

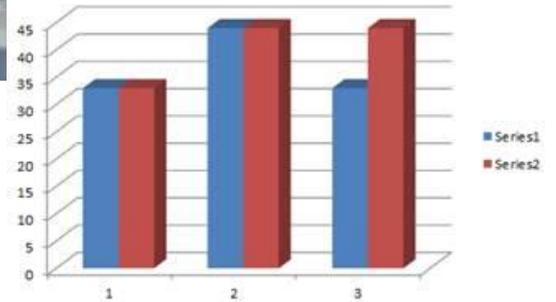
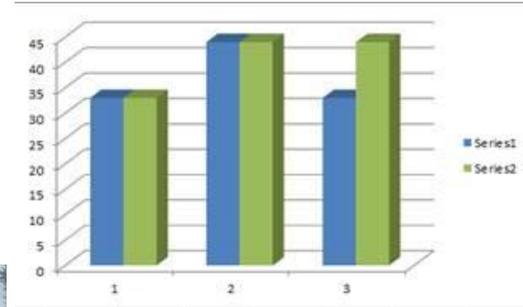
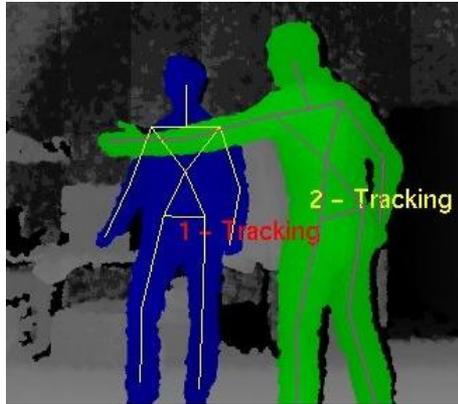
Modeling Multilingual Grounded Language

BILL DOLAN

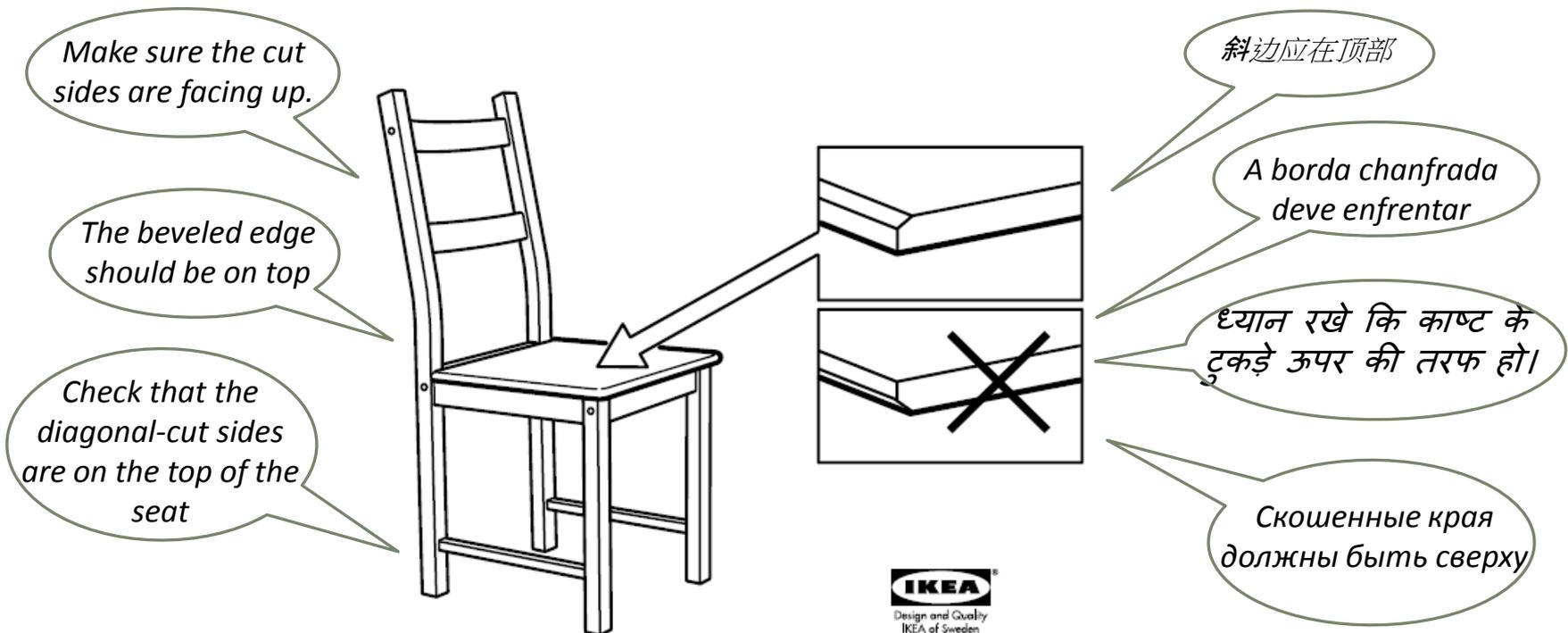
NATURAL LANGUAGE PROCESSING GROUP, MICROSOFT RESEARCH

XLITE: CROSS-LINGUAL TECHNOLOGIES

DECEMBER 7, 2012



Same Intent, Different Words



Grounded Multilingual Semantics

Goal: a pragmatic, application-driven model of natural language semantics that is:

- Able to handle both monolingual and multilingual alternations in a robust way
- Data-driven
- Grounded, actionable: captures a mapping between language and objects/actions in the world



What clusters of utterances seem to “mean the same thing”, as far as some desired action in the world is concerned?

Agenda

Translation

- Statistical machine translation as a semantic modeling framework

Paraphrase as Translation

- A statistical machine translation model of paraphrase generation

Grounding as Translation

- Learning to map language to objects, states, events, and actions

Statistical Machine Translation

If you have a hammer...

The screenshot shows the Bing Translator interface. At the top, there are navigation links: WEB, IMAGES, VIDEOS, MAPS, TRANSLATOR, and MORE. The Bing logo and 'MS Beta' are in the top left. Below the navigation, there are 'TRANSLATE' and 'HELP' buttons. The main heading is 'The Pivot'. The 'Translate from' dropdown is set to 'Russian (Auto-Detected)' and the 'Translate to' dropdown is set to 'English'. An orange 'Translate' button is to the right of the dropdowns. Below the dropdowns, there are two text boxes. The left box contains the Russian text 'Во времена Диккенса и Достоевского |' and has a placeholder 'Enter text or webpage URL' at the bottom. The right box contains the English translation 'During the times of Dickens and Dostoevsky'. At the bottom right of the right box, there is a 'Rate this translation' section with checkmark, X, and refresh icons. A 'Clear All' button is located at the bottom right of the left text box.

Translation Equivalence as Shallow Semantics

Vivre près d'un garage ou une station d'essence pourrait quadrupler le risque de leucémie infantile, suggère une étude française

Vivir cerca de una gasolinera puede llegar a cuadruplicar el riesgo de leucemia en niños.

A child who lives near a petrol (gas) station is four times more likely to develop leukemia than a child who lives far away from one, according to a new study



Bilingual Alignment

	four	times	more	likely	to	develop	leukemia
quadrupler	■	■	□	□	□	□	□
le	□	□	■	■	□	□	□
risque	□	□	■	■	□	□	□
de	□	□	□	□	■	■	□
leucémie	□	□	□	□	□	□	■
infantile	□	□	□	□	□	□	□

The Pivot

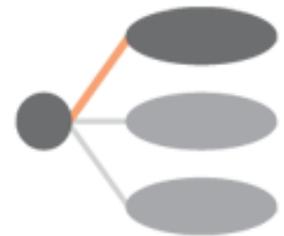
Translators can (and do) translate the same phrases in one language into different phrases in another.

- Translators differ
- Contexts differ

Statistical machine translation models utilize parallel corpora created by human translators.

- Two or more sentences, pivoted on a single semantic “intent”
- Large datasets are potentially available

Paraphrase as Translation



Defining Paraphrase

A paraphrase is a pair of formally different expressions that can be interpreted as semantically equivalent in some context

- A good approximation of semantic equivalence is ordinarily sufficient
- Exact synonymy or semantic equivalence (as in entailment relations) is usually of less importance.

Note: paraphrase is mostly accidental

- Paraphrase usually arises because humans independently approach the same information using different forms of expression.

Paraphrase as Translation

Vivre près d'un garage ou une station d'essence pourrait quadrupler le risque de leucémie infantile, suggère une étude française

Vivir cerca de una gasolinera puede llegar a cuadruplicar el riesgo de leucemia en niños.

A child who lives near a petrol (gas) station is four times more likely to develop leukemia than a child who lives far away from one, according to a new study

Living near to a petrol station or garage may increase the risk of acute childhood leukaemia by 400%.

Children who live in close proximity to gas stations and auto body shops have a dramatically higher rate of leukemia, according to a new study.

Living near a petrol station may quadruple the risk for children of developing leukaemia, new research says.

Children who live near petrol stations may be four times more susceptible to leukaemia.

Monolingual Alignment

	four	times	more	likely	to	develop	leukemia
quadruple	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
the	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
risk	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
for	<input type="checkbox"/>						
children	<input type="checkbox"/>						
of	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
developing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
leukemia	<input type="checkbox"/>	<input checked="" type="checkbox"/>					

An Uncommon Event

A four-metre python exploded when it tried to swallow a two-metre alligator whole in Florida.

Animal experts in the US are baffled after a huge snake exploded when trying to snack on a giant alligator.

Scientists in Florida are puzzling over a Burmese python that scarfed down a six-foot alligator before its stomach ruptured.

A python in Florida's Everglades clashed with a 6-foot-long alligator. And won. But then tried to swallow him whole. And exploded.

Scientists in Florida still aren't quite sure how a 13-foot Burmese python managed to devour a six-foot alligator in the Everglades.

A 13-foot Burmese python recently burst after it apparently tried to swallow a live, six-foot alligator whole, authorities said.

SMT for Paraphrase

Use news text to build statistical translation models of paraphrase:

- *Quirk et al., 2004*
- *Dolan et al., 2004*

Microsoft Research Paraphrase Corpus (2005):

- 5800 paraphrase pairs extracted from news articles
- Download from: <http://research.microsoft.com/en-us/downloads/607d14d9-20cd-47e3-85bc-a2f65cd28042/default.aspx>
- *Dolan & Brockett, 2005*
- *Brockett & Dolan, 2005*

A Stumbling Block

Naturally-occurring paraphrase data is hard to obtain

- There is no such profession as a “paraphraser”
- Full-sentence paraphrases are relatively rare
 - Most paraphrases, even in news data, are fragmentary
 - Usually a single source, edited
- People find it difficult to paraphrase

A major challenge for ML approaches to the problem of modeling monolingual alternations.

The Pivot

Translators can (and do) translate the same phrases in one language into different phrases in another.

- Translators differ
- Contexts differ

Statistical machine translation models utilize parallel corpora created by human translators.

- Two or more sentences, pivoted on a single semantic “intent”
- Large datasets are potentially available.

Aligned data intended for use in SMT models can be a rich source of paraphrase data, if we pivot on the foreign language phrases.

- *Bannard & Callison-Burch, 2005*
- *Callison-Burch et al. 2006*
- *Callison-Burch, 2008*
- *Zhao, et al., 2008*

Pivots as Bipartite Graphs

JAPANESE - ENGLISH

セルホン cellphone
 セルホン cellular
 セルホン cellular phone
 携帯電話 cell
 携帯電話 cell phone
 携帯電話 cellphone
 携帯電話 cellular phone
 携帯電話 mobile
 携帯電話 mobile phone
 携帯電話 phone

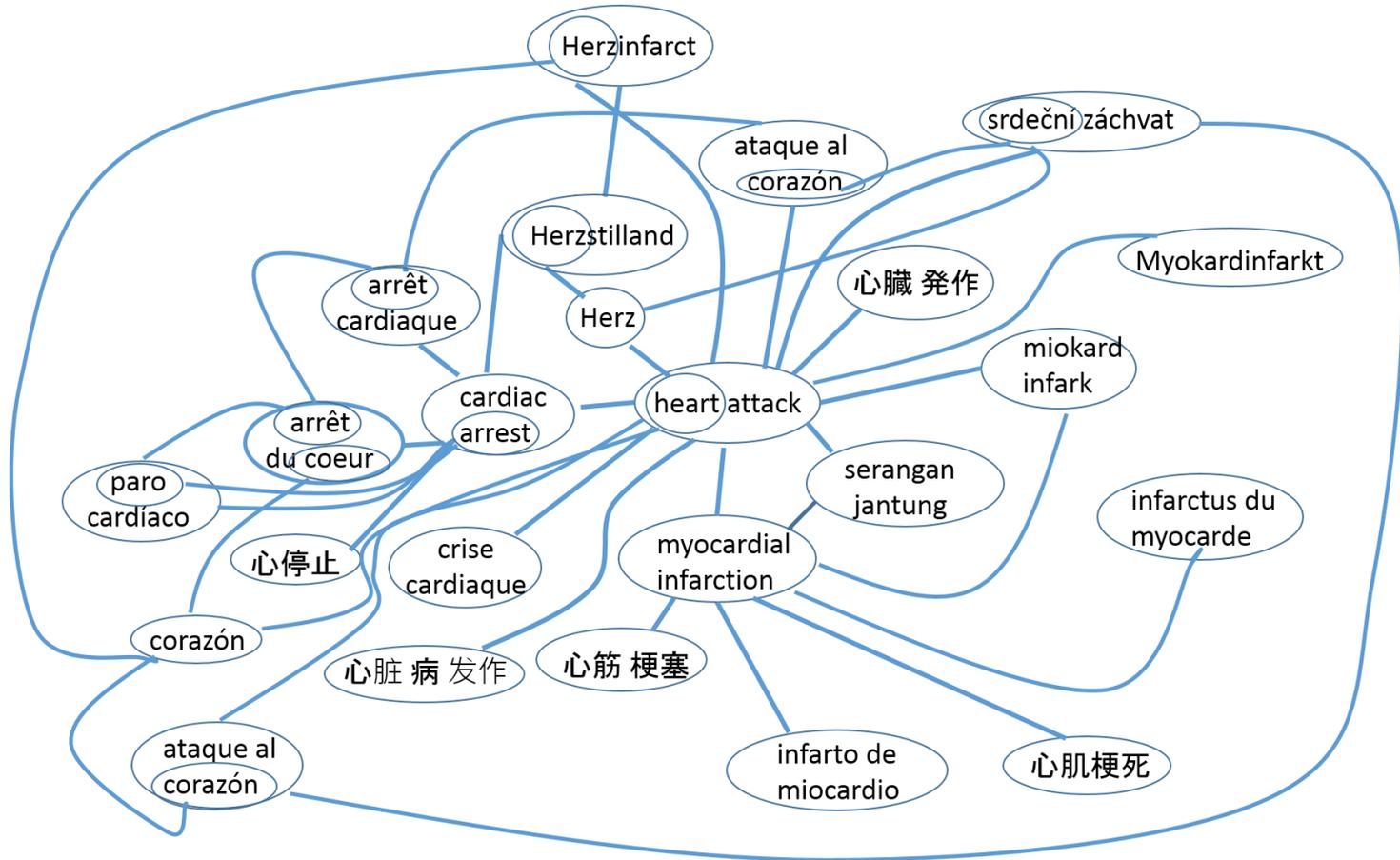
CHINESE - ENGLISH

手机 cell
 手机 cell phone
 手机 cellphone
 手机 cellular phone
 手机 mobile
 手机 mobile phone

Under the Hood

携帯 " , for mobile	携帯 " from mobile	携帯 " when cell
携帯 " applicant for registration "	携帯 " how cheap mobile	携帯 " which controls home appliance using cellular
携帯 " applicant for registration	携帯 " in the mobile	携帯 " wireless
携帯 " applicant for	携帯 " keitai credit	携帯 " you hear the phone
携帯 " applicant	携帯 " keitai credit	携帯 # on your mobile
携帯 " campaign so that mobile	携帯 " keitai	携帯 & cell
携帯 " cell	携帯 " mobile carriers to	携帯 & mobile
携帯 " cellphones	携帯 " mobile carriers	携帯 > cell
携帯 " cellular phone	携帯 " mobile	携帯 > regional information - mobile
携帯 " diamond " mobile	携帯 " newsy	携帯 ' a mobile
携帯 " encouraging " mobile carriers to	携帯 " on your	携帯 ' awareness
携帯 " encouraging " mobile carriers	携帯 " portability	携帯 ' cell
携帯 " encouraging " mobile	携帯 " portable	携帯 ' cellphones and
	携帯 " the mobile	
	携帯 " the service promises to provide mobile	

Bilingual Data, Multilingual Space



A Random Walk Model

Treat the pivot data as a multipartite graph

- *Kok & Brockett, NAACL/HLT 2010*
- Implement a random walk over the graph
- MT pivot meets graphical models for search (*Mei, et al., CIKM 2008*)
- Use translations among other languages as “bridge” languages

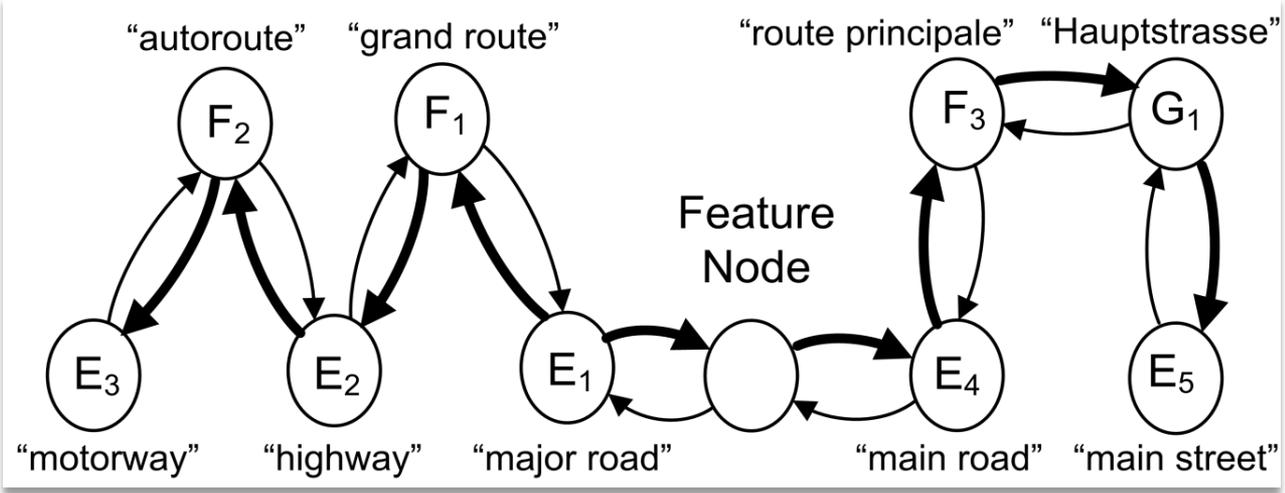
Phrase tables from Europarl parallel data:

- 6 languages, 15 language pairs, 1M sentences ea.

Results from random 100 phrase sample (1...4-grams)

- Found paraphrases for all (vs. 49 in CCB 2008)
- 0.61 good at top-1
- 0.42 good in top-10 (~3 x CCB 2008)

Bridge Languages



SMT for Paraphrase Generation

Desiderata:

- Generate paraphrases, up to and including whole sentences
- Provide a robust delivery system
- Add in data from other sources
- Generate new contextual paraphrase candidates – not just memorized ones

SMT for Paraphrase Generation

Solution:

- Exploit our existing phrasal SMT pipeline and large language models
- Use bipartite phrasal mappings as training corpus
 - Currently: 350M paired phrases from 17 language pairs
 - Future: 560 M paired phrases from 24 language pairs
 - Require each phrase to have been seen in at least 2 language pairs
- Build a monolingual phrasal machine translation system
 - Train weights using Max BLEU on 2K sentence pair data set selected for diversity
- Add features derived from original translation scores

The Quality Bar

SMT meets an information need that cannot otherwise be met

- A user may be satisfied with low-grade readable output
- n-best lists are exposed only in special cases

With paraphrase, the grammatical quality bar may be higher

- The human user knows what the input means, is a fluent speaker
- We will be exposing the n-best list, not normal in SMT

Present initial prototype as a contextual thesaurus

DEMO

Home | Users | Developers | Webmasters | Tools | Partners | Labs

Contextual Thesaurus

[Provide Feedback](#)

Translate from English to English to explore alternate ways of expressing the same idea. ([Learn more ...](#))

enter your phrase

please feel free to ask for help

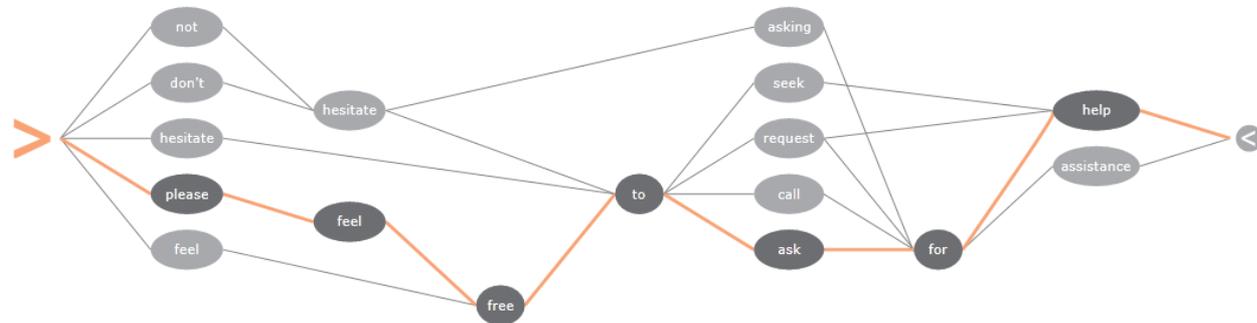
Hint: Enter short phrases (about 4-8 words) in a business or formal style. To see another random example, refresh your browser (F5).

clear

suggestions 

please feel free to ask for help

- feel free to ask for assistance
- feel free to call for help
- hesitate to ask for assistance
- don't hesitate to ask for assistance
- hesitate to call for help
- don't hesitate to call for help
- hesitate to seek help



Contextual Thesaurus

Characteristics of the English-English phrasal MT system

- It tends to prefer more common words and sequences
- It has no grammatical knowledge
- Output tends to be monotone, with only local inversions
- Output is best on short sequences of words -- 3-8 words typically
- Better in certain domains: business, finance, government

Paraphrase Public API

Contextual Thesaurus web page:

<http://labs.microsoftTranslator.com/thesaurus>

Information about public Paraphrase API:

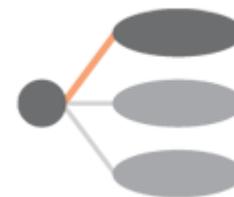
- <http://msdn.microsoft.com/en-us/library/hh847648.aspx>

Information about subscribing to the API with an Azure subscription:

- <http://msdn.microsoft.com/en-us/library/hh847649>

Contact to boost subscription limit for research projects:

- mtlic@Microsoft.com



Lexical Diversity

Decoding in SMT is NP-complete problem (Knight 1999)

- Use heuristic limits on search to limit the search space
- SMT traditionally seeks single best translation -- usually by optimizing BLEU

Paraphrase generation is premised on diversity

- Most naturally-occurring paraphrases are similar to each other
- In paraphrase, we often want to expose to the user a selection of alternatives
- The number of potential alternatives is huge – a 10 word sentence with 10 synonyms per word yields 10^{10} combinations
- Risky to expose to user because lower ranked alternatives may be of poorer grammatical and semantic quality

Stress Testing

tend to eat under stress
tend to overeat under stress
tend to over-eat under stress
tends to overeat under stress
tended to overeat under stress
tend to overeat in stress
usually overeat under stress
tend to overfeed under stress
tendency to overeat under stress
tend to graze under stress
tend to eating under stress
inclined to eat under stress
tends to overeat in stress
tend to overeat on stress
tend to eat too much under stress
inclined to overeat under stress

tend to over-eat in stress
tend to overeat with stress
tend to overeat by stress
tend to binge under stress
tend to overeat under tension
tend to overeat under pressure
tendency to over-eat under stress
tends to overfeed under stress
tended to overeat in stress
tend to eat too under stress
are inclined to eat under stress
tend to over-eat under tension
tend to over-eat under pressure
usually eating under stress
tends to overeat on stress
tend to overeat the stress
tend to overeat during stress

tendency to overeat in stress
tend to over-eating under stress
tends to graze under stress
tends to over-eat under stress
tend overeating under stress
tends to overeat with stress
generally overeat under stress
tends to overeat by stress
prone to overeat under stress
often overeat under stress
tend to overeat under strain
tend eating under stress
tends to overeat under tension
generally eating under stress
tends to overeat under pressure
tend to eat in stress
inclined to eating under stress

The Office Nuisance

a sure way to annoy one's boss
one sure way to irritate one 's boss
a sure way to irritate one 's boss
a surefire way to irritate one 's boss
the sure way to irritate one 's boss
surefire way to irritate one 's boss
one sure way to annoy one 's boss
a sure-fire way to irritate one 's boss
a sure fire way to irritate one 's boss
one sure way to irritate one 's the boss
a guaranteed way to irritate one 's boss
one sure way to irritate one 's head
a sure way to annoy one 's boss
a safe way to irritate one 's boss
a certain way to irritate one 's boss
one sure way to antagonize one 's boss
one sure way to upset one 's boss
surefire ways to irritate one 's boss

a sure way to irritate one 's the boss
is sure way to irritate one 's boss
sure-fire way to irritate one 's boss
is a way to irritate one 's boss
one sure way to disturb one 's boss
for sure way to irritate one 's boss
a sure way to irritate one 's head
one sure way to annoy your boss
a way to irritate one 's boss
assured way to irritate one 's boss
sure fire way to irritate one 's boss
a secure way to irritate one 's boss
a sure way to antagonize one 's boss
a sure way to upset one 's boss
a surefire way to irritate one 's the boss
one sure way to annoy his boss
reliable way to irritate one 's boss
the sure way to annoy one 's boss
one sure way to tease one 's boss
confident way to irritate one 's boss

certain way to irritate one 's boss
safe way to irritate one 's boss
a sure way to disturb one 's boss
the sure way to irritate one 's the boss
a surefire way to irritate one 's head
unsure way to irritate one 's boss
a sure way to annoy your boss
secure way to irritate one 's boss
a surefire way to annoy one 's boss
one sure way to irritate one 's superiors
a reliable way to irritate one 's boss
surefire way to irritate one 's the boss
a surefire way to antagonize one 's boss
one sure way to annoy one 's the boss
one sure way to harass one 's boss

Context is Important

RETIREMENT ADVICE

retirement tips
pension advice
pensions advice
retirement counseling
retirement guidance
retirement consulting
retirement counselling
pension tips
pensions tips
pension counseling
resignation advice
superannuation advice
retirement consultation

TAX ADVICE

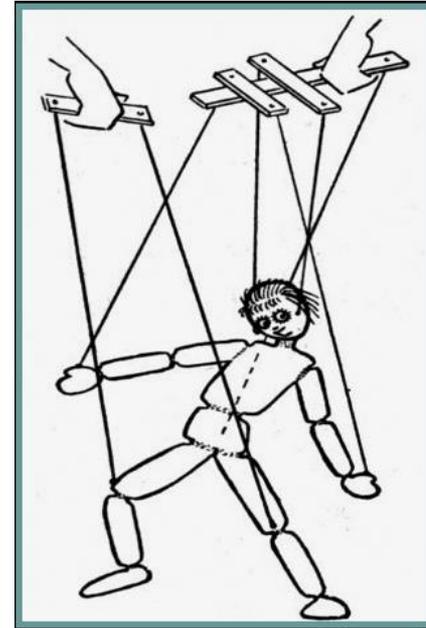
tax tips
tax counselling
taxation advice
tax consulting
tax counseling
tax advising
tax consultancy
tax advise
tax consultation
fiscal advice
tax guidance
tax advisory
vat advice

Another kind of Pivot: Query/URL

how to assemble ikea chair
assemble ikea furniture
assemble and disassemble ikea furniture
how to put ikea furniture together
how to put together ikea's furniture
how difficult is it to assemble ikea furniture
how to put together ikea
ikea directions on how to put ikea furniture together
Ikea do I need tools for self assembly
tools required for assembly of ikea furniture
best way to assemble ikea furniture
directions on how to put together ikea roger chairs
assebling ikea furniture
best way to assemble ikea furniture
is it difficult to build ikea furniture
put together ikea furniture



Grounding as Translation



Non-Linguistic Grounding

So far, we have pivoted on linguistic objects

- Contextual Thesaurus → phrases in other languages
- Queries -> documents

To make linguistic inputs actionable, though, we need data grounded in the world. And we'd like it in many languages.

And it doesn't exist.

One approach: crowdsourcing

Grounding in Video

Chen & Dolan, 2011

Use short video clips to elicit descriptions from people around the world

Parallel descriptions of same activities, in many languages



Translation data

Parallel descriptions in the same language



Paraphrase data (train SMT models on cross-product)

Pivoted on the action in the video

Crowdsourced Video Captions

A man eats spaghetti sauce.
A man is eating food.
A man is eating from a plate.
A man is eating something.
A man is eating spaghetti.
The man tried his pasta and sauce.
A man is eating spaghetti.
A man is eating.
The man ate some pasta from a bowl.
The man is eating.



A man is eating spaghetti from a large bowl while standing.
A man is eating spaghetti out of a large bowl.
A man tasting some food in the kitchen is expressing his satisfaction.

...in many languages

Slovene: Moški je špagete z vilico.

German: Ein Mann isst Spagetti

Romanian: Un barbat mananca paste.

Un barbat mananca spaghetti.

Un bucatar mananca ce a preparat.

French: Un homme mande des pates.

Spanish: Un gordo saborea un plato de pasta

Serbian: Čovek jede špagete.

Filipino: Linasahan ng kusinero ang kanyang pagkain.

Dutch: De luie kok neemt gulzig een hap van zijn bord spaghetti met worstjes.

Russian: Мужчина что-то ест из тарелки.

Tamil: ஒருவர் சாப்பிட்டுக்கொண்டு இருக்கிறார்.

ஒருவர் முள் கரண்டியால் உணவை சாப்பிடுகிறார்

மனிதன் சாப்பிட்டு கொண்டு இருக்கிறான்.



Language Distribution

English	85550 (33855)	Spanish	1883
Hindi	6245	Gujarati	1437
Romanian	3998	Russian	1243
Slovene	3584	French	1226
Serbian	3420	Italian	953
Tamil	2789	Georgian	907
Dutch	2735	Polish	544
German	2326	Chinese	494
Macedonian	1915	Malayalam	394

Also: Tagalog, Portuguese, Norwegian, Filipino, Estonian, Turkish, Arabic, Urdu, Hungarian, Indonesian, Malay, Bulgarian, Danish, Bosnian, Marathi, Swedish, Albanian

Dataset Publicly Available

<http://research.microsoft.com/en-us/downloads/38cf15fd-b8df-477e-a4e4-a4680caa75af/default.aspx>

Microsoft Research

Search Microsoft Research

Videos Projects Publications People Downloads

Home Our Research Connections Careers

Worldwide Labs Research Areas Research Groups

Find us on Facebook Follow us on Twitter

Home > Downloads > Microsoft Research Video Description Corpus

Microsoft Research Video Description Corpus

This data consists of about 120K sentences collected during the summer of 2010. Workers on Mechanical Turk were paid to watch a short video snippet and then summarize the action in a single sentence. The result is a set of roughly parallel descriptions of more than 2,000 video snippets. Because the workers were urged to complete the task in the language of their choice, both paraphrase and bilingual alternations are captured in the data. We expect this data to be useful for training and testing translation and paraphrase algorithms. A paper describing how the data was created and used is in progress.

Download Details	
File Name:	VideoDescriptions.msi
Version:	1.0
Date Published:	12 November 2010
Download Size:	2.54 MB

[Download](#)

Note: By installing, copying, or otherwise using this software, you agree to be bound by the terms of its license. [Read the license.](#)

Contact Us | Terms of Use | Trademarks | Privacy Statement

©2011 Microsoft Corporation. All rights reserved.

Microsoft

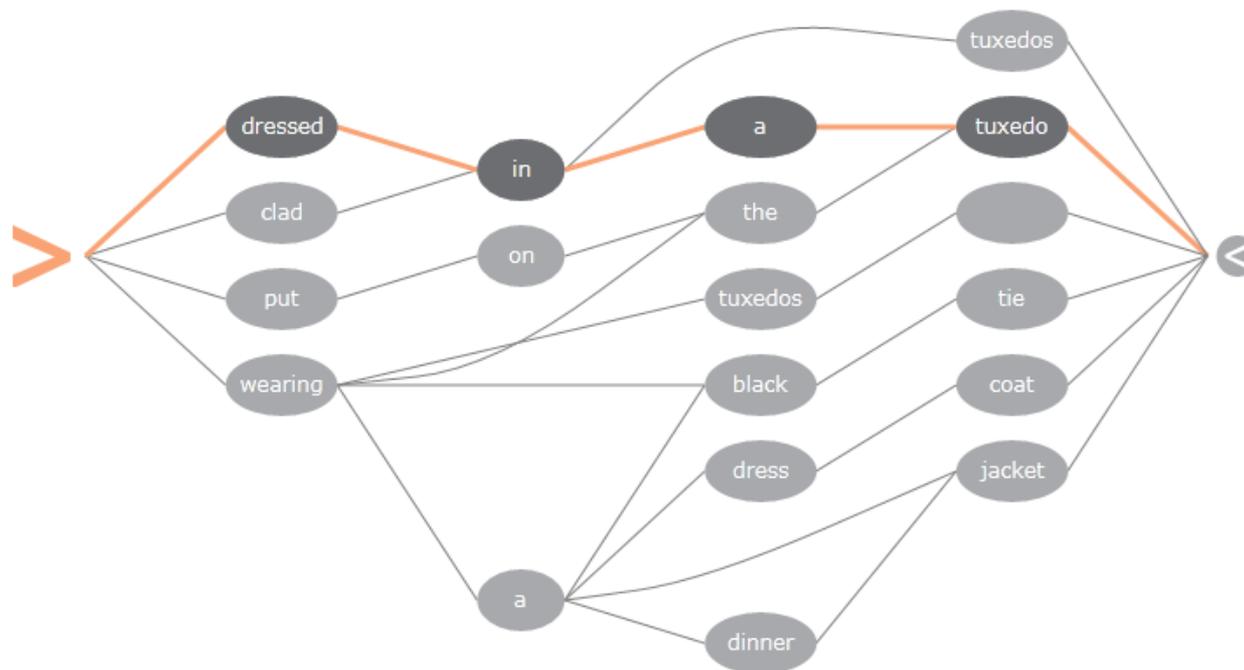
~2K videos
~120K captions
~\$5K crowd costs

Grounding in an Inventory of Virtual Objects

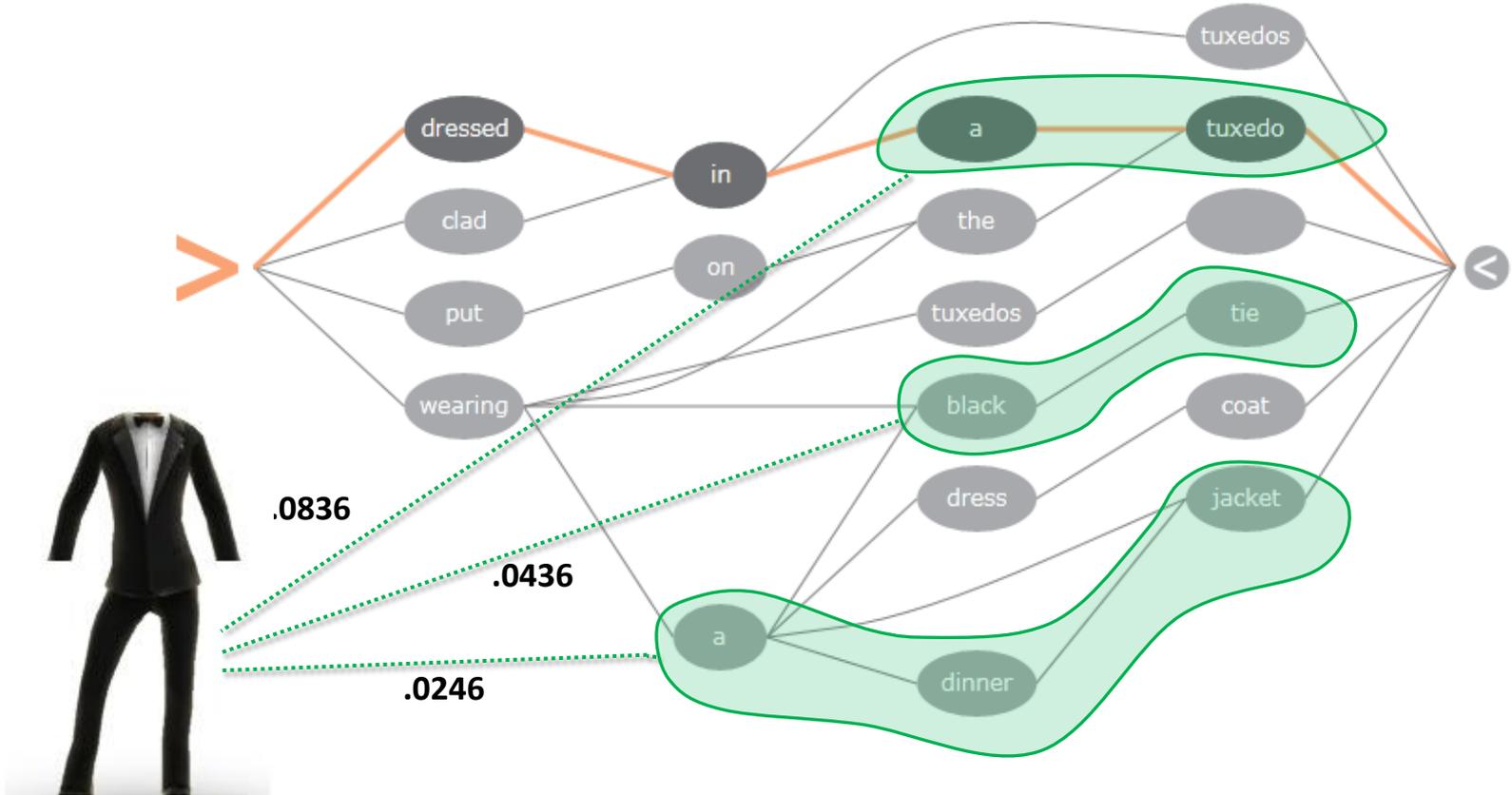


green color **flower design** hat
green and yellow **winter hat**, **daisy design** on hem and pom-pom on top
green and brown **striped beanie** with **white daisies** on brim and **pom-pom** on top
light brown **striped** hat with **flower pattern**.
green color cute hat with **white color flowers**
dark and light green **winter hat** with **white and yellow flower** on border
green, **flower print** elastic clothing hat
round green **striped beanie with flowers on around sides**
green striped **beanie cap with daisies around bottom** and green **puffball**
brown hat with yellow bands and **flower pattern**
green and brown mixed color round hat, **round ball** on top
green striped knit **beanie** with a **green ball** on top and **daisy pattern** rim
drab and yellow **beanie** with flowers
a green striped **beanie** with a **knob** on top and **sunflowers around the edge**
poop green and yellow **striped woolen hat** with **daisies**
adorable brown and mustard **striped skull cap** with **ball** top and **daisies on the rim**
green **bucket hat** with white & **yellow flower design** upon hat
green and beige **striped beanie** with **pom-pom** on top and **flowers on brim**
a bright **toboggan** with **daisies around** the hat and a **pom-pom**.
a yellow knit **cap** with white **stars around** the bottom edge
green and dark yellow striped snow hat with **daisies going around the rim**
green **toboggan** with **flowers** and **stripes**

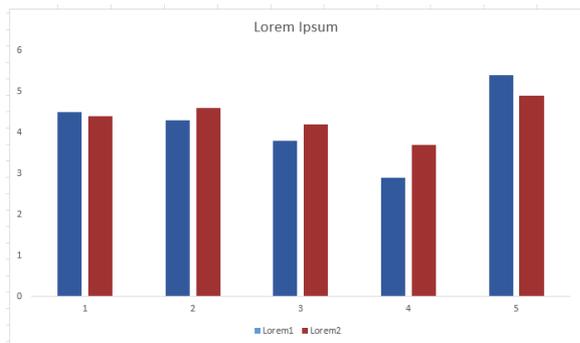
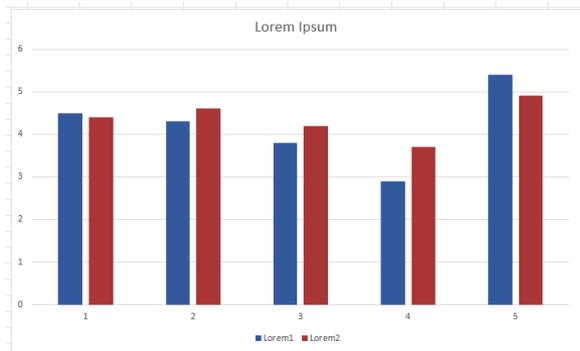
Linguistic “Neighborhoods”



Grounded Mappings

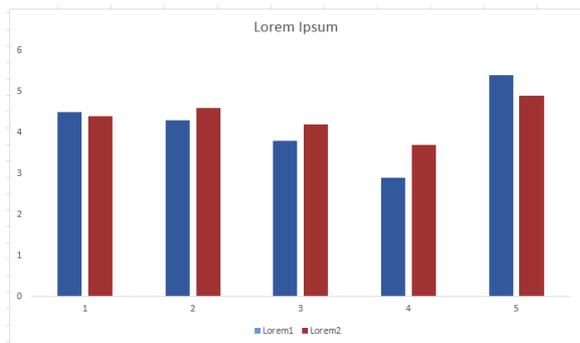
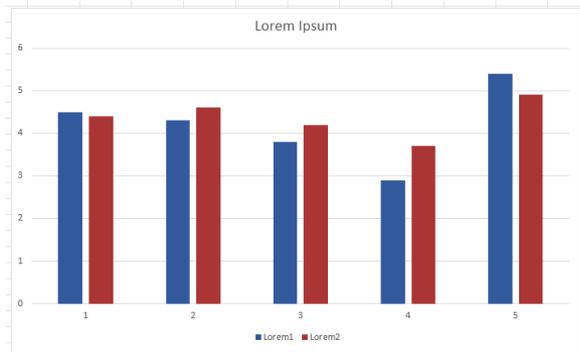


Demanding Change



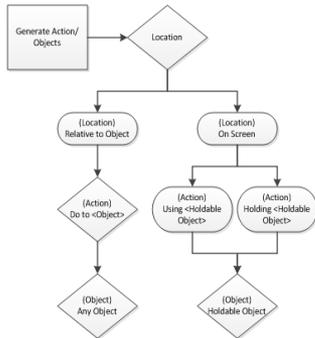
- Remove measurement lines from background
- remove graph lines
- Remove grid lines
- Remove the rows within the chart
- remove all grid lines .
- Remove horizontal lines.
- remove graph lines from chart
- remove rulings
- Remove lines from background
- Remove grid lines from chart.
- delete the background lines
- no lines
- Remove horizontal grid lines.
- Take away the horizontal graph lines that show in the background.
- Get rid of the lines in the background.
- Remove horizontal lines in graph
- remove the segment lines
- Remove vertical background lines.
- take out grid line
- remove all grid lines from the chart
- remove horizontal lines from behind the bars

Negotiating Change



Please hide the grid on the chart
 Let's see how it looks without grid lines.
 I think it would look better without the grid lines.
 I think those lines in the background make it look cluttered
 Please take away the vertical lines that represent the values on the left
 We need to remove the background lines from the graph.
 What about removing the grid lines?
 Can we remove all background lines except for the bottom line?
 What if we remove the lines?
 Can we remove the background grid?
 Let's see how it looks without the background lines
 Can you remove result number delineations?
 The grid lines aren't really needed, let's get rid of them.
 Would no horizontal lines look better ?
 Would it be better to remove the lines?
 Maybe you could take away the lines because they aren't very helpful anyway, and are distracting.
 It will look cleaner without the increment lines
 Try deleting the grid lines.
 Can you remove the scale lines going across on the background?

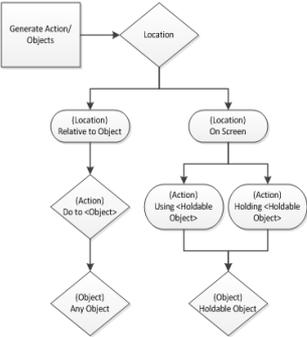
Grounding in Parameterized Video



Action			Object		
<i>On screen</i>	<i>On screen</i>	<i>Relative to object</i>	<i>Big</i>	<i>Small</i>	
<i>Using <Holdable object></i>	<i>Holding <Holdable object></i>	<i>Do to <Object></i>	Fire truck	<i>Other</i>	<i>Holdable</i>
Hit	Jump	Kick	Train	Oil can	Frying pan
Throw	Run	Punch	Car	Keyboard	Fork
Wave	Skip	Dodge	Airplane	Backpack	Spoon
Poke	Crawl	Throw	Bicycle	Lamp	Coffee mug
Drop	Walk	Pick up	Arrow	Ladybug	Juice box
Shoot	Dance	Smash	House	Disco ball	Cheese
Hug	Sit	Hug	Table	Skateboard	Soda can
Kiss	Spin in a circle	Kiss	Computer	Roller skates	Candy bar
Smell	Cough	Poke	TV	Butterfly	Lollipop
Hold high	Sneeze	Dust	Couch		Ice cream cone
Hold to the right	Blush	Hit	Chair		Pizza slice
Hold to the left	Shrug	Lump over	Bed		Hamburger



Code -> Animation



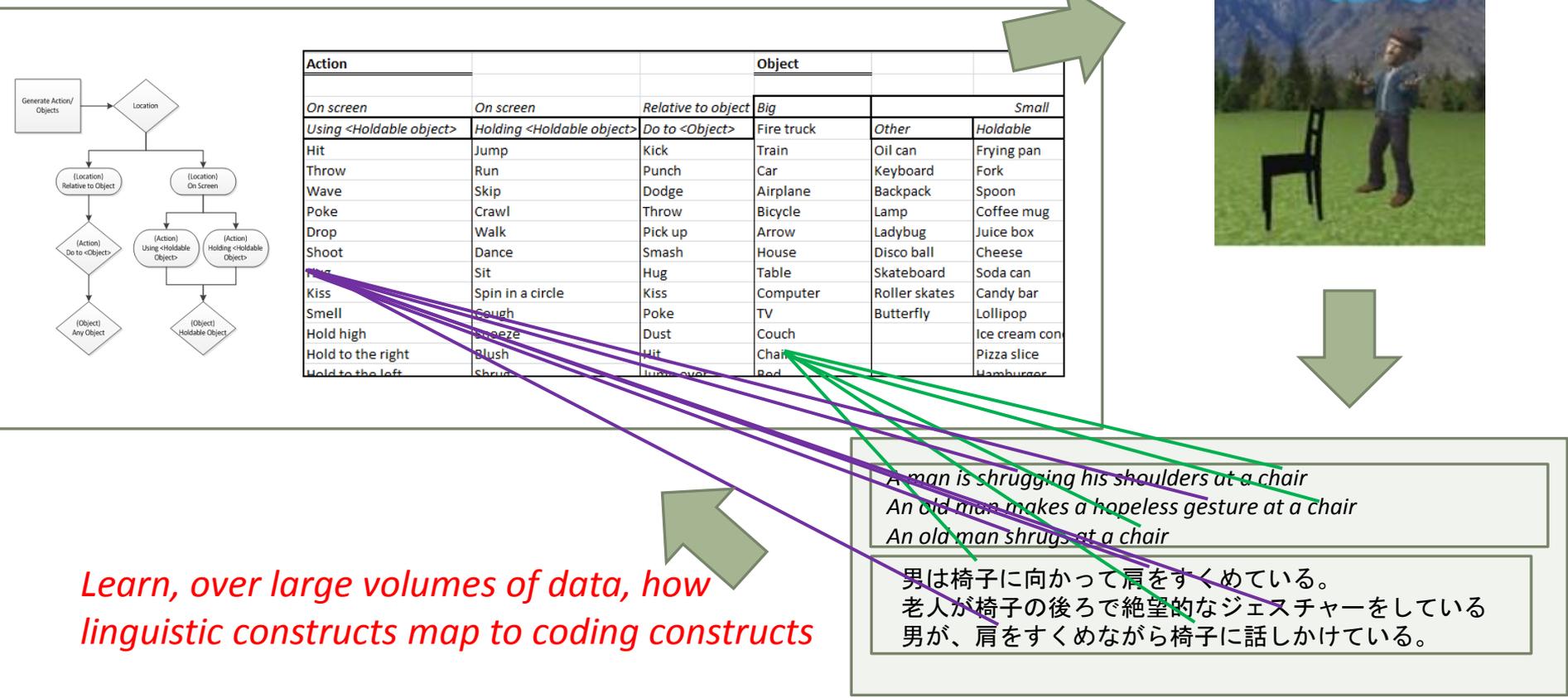
Action			Object		
On screen	On screen	Relative to object	Big	Small	
Using <Holdable object>	Holding <Holdable object>	Do to <Object>	Fire truck	Other	Holdable
Hit	Jump	Kick	Train	Oil can	Frying pan
Throw	Run	Punch	Car	Keyboard	Fork
Wave	Skip	Dodge	Airplane	Backpack	Spoon
Poke	Crawl	Throw	Bicycle	Lamp	Coffee mug
Drop	Walk	Pick up	Arrow	Ladybug	Juice box
Shoot	Dance	Smash	House	Disco ball	Cheese
Hug	Sit	Hug	Table	Skateboard	Soda can
Kiss	Spin in a circle	Kiss	Computer	Roller skates	Candy bar
Smell	Cough	Poke	TV	Butterfly	Lollipop
Hold high	Sneeze	Dust	Couch		Ice cream cone
Hold to the right	Blush	Hit	Chair		Pizza slice
Hold to the left	Shrug	Jump over	Red		Hamburger



A man is shrugging his shoulders at a chair
An old man makes a hopeless gesture at a chair
An old man shrugs at a chair

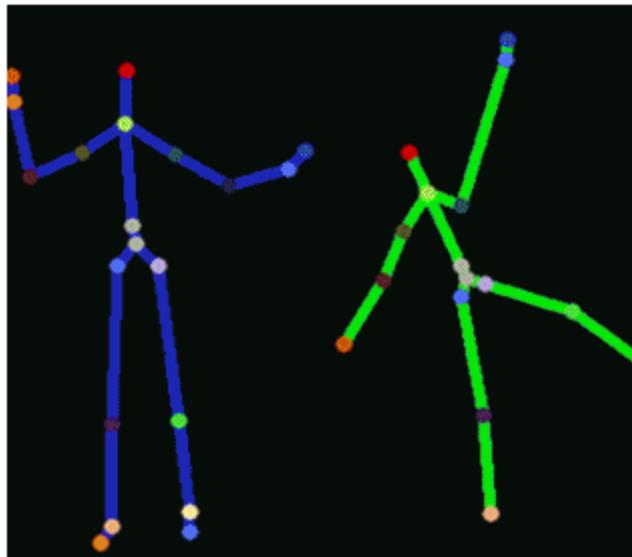
男は椅子に向かって肩をすくめている。
 老人が椅子の後ろで絶望的なジェスチャーをしている
 男が、肩をすくめながら椅子に話しかけている。

Animation -> Language -> Code



Learn, over large volumes of data, how linguistic constructs map to coding constructs

Language - Skeletal Tracking Data

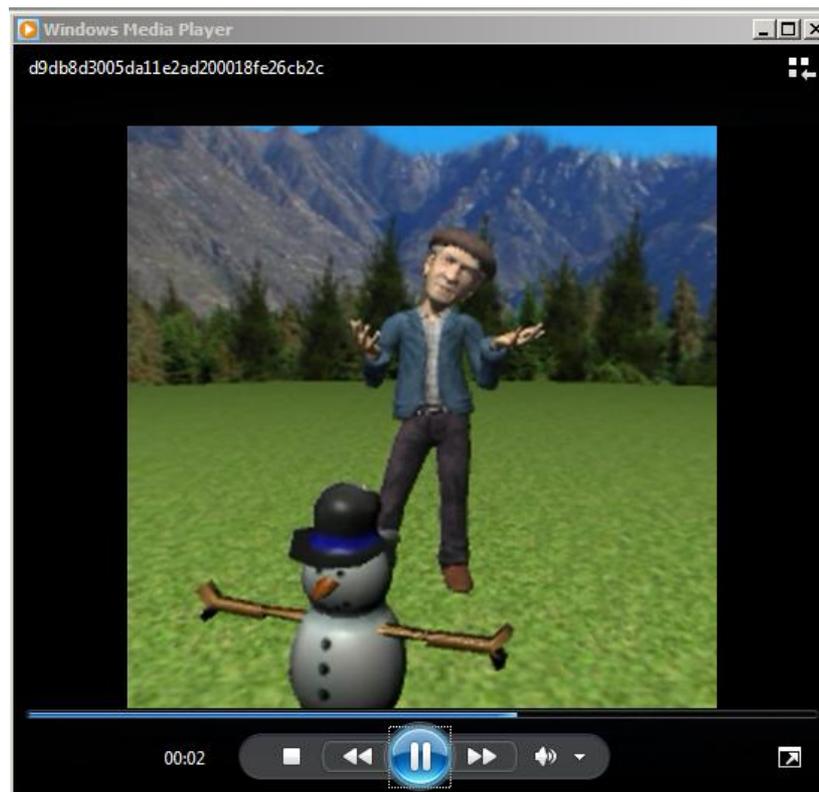


老人が雪だるまの後ろで絶望的なジェスチャーをしている。

男は雪だるまに向かって肩をすくめている。

A man is shrugging his shoulders at a snowman.

An old man is making a hopeless gesture behind a snowman.



Animation Captioning

bares her teeth and looks confrontational.
does not wave, and steps forward.
doesn't wave then stomps foot and looks angry
gets mad and steps forward.
had angry stance
has a growling face
has more dramatic facial expressions.
is acting aggressive.
is acting angry
is acting angry
is acting angry.
is acting very angry.
is angry and unfriendly
is angry looking.
is angry, not friendly
is demonstrating aggression
is flexing.
is gesturing angrily.
is growling and clinching
is making angry expressions
is making wrestling moves and an angry face.
is more aggressive
is pretending to be a sumo wrestler.
right one moved his right arm. Left moved both arms and legs

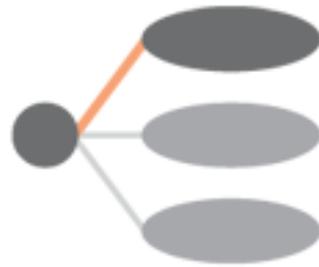


is ready to fight
is showing aggression.
is singing in an angry manner with hands down low
is standoffish and gestures angrily.
is telling about something that made her angry.
is threatening and aggressive
is upset which shows by her body language and facial expressions.
is very angry
is waving.
is wearing a ring
looks an angry and moves differently.
looks ready for a fight.
looks very angry.
makes angry gestures with mouth and fists.
makes an angry face
Quiet, then angry.
seems angry
seems angry.
she is mad.
shows her teeth looking angry
steps up and back and looks angry.
takes a step forward.
very angry

Grounded Multilingual Models

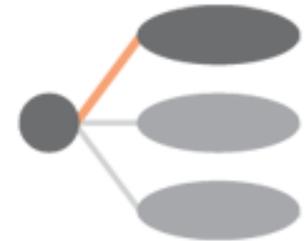
Goal: a data-driven, ML framework that simultaneously captures:

- Multilingual + monolingual “neighborhoods” tied to shared semantic “intent”
 - Language grounded in objects/actions/code state
-
- Statistical Machine Translation is a natural fit
 - Developed to capture meaning-preserving string transformations
 - Allows a unified approach to modeling bilingual + monolingual data
 - Can be readily adapted to add grounding features
 - Active research community feeds rapid development



Thanks

Query Expansion as Translation



Query Expansion as SMT

D: best home remedies for cold and flu

Q: how to deal with stuffy nose

Documents and queries as different languages

Cast QE as bridging the language gap via translation

- E.g. rewrite/translate a query Q to a title of a doc that is relevant to Q

Three lexicon models trained on billions of Q-D pairs mined from clickthrough logs