Connectives Types and grammatical categories
- **Arguments representing AOs**
Types, orders and what can not be arguments
- **Discourse Relations**
Definitions and examples

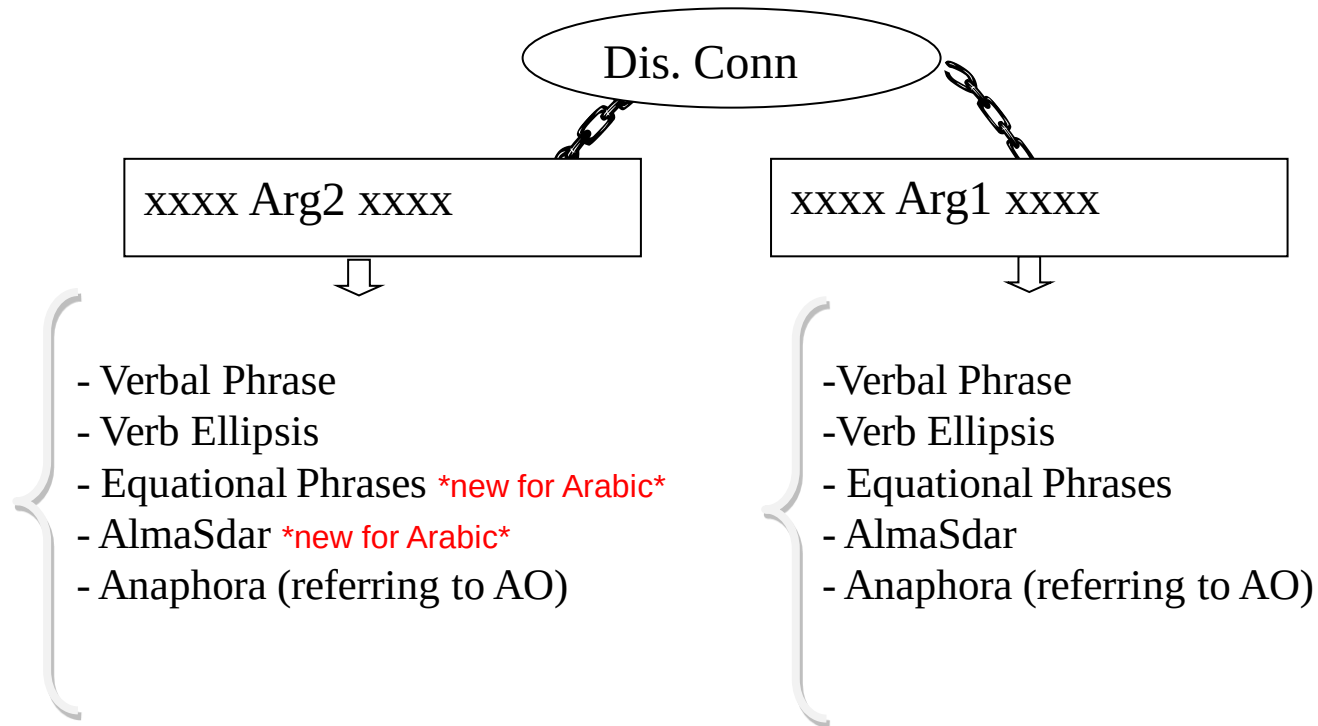
- Simple connectives
 - Ex: *‘لأن’ because, ‘بعدها’ after, ‘و’ and, ‘على النقيض’ in contrast, ‘نتيجة ل’ as result of, ‘من جهة’ besides*
- Paired Connectives
 - Ex: *‘إذا..ف..’ if.. Then, ‘رغم..ف’ although.., ‘ما لبث .. حتى..’*
- Modified connectives
 - Ex: *‘حتى لو’ even if, ‘بعدها’ after, ‘و ايضا’ and also*
- Combined connectives
 - Ex: *‘الا بعد’ except after, ‘ولكن’ and but*

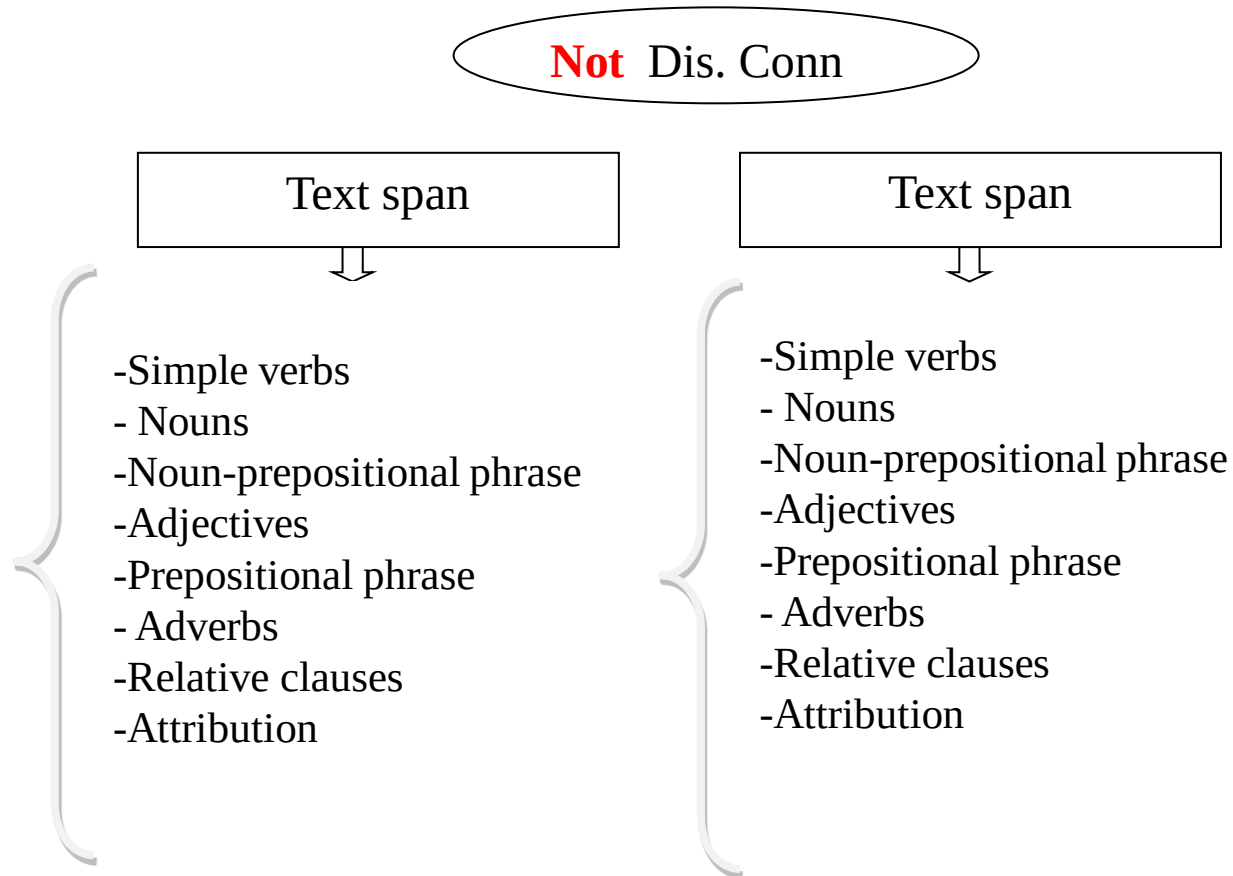
Discourse Connectives: Syntactic Categories



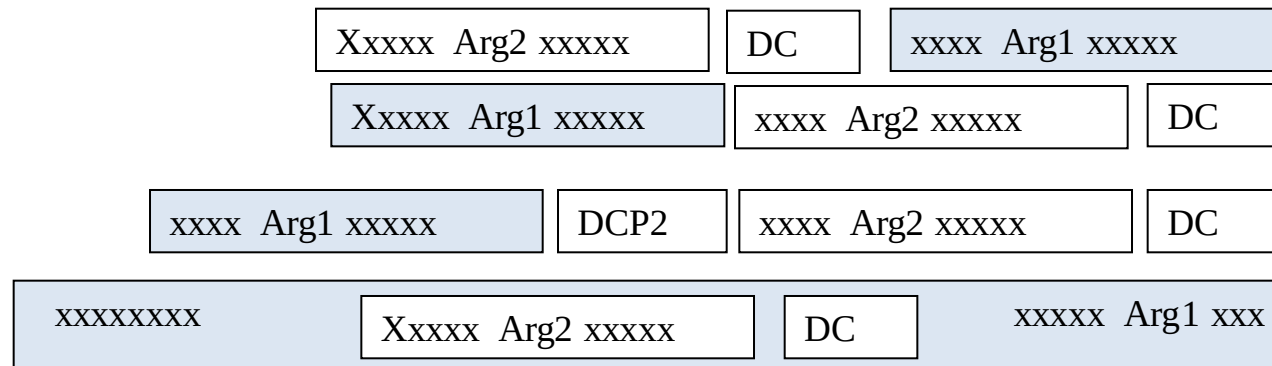
UNIVERSITY OF LEEDS

- Coordinating Conjunctions
- Subordinating Conjunctions
 - Simple
 - Paired
- Adverbs & prepositional phrases
 - Ex: ‘لذلك’/therefore - ‘بالتالي’/consequently - ‘بسبب’/because of
- Prepositions * new for Arabic *
 - Ex: ‘ب’ by/of/using - ‘ل’/l/to/due to - ‘جراء’/^grA/because of
- Nouns: * new for Arabic *
 - ‘بغية’/bgyp/desire - ‘بيد’/byd/





- ❑ Adjacent and non-adjacent arguments
- ❑ Flexible order



Al-maSdar or verbal nouns

- Al-masdar is a noun forming an action denoted by a corresponding verb such as وصول/arrival from the verb وصل/to arrive and محاولة/attempt from the verb حاول/to try
- The common morphological patterns of al-MaSdar are around 60 patterns in major Arabic grammar
- MaSdar nouns do not fit on one grammatical category in English

MaSdar	Morph. pattern	English gloss
سباحة	فعالة	swimming
انعكاس	انفعال	reflection
تجربة	تفعلة	experiment
حرب	فعل	war

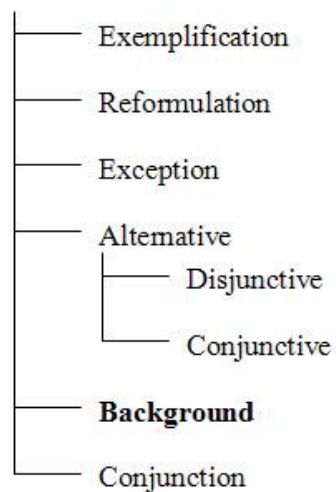
Discourse Relations – Senses



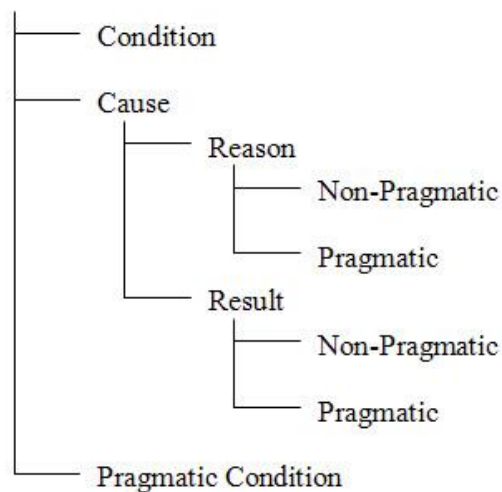
UNIVERSITY OF LEEDS

- 17 discourse relations

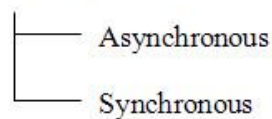
EXPANSION



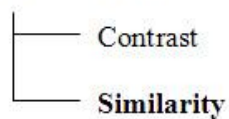
CONTINGENCY



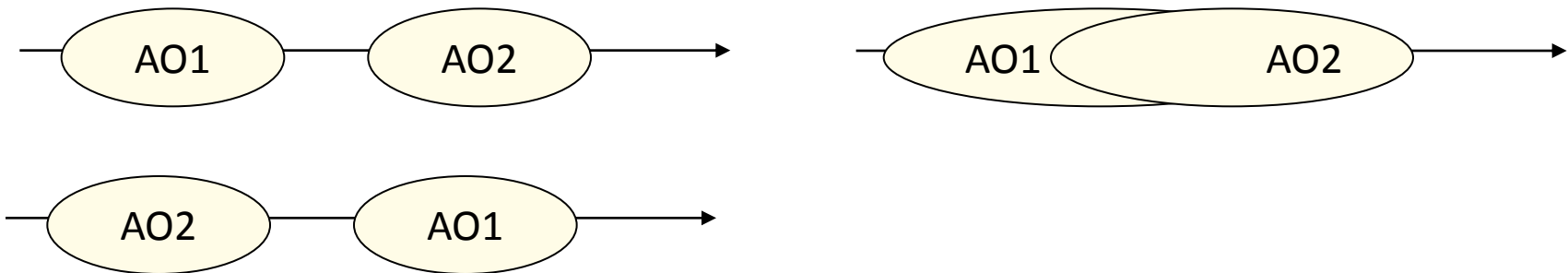
TEMPORAL



COMPARISON



- TEMPORAL Relations



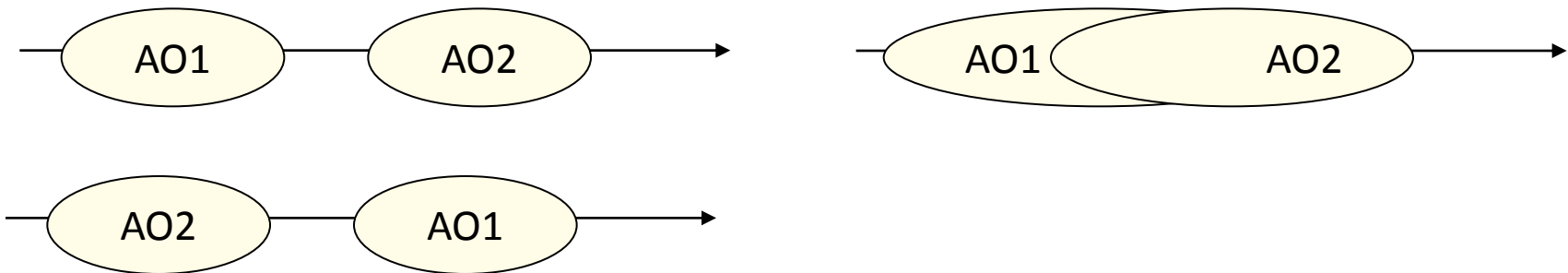
Asynchronous زمنية ترتيبية

Synchronous زمنية توافقية

Examples:

تم افتتاح المعرض الثقافي العالمي يوم الأحد مع وصول جميع ممثلين اللجان الثقافية في البلدان المشاركة
{The international cultural exhibition was opened on Sunday with the arrival of all cultural committees' representatives in participating countries}

- TEMPORAL Relations



Asynchronous زمنية ترتيبية

Synchronous زمنية توافقية

Examples:

واعربت عن قلقها للاستخدام المفرط للقوة "من قبل إسرائيل منذ بدء المواجهات في 28 ايلول/سبتمبر
She expressed a concern at *the excessive use of force by Israel* since **starting of the**
clashes on September 28

CONTINGENCY.Cause:

The situations in Arg1 and Arg2 are causally influenced and the two **are not in a conditional relation**

- Reason سببية
 - Pragmatic مجازي
 - Non-pragmatic مباشر
- Result نتيجة
 - Pragmatic مجازي
 - Non-pragmatic مباشر

CONTINGENCY.Cause. Reason. NonPragmatic

بلغت استراليا الدور النهائي لمسابقة كأس ديفيس لكرة المضرب يتقدمها على البرازيل 3-صفر

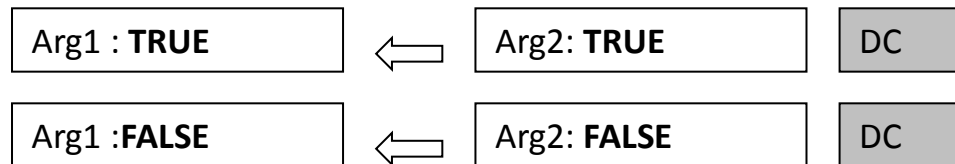
{Australia reached the final round of the Davis Cup Tennis Tournament because of her progress against Brazil 3 – zero}

CONTINGENCY.Cause. Result. NonPragmatic

ادخل المستشفى سبعة فلسطينيين من بينهم ثلاثة صحافيين اليوم السبت و استنكرت لجنة المصور الصحفي الفلسطيني اليوم السبت في بيان "الاعتداء الحاقد من قبل المستوطنين على الصحفيين"

Seven Palestinians, including three journalists were sent to the hospital on Saturday.
(and) the Committee of Palestinian press photographers condemned on Saturday in a statement: “vindictive attack by settlers on the journalists”

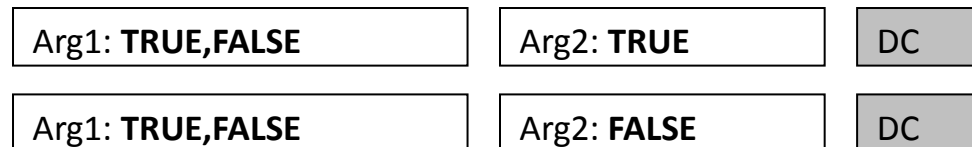
CONTINGENCY. Condition



ان اجلاء الطاقم سيبدأ بعد الظهر إذا تأكد تحسن الاحوال الجوية

The evacuation of the crew will happen this afternoon if weather conditions improve

CONTINGENCY. Pragmatic Condition



إذا دخلت المطبخ هناك علبة بسكويت على الطاولة

If you get in the kitchen, there is a box of biscuit on the table

COMPARISON. Contrast

نجح أحمد في الإمتحانات بينما توقع المدرس فشله
Ahmad succeeded in the exam while his teacher expected him to fail

COMPARISON. Similarity

New for Arabic

صوت النائب الثاني لصالح الحزب الديموقراطي كما فعل الكثير في العام المنصرم
The vice president voted for the Democratic Party as **many people did last year**

- EXPANSION الاستزادة
 - Exemplification التمثيل
 - Alternative التخيير
 - Disjunctive المتعارض
 - Conjunctive المتوافق
 - Exception الاستثناء
 - Background الخلفية
 - Reformulation إعادة الصياغة
 - Conjunction الربط

EXPANSION. Background

New for Arabic

- غادر الرئيس جورج بوش العراق بخيبة أمل من إيجاد حل سياسي للحرب على الإرهاب في العراق. وقد بدأت الحرب في العراق عام 2005 اثر مزاعم امريكية بنية العراق امتلاك السلاح النووي

President George W. Bush, left Iraq with disappointing feeling for finding a political solution to the terrorism war in Iraq. (and qad) The war in Iraq began in 2005 after U.S. allegations that Iraq has intention of acquiring nuclear weapons

Discourse Annotation Tool



UNIVERSITY OF LEEDS

Arabic Discourse Annotation Tool

ملف الروابط المحذوف مساعدة

رابط جديد

العائلة الرابط

الحدث الثاني الحدث الأول

ملاحظة:

الجزء الثاني رابط ثاني؟ ☐

ملف التمثيل

الروابط الدالة في النص الروابط المقترحة تمست روابط

>>> <<<

>>> <<<

Discourse Annotation Tool



UNIVERSITY OF LEEDS

Arabic Discourse Annotation Tool

20000815_AFP_ARB_0057

AFP_ARB.0057_20000815

القدس 4 0067 ش 0169 اسر / اقب-حرر 88 فلسطين/اسرائيل/القدس

معنى القدس يتجدد قوى تجريم حصول فلسطيني المدينة على الجنسية الاسرائيلية

القدس 8-15 (أف ب) - جدد معنى القدس والادار الفلسطينية الشيخ عكرمة صيري اليوم الثلاثاء فؤاد تجريم حصول فلسطيني المدينة المقدسة على الجنسية الاسرائيلية.

وقال المفتي في حديث لوكالة فرانس برس " لقد اصدرنا قوى لهذه الخصوص قبل عشر سنوات لكننا رايانا التاكيد مرة اخرى على جرمه التجنس بالجنسية الاسرائيلية لفلسطينيين من سكان القدس".

بأوضح صيري ان تجديد الفتوى جاء بعد صدور تقارير عن وزارة الداخلية الاسرائيلية تفيد ان 183 فلسطينيا مقدسيا تقدموا للحصول على الجنسية الاسرائيلية خلال العام الحالي مقابل 98 خلال نفس الفترة من العام الماضي.

وأعبر صيري ان قيام عرب من القدس بطلب الجنسية الاسرائيلية "يعني الاعتراف بالوجود الاسرائيلي في المدينة المقدسة وزيادة عدد الاسرائيليين فيها مقابل تقليص عدد سكانها العرب" وقال " وهذا جرم شنيعا".

وكانت قمة كامب ديفيد الفلسطينية الاسرائيلية الاخيرة فشلت بسبب خلافات الطرفين الكبيرة حول قضايا الوضع النهائي ولا سيما القدس والملاجئين والحدود.

رابط جديد

المقالة: الرابط:

المصدر الأول: المصدر الثاني:

ملاحظة:

الجزء الثاني ☐ رابط عشوائي

حفظ التعليل

روابط الدلالة في النص

روابط المقترحة

لوحة روابط

Loc: 56..57 : ف

Loc: 57..58 : ل

Loc: 99..100 : ف

Loc: 115..116 : ف

Loc: 116..117 : ل

Loc: 174..175 : ب

Loc: 194..195 : ر

Loc: 245..246 : ف

Loc: 251..252 : ب

Loc: 263..264 : ف

Loc: 264..265 : ل

>>

<<

16/04/2011

Discourse Annotation Tool



UNIVERSITY OF LEEDS

Arabic Discourse Annotation Tool

File Connectives Align Help

20000815_AFP_ARB_0110

Press CTRL when select more than one relation

Discourse Relations for Arabic العلاقات

- TEMPORAL الزمنية
 - Asynchronous غير الزمنية
 - Synchronous الزمنية
- CONTINGENCY الاستمرارية
 - Cause السببية
 - Reason السبب
 - NonPragmatic سبب غير عملي
 - Pragmatic سبب عملي
 - Result النتيجة
 - NonPragmatic نتيجة غير مباشرة
 - Pragmatic نتيجة مباشرة
 - Condition الشرطية
 - Pragmatic Condition الشرطية العملية
- COMPARISON المقارنة
 - Contrast التباين
 - Similarity التشابه
- EXPANSION الإضافة
 - Exemplification التمثيل

Connective

Arg 1

Comment

Paired Cons?

Loc: 82..83 : ب

Loc: 120..121 : ج

Loc: 159..160 : د

Loc: 168..169 : هـ

Loc: 193..194 : و

Loc: 208..209 : ز

Loc: 302..304 : ح

Loc: 352..353 : ط

Loc: 381..382 : ي

Loc: 442..443 : ك

Loc: 463..465 : ل

مسح/نسخ/لصق / ج ب 261 افب

واضاف ان اسرائيل كانت شددت على ان "يبقى 11 حيا يهوديا في المدينة المقدسة تحت سيادتها وان تسحب هذه السيادة الى مستوطنات مجاورة في الضفة الغربية وخصوصا غابات زلفي وعالي ادوميم"، من دون ان يوضح مصير احياء القدس ذات الغالبية العربية.

واكد بن عامي من جهة اخرى ان اسرائيل اغلقت "ابوابها لا يمكن ان تتخلى عن ارضها في جبل الهيكل في القدس" ووجد الحرم القدسي.

واوضح الوزير ان مجموعات المستوطنات التي تعزم اسرائيل ضمها للاحتفاظ بـ 80% من المستوطنين في الضفة الغربية "ستضم خصوصا قطاعات ارييل وعوش اريون ومناطق محيطه بالقدس باتجاه البحر الميت".

ولم يشر بن عامي لا الى مستوطنات قطاع غزة ولا الى الحيا اليهودي التي تسيطر عليه اسرائيل في الخليل حيث الحرم الابراهيمي، احد الاماكن المقدسة لدى اليهود والمسلمين.

واعتر ان ضم هذه المناطق والاعتراف الدولي بالسيادة الاسرائيلية على احياء يهودية اقيمت منذ 1967 في القدس الشرقية، سيسبب "تراجعا كبيرا".

واخيرا، اعرب بن عامي عن الامل في ان لا يقول اجتماع لجنة القدس التابعة لمنظمة المؤتمر الاسلامي في 28 آب/اغسطس في المغرب "الباب امام التفاوض".

- Task1: Is Potential Connective a Discourse Connective?
- Task2: Do both annotators assign the same Discourse Relation?
- Task3: Do they agree on Argument1 and Argument2 spans?

Disagreement studies

Why? Who is correct? to which depth they agree?

Task1: Discourse Connective?

ان اسرائيل تعتبر ان ايران ستبدأ بالحصول على قدرات نووية

لِاغراض عسكرية

Israel believes Iran is to begin acquiring a nuclear capability **for** the purposes of military

Task2: Discourse Relation/Relations?

سنشن هجمات شرسة على التجمعات العسكرية في كشمير المحتلة

حتى نجعل من يوم الاستقلال يوما للدمار

We will launch a fierce attack on military gatherings in occupied Kashmir

In order to make the Independence Day a day of destruction

{Temporal. Synchronous / Temporal. Asynchronous / Cause. Reason. NonPragmatic / Cause. Result.
NonPragmatic}

عجزوا عن فرض السيطرة على المجريات إذ اهتزت شبكاتهم بشكل مبكر

They were unable to impose control over the events, as their goal network was hit early

{Cause.Reason.Pragmatic / Temporal. Asynchronous /Temporal. Synchronous}

Task3: Argument boundaries?

تمثل قضية مستقبل القدس الشرقية التي **احتلتها اسرائيل** و**ضممتها عام 1967** مستقبل الفلسطينيين

The issue of the future of East Jerusalem, which **Israel occupied** and **annexed in 1967**, the future of the Palestinians

هي اكبر صفقة في تاريخ الانتقالات السعودية حتى الان والتي **استفاد منها الفريق كثيرا** **لـ**تجديد عقود ابرز اللاعبين الاخرين.

It is the biggest transfer deal in the football history of Saudi Arabia so far, (and) which **benefited the team greatly** **to** **renew the contracts of famous other players.**

- Arg2 is more than one sentence:

الجنرال سوموزا يتخلى عن الحكم
بعد معارك استمرت خمسة اسابيع ضد الساندينينيين واسفرت عن مقتل حوالي 40 الف مدني

General Somoza resigns from power

After fighting lasted five weeks against the Sandinistas, (and) killed around 40 thousand civilians



- Automatic correction:
 - Simple technical disagreement such as inclusion of punctuations, function words (ان/An) and attributions in the arguments
 - Assigning Conjunction relations to و/w connective at Beginning of Paragraph
- Manual judgments by third expert of all disagreed instances of discourse connectives, relations and arguments

All connectives

Observed agreement (conn)	95%
---------------------------	-----

Kappa (conn)	0.88
--------------	------

Observed agreement (rel)	86%
--------------------------	-----

Kappa (rel)	0.8
-------------	-----

Excluding wa at BOP

Observed agreement (conn)	94%
---------------------------	-----

Kappa (conn)	0.83
--------------	------

Observed agreement (rel)	76%
--------------------------	-----

Kappa (rel)	0.71
-------------	------

	Arg1	Arg2
Exact match =1	2478 (42%)	4186 (76%)
Exact match =0	677 (12%)	4
Partial match	2386 (43%)	1351 (24%)
Average $agr(ann1,ann2)$	78%	95%

Files	534
Discourse connectives	6328
Distinct connectives	80
Distinct relations	17 (+ 38 combined)
Clitic connectives	4779 (76%)

Ambiguity in Discourse connectives

discourse vs. non-discourse usage

ambiguous in usage	46
unambiguous in usage	34
Total connectives	80

Ambiguity in Discourse Relations

Ambiguous- signalled by ≤ 5 conn	5
Ambiguous- signalled by > 5 conn	11
unambiguous	1
Total relations	17

Most frequent connectives in the LADTB

Connective	# Rel	Common relations	Dis.Conn	Non-conn	Total
w_و	31	{76%:EXPANSION.Conjunction(3070)}; {7%:EXPANSION.Reformulation(287)}; {4%:EXPANSION.Background(184)}; {3%:CONTINGENCY.Cause.Result.NonPragmatic(134)};	3999 (55%)	3376 (45%)	7375
l_ل	4	{93%:CONTINGENCY.Cause.Reason.NonPragmatic(437)}; {5%:CONTINGENCY.Cause.Result.NonPragmatic(25)};	468 (11%)	3838 (89%)	4306
lkn_لكن	5	{97%:COMPARISON.Contrast(198)};	204 (98.5%)	3 (1.5%)	207
bEd_بعد	7	{51%:TEMPORAL.Asynchronous(100)}; {39%:CONTINGENCY.Cause.Reason.NonPragmatic/TEMPORAL.Asynchronous(76)};	194 (62%)	121 (38%)	315
xIAI_خلال	1	{100%:TEMPORAL.Synchronous(102)}	102 (81%)	24 (19%)	126
f_ف	13	{29%:CONTINGENCY.Cause.Result.NonPragmatic(29)}; {20%:CONTINGENCY.Cause.Reason.NonPragmatic(20)}; {1%:CONTINGENCY.Cause.Reason.Pragmatic/EXPANSION.Background(1)}; {1%:CONTINGENCY.Cause.Reason.NonPragmatic/EXPANSION.Exemplification(1)}	99 (6.5%)	1426 (93.5%)	1525
b_ب	4	{89%:CONTINGENCY.Cause.Reason.NonPragmatic(86)}; {5%:TEMPORAL.Synchronous(5)};	96 (2%)	4072 (98%)	4168
qbl_قبل	2	{98%:TEMPORAL.Asynchronous(83)};	84 (52%)	77 (48%)	161
lAn_لان	1	{100%:CONTINGENCY.Cause.Reason.NonPragmatic(77)}	77 (73%)	29 (27%)	106

Leeds Discourse Arabic Treebank

LADTB v.1 will be distributed soon via LDC. The tool and annotation scheme will be available online next month

UNIVERSITY OF LEEDS

```
74 ##### Explicit Conn #####  
75 بعد; bEd  
76 486..489; Sentence_Seq:2; Token_Seq:{19}; (1, 5, 0)  
77 ##### FEATURES #####  
78 PREP; Simple; TEMPORAL.Asynchronous  
79 Arg1_Conn_Arg2  
80 ##### ARG1 #####  
81 410..484  
82 Sentence_Seq:2; Token_Seq:{2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17}; (1, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1, 0), (1, 1, 1, 1, 1), (1, 1, 1,  
2), (1, 2, 0), (1, 2, 1, 0), (1, 3, 0), (1, 4, 0, 0), (1, 4, 0, 1, 0), (1, 4, 0, 2, 0), (1, 4, 0, 2, 1, 0, 0), (1, 4, 0, 2, 1, 0, 1, 0), (1, 4, 0,  
2, 1, 1, 0), (1, 4, 0, 2, 1, 1, 1, 0, 0), (1, 4, 0, 2, 1, 1, 1, 1, 0)  
83 شمال شرق -LRB- مساء الاربعاء المدينة متوجها الى ولاية اوهايو -RRB- عام 45 -LRB- غادر كنت  
84 gAdr knt -LRB- 45 EAmA -RRB- msA' AlArbEA' Almdynp mtwjha ALY wLAyp AwhAYw -LRB- $mAl $rq  
85 ##### ARG2 #####  
86 490..562  
87 Sentence_Seq:2; Token_Seq:{20,21,22,23,24,25,26,27,28,29,30,31}; (1, 5, 1), (1, 5, 2, 0, 0), (1, 5, 2, 0, 1, 0), (1, 5, 2, 0, 2, 0), (1, 5, 2, 0, 2,  
1, 0), (1, 5, 2, 0, 2, 1, 1, 0, 0), (1, 5, 2, 0, 2, 1, 1, 1, 0), (1, 5, 2, 0, 2, 1, 1, 2, 0), (1, 5, 2, 0, 2, 1, 1, 3, 0, 0), (1, 5, 2, 0, 2, 1, 1,  
3, 1, 0, 0), (1, 5, 2, 0, 2, 1, 1, 3, 1, 0, 1, 0), (1, 5, 2, 0, 2, 1, 1, 3, 1, 0, 2, 0), (1, 5, 2, 0, 2, 1, 1, 3, 1, 0, 2, 1), (1, 5, 2, 0, 2, 1, 1,  
3, 1, 0, 2, 2)  
88 ان استقل احد باصات شركة غريهوند الشهيرة التي تجوب كل الولايات الاميركية  
89 An Astql AHd BASAt $rqp gryhAwnd Al$hyrp AltY tjwb kl AlwLAyAt AlAmyrkyp  
90 _____  
91 _____  
92 ##### Explicit Conn #####  
93 لا!ان; AlA#An  
94 592..598; Sentence_Seq:3; Token_Seq:{8,9}; (1, 4, 0), (1, 4, 1)  
95 ##### FEATURES #####  
96 ADV#FUNCuWORD; MoreThanToken; COMPARISON.Contrast  
97 Arg1_Conn_Arg2  
98 ##### ARG1 #####  
99 567..590  
100 Sentence_Seq:3; Token_Seq:{2,3,4,5,6}; (1, 0), (1, 1, 0), (1, 1, 1), (1, 2, 0), (1, 2, 1)  
101 بدا ستيفن كنت نحيلا جدا  
102 bdA styfn knt nHylA jdA  
103 ##### ARG2 #####  
104 598..657  
105 Sentence_Seq:3; Token_Seq:{10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25}; (1, 4, 2, 0, 0, 0), (1, 4, 2, 0, 1, 0), (1, 4, 2, 0, 1, 1, 0), (1, 4,  
2, 1), (1, 4, 2, 2, 0, 0), (1, 4, 2, 2, 0, 1, 0), (1, 4, 2, 2, 0, 2, 0), (1, 4, 2, 2, 0, 2, 1), (1, 4, 2, 2, 0, 3, 0), (1, 4, 2, 2, 0, 3, 1, 0, 0),  
(1, 4, 2, 2, 0, 3, 1, 0, 1), (1, 4, 2, 2, 0, 3, 1, 1, 0), (1, 4, 2, 2, 0, 3, 1, 1, 1, 0, 0), (1, 4, 2, 2, 0, 3, 1, 1, 1, 1, 0, 0), (1, 4, 2, 2, 0,  
3, 1, 1, 1, 1, 0, 0), (1, 4, 2, 2, 0, 3, 1, 1, 1, 1, 1, 1, 0), (1, 4, 2, 2, 0, 3, 1, 1, 1, 1, 1, 1, 2, 0), (1, 4, 2, 2, 0, 3, 1, 1, 1, 1, 1, 1, 2, 1,  
0, 0), (1, 4, 2, 2, 0, 3, 1, 1, 1, 1, 2, 1, 0, 1, 0), (1, 4, 2, 2, 0, 3, 1, 1, 1, 1, 1, 1, 2, 1, 0, 2, 0)  
106 ه اغتسل وحلق ذقن ه ل لمرة الاولى منذ فترة لا بد ان تكون طويلة  
107 h Agtstl w Hlq *qn h l Almrrp AlAwly mn* ftrp lA bd An tkwn Twylyp  
108
```

Thank you ..



UNIVERSITY OF LEEDS

- **British Academy** for funding the annotation
- **Imam University** for sponsoring the project
- **Dr. Hussein Abdul-Raof** for linguistic advice
- The annotators:
 - **Latifa Alsulaiti** and **Abdul-baqi Sharif** in the actual annotation
 - **Basmah Al-Soli**, **Boshra Al-shyban** and **Maryam Al-Gawi** in the pilot annotation.
- **PDTB** team for helpful discussions
- My email: assaif@comp.leeds.ac.uk