Coursework discussion

Coursework discussion

- 5636915

Coursework discussion

- 5636915
- Update to Inuktitut baselines

Coursework discussion

- 5636915
- Update to Inuktitut baselines
- University plagiarism policy

Bits and bobs

What do we care about?

- Models are trained to optimize **perplexity**.
- Models are (too frequently) evaluated using **BLEU**.
- Is this a problem?

Minimum risk

maximum likelihood

$$\hat{\boldsymbol{\theta}}_{\mathrm{MLE}} = \operatorname*{argmax}_{\boldsymbol{\theta}} \left\{ \mathcal{L}(\boldsymbol{\theta}) \right\}$$

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{s=1}^{S} \log P(\mathbf{y}^{(s)} | \mathbf{x}^{(s)}; \boldsymbol{\theta})$$

$$= \sum_{s=1}^{S} \sum_{n=1}^{N^{(s)}} \log P(\mathbf{y}_n^{(s)} | \mathbf{x}^{(s)}, \mathbf{y}_{<n}^{(s)}; \boldsymbol{\theta})$$

Minimum risk

minimum risk

$$\hat{\boldsymbol{\theta}}_{\mathrm{MRT}} = \operatorname*{argmin}_{\boldsymbol{\theta}} \Big\{ \mathcal{R}(\boldsymbol{\theta}) \Big\}.$$

$$\mathcal{R}(\boldsymbol{\theta}) = \sum_{s=1}^{S} \mathbb{E}_{\mathbf{y}|\mathbf{x}^{(s)};\boldsymbol{\theta}} \Big[\Delta(\mathbf{y}, \mathbf{y}^{(s)}) \Big]$$

$$= \sum_{s=1}^{S} \sum_{\mathbf{y} \in \mathcal{Y}(\mathbf{x}^{(s)})} P(\mathbf{y}|\mathbf{x}^{(s)}; \boldsymbol{\theta}) \Delta(\mathbf{y}, \mathbf{y}^{(s)})$$

Minimum risk

minimum risk

$$\hat{\boldsymbol{\theta}}_{\mathrm{MRT}} = \operatorname*{argmin}_{\boldsymbol{\theta}} \Big\{ \mathcal{R}(\boldsymbol{\theta}) \Big\}.$$

$$\tilde{\mathcal{R}}(\boldsymbol{\theta}) = \sum_{s=1}^{S} \mathbb{E}_{\mathbf{y}|\mathbf{x}^{(s)};\boldsymbol{\theta},\alpha} \Big[\Delta(\mathbf{y}, \mathbf{y}^{(s)}) \Big]$$

$$= \sum_{s=1}^{S} \sum_{\mathbf{y} \in \mathcal{S}(\mathbf{x}^{(s)})} Q(\mathbf{y}|\mathbf{x}^{(s)};\boldsymbol{\theta},\alpha) \Delta(\mathbf{y}, \mathbf{y}^{(s)})$$

Minimum risk

Minimum risk

System	Training	MT06	MT02	MT03	MT04	MT05	MT08
MOSES	MERT	32.74	32.49	32.40	33.38	30.20	25.28
RNNSEARCH	MLE	30.70	35.13	33.73	34.58	31.76	23.57
	MRT	37.34	40.36	40.93	41.37	38.81	29.23

Table 3: Case-insensitive BLEU scores on Chinese-English translation.

Minimum risk

System	Training	MT06	MT02	MT03	MT04	MT05	MT08
MOSES	MERT	32.74	32.49	32.40	33.38	30.20	25.28
RNNSEARCH	MLE	30.70	35.13	33.73	34.58	31.76	23.57
	MRT	37.34	40.36	40.93	41.37	38.81	29.23

Table 3: Case-insensitive BLEU scores on Chinese-English translation.

In practice: seems to be tuning for brevity penalty.

What is MT good at?

- Goodhart's law: When a measure becomes a target, it ceases to be a good measure.
- BLEU has been a target for over 15 years.
- What does BLEU tell us about accuracy?
- What doesn't it tell us?

Contrastive pairs

Sennrich 2017

Contrastive pairs

Sennrich 2017

- Basic idea: given a sentence and its correct translation, systematically corrupt translation to reflect some phenomenon that is known to be difficult to (machine) translate.

Contrastive pairs

Sennrich 2017

- Basic idea: given a sentence and its correct translation, systematically corrupt translation to reflect some phenomenon that is known to be difficult to (machine) translate.

- Since MT is just conditional language modeling, ask it which sentence is more probable.

Contrastive pairs

Sennrich 2017

- Basic idea: given a sentence and its correct translation, systematically corrupt translation to reflect some phenomenon that is known to be difficult to (machine) translate.

- Since MT is just conditional language modeling, ask it which sentence is more probable.

- Why doesn't this work for phrase-based MT?

Contrastive pairs

Sennrich 2017

- German determiners must agree with head noun in case, number, and gender.

English	… **of the** American **Congress**
correct	… **des** amerikanischen **Kongresses**
contrastive	* … **der** amerikanischen **Kongresses**

Contrastive pairs

Sennrich 2017

- Subjects and verbs must agree with one another in grammatical number and person.

English	… that the **plan will** be approved
correct	… dass der **Plan** verabschiedet **wird**
contrastive	* … dass der **Plan** verabschiedet **werden**

Contrastive pairs

Sennrich 2017

- Reverse polarity by deleting or inserting the negative particle *nicht* (`not'), swapping the determiner *ein* (`a') and is negative counterpart *kein* (`no'), or deleting/ inserting the prefix *un-*

English	… the timing is **uncertain**
correct	… das Timing ist **unsicher**
contrastive	* … das Timing ist **sicher**

Contrastive pairs

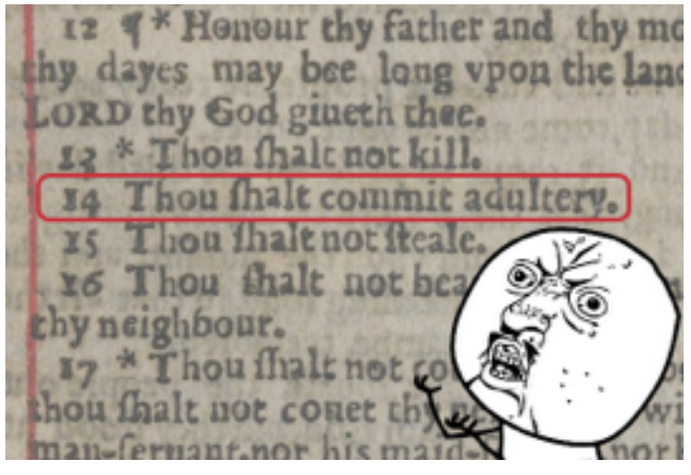
Sennrich 2017

- Reverse polarity by deleting or inserting the negative particle *nicht* (`not'), swapping the determiner *ein* (`a') and is negative counterpart *kein* (`no'), or deleting/ inserting the prefix *un-*

English … the timing is **uncertain**

correct … das Timing ist **unsicher**

contrastive * … das Timing ist **sicher**



Contrastive pairs

Sennrich 2017

- Transliteration: subword and character-level models should be able to copy or transliterate names. For unseen names, swap adjacent characters.

English		Mr. **Ensign's** office
correct		Senator **Ensigns** Büro
contrastive	*	Senator **Enisgns** Büro

Contrastive pairs

system (category and size→)	agreement: noun phrase 21813	agreement: subject-verb 35105	verb particle 2450	polarity (negation): insertion 22760	polarity (negation): deletion 4043	transliteration 3490
BPE-to-BPE	**95.6**	**93.4**	**91.1**	97.9	**91.5**	96.1
BPE-to-char	93.9	91.2	88.0	**98.5**	88.4	**98.6**
char-to-char	93.9	91.5	86.7	**98.5**	89.3	**98.3**
(Sennrich et al., 2016a)	98.7	96.6	96.1	98.7	92.7	96.4
human	99.4	99.8	99.8	99.9	98.5	99.0

Table 4: Accuracy (in percent) of models on different categories of contrastive errors. Best single model result in bold (multiple bold results indicate that difference to best system is not statistically significant).

Contrastive pairs

system	sentence	cost
source	Since then we have only played in the Swedish **league which is** not the same level.	
reference	Seitdem haben wir nur in der Schwedischen **Liga** gespielt, **die** nicht das gleiche Niveau **hat**.	0.149
contrastive	Seitdem haben wir nur in der Schwedischen **Liga** gespielt, **die** nicht das gleiche Niveau **haben**.	0.137
1-best	Seitdem haben wir nur in der schwedischen **Liga** gespielt, **die** nicht die gleiche Stufe **sind**.	0.090
source	FriendsFest: the **comedy show that taught** us serious lessons about male friendship.	
reference	FriendsFest: die **Comedy-Show, die** uns ernsthafte Lektionen über Männerfreundschaften **erteilt**	0.276
contrastive	FriendsFest: die **Comedy-Show, die** uns ernsthafte Lektionen über Männerfreundschaften **erteilen**	0.262
1-best	FriendsFest: die **Komödie** zeigt, **dass** uns ernsthafte Lehren aus männlichen Freundschaften	0.129
source	Robert Lewandowski **had** the best opportunities in the first half.	
reference	Die besten Gelegenheiten in Hälfte eins **hatte** Robert Lewandowski.	0.551
contrastive	Die besten Gelegenheiten in Hälfte eins **hatten** Robert Lewandowski.	0.507
1-best	Robert Lewandowski **hatte** in der ersten Hälfte die besten Möglichkeiten.	0.046

Table 6: Examples where char-to-char model prefers contrastive translation (subject-verb agreement errors). 1-best translation can make error of same type (example 1), different type (translation of *taught* is missing in example 2), or no error (example 3).

Who is MT for?

Who is MT for?

- Dissemination of information to other languages.

Who is MT for?

- Dissemination of information to other languages.
- Assimilation of information from other languages.

Who is MT for?

- Dissemination of information to other languages.
- Assimilation of information from other languages.

Not surprisingly, what the very word “translation” means for these two sets of people is entirely different. And I just would like to hope that you, the computational linguists of the future, will keep in mind the needs of both of these very worthy communities.

Martin Kay (2005)

Who is MT for?

- Dissemination of information to other languages.
- Assimilation of information from other languages.
- Translators

Who is MT for?

- Dissemination of information to other languages.
- Assimilation of information from other languages.
- Translators
  - Does MT make translators more efficient?

Who is MT for?

- Dissemination of information to other languages.
- Assimilation of information from other languages.
- Translators
  - Does MT make translators more efficient?
  - Do they like it?

Who is MT for?

- Dissemination of information to other languages.
- Assimilation of information from other languages.
- Translators
  - Does MT make translators more efficient?
  - Do they like it?
- Field linguists documenting languages

Neural interactive MT

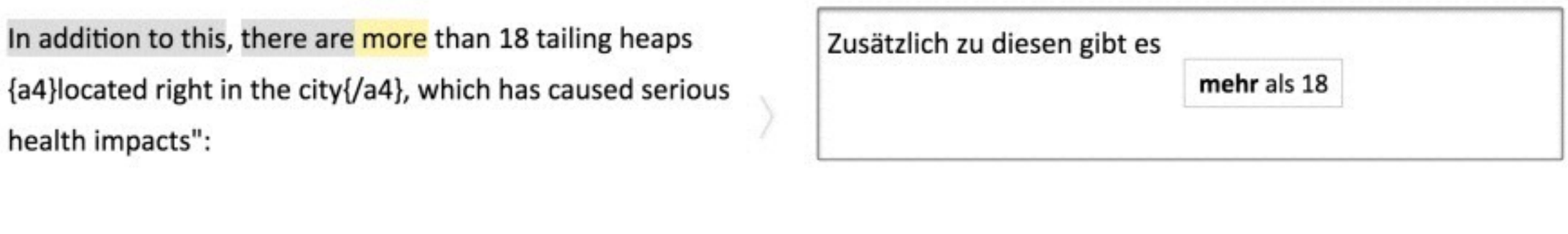
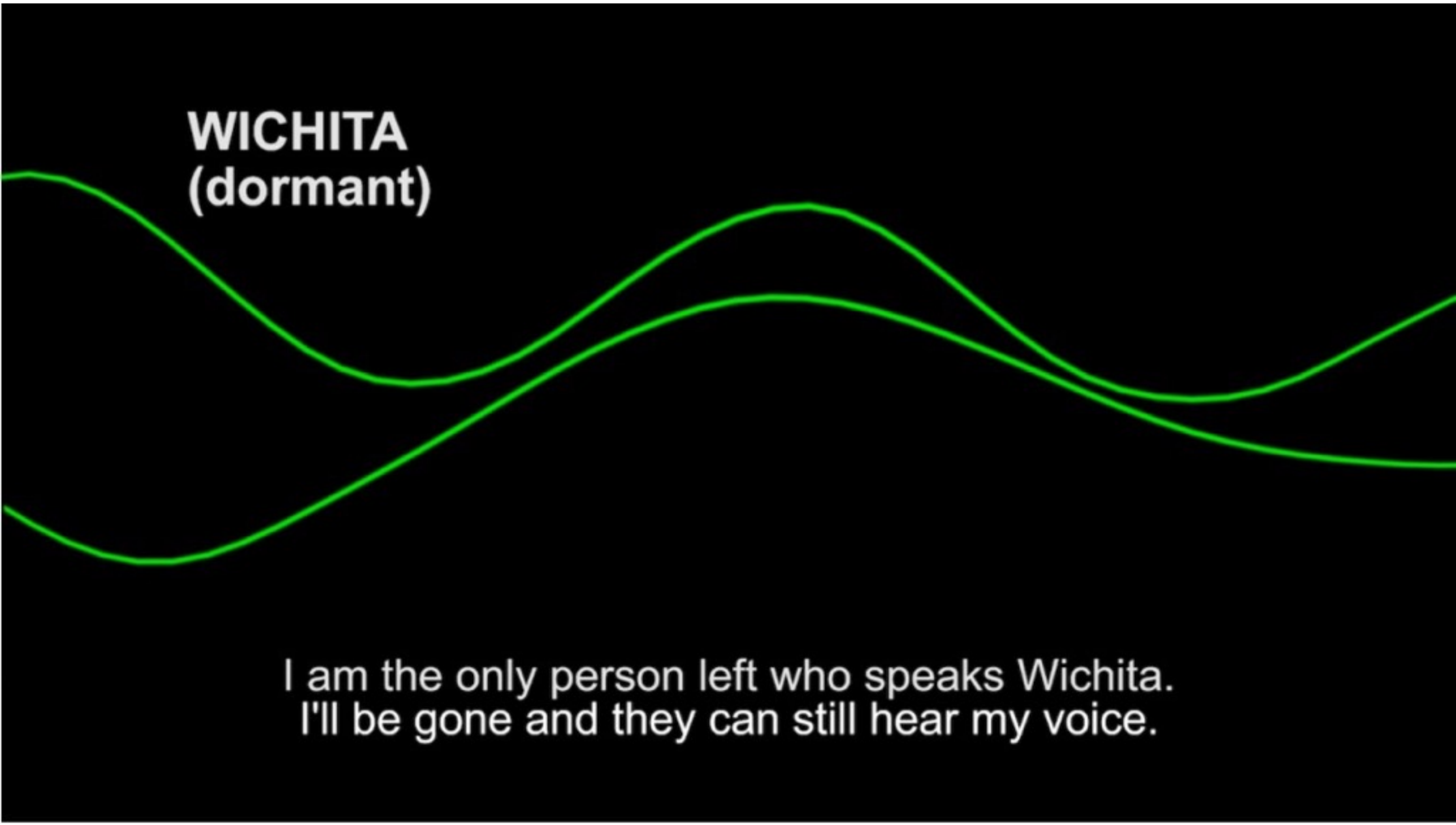


Figure 1: Interactive translation prediction in CASMACAT: The system suggests to continue the translation with the words *mehr als 18*, which the user can accept by pressing the TAB key.

Neural interactive MT

Input: *Das Unternehmen sagte, dass es in diesem Monat mit Bewerbungsgesprächen beginnen wird und die Mitarbeiterzahl von Oktober bis Dezember steigt.*

	Correct	Prediction	Prediction probability distribution
✓	the	the	**the (99.2)**
✓	company	company	**company (90.9)**, firm (7.6)
✓	said	said	**said (98.9)**
✓	it	it	**it (42.6)**, this (14.0), that (13.1), job (2.0), the (1.7), ...
✓	will	will	**will (77.5)**, is (4.5), started (2.5), 's (2.0), starts (1.8), ...
✓	start	start	**start (49.6)**, begin (46.7)
	inter@@	job	job (16.1), application (6.1), en@@ (5.2), out (4.8), ...
✘	viewing	state	state (32.4), related (5.8), **viewing (3.4)**, min@@ (2.0), ...
✘	applicants	talks	talks (61.6), interviews (6.4), discussions (6.2), ...
✓	this	this	**this (88.1)**, so (1.9), later (1.8), that (1.1)
✓	month	month	**month (99.4)**
✘	,	and	and (90.8), **, (7.7)**
✘	with	and	and (42.6), increasing (24.5), rising (6.3), **with (5.1)**, ...
✓	staff	staff	**staff (22.8)**, the (19.5), employees (6.3), employee (5.0), ...
✘	levels	numbers	numbers (69.0), **levels (3.3)**, increasing (3.2), ...
✘	rising	increasing	increasing (40.1), **rising (35.3)**, climbing (4.4), rise (3.4), ...
✓	from	from	**from (97.4)**
✓	October	October	**October (81.3)**, Oc@@ (12.8), oc@@ (2.9), Oct (1.2)
✘	through	to	to (73.2), **through (15.6)**, until (8.7)
✓	December	December	**December (85.6)**, Dec (8.0), to (5.1)
✓	.	.	**. (97.5)**



Susan Hiller, “Lost and Found,” video still (image courtesy the artist and Lisson Gallery, London)

Language documentation

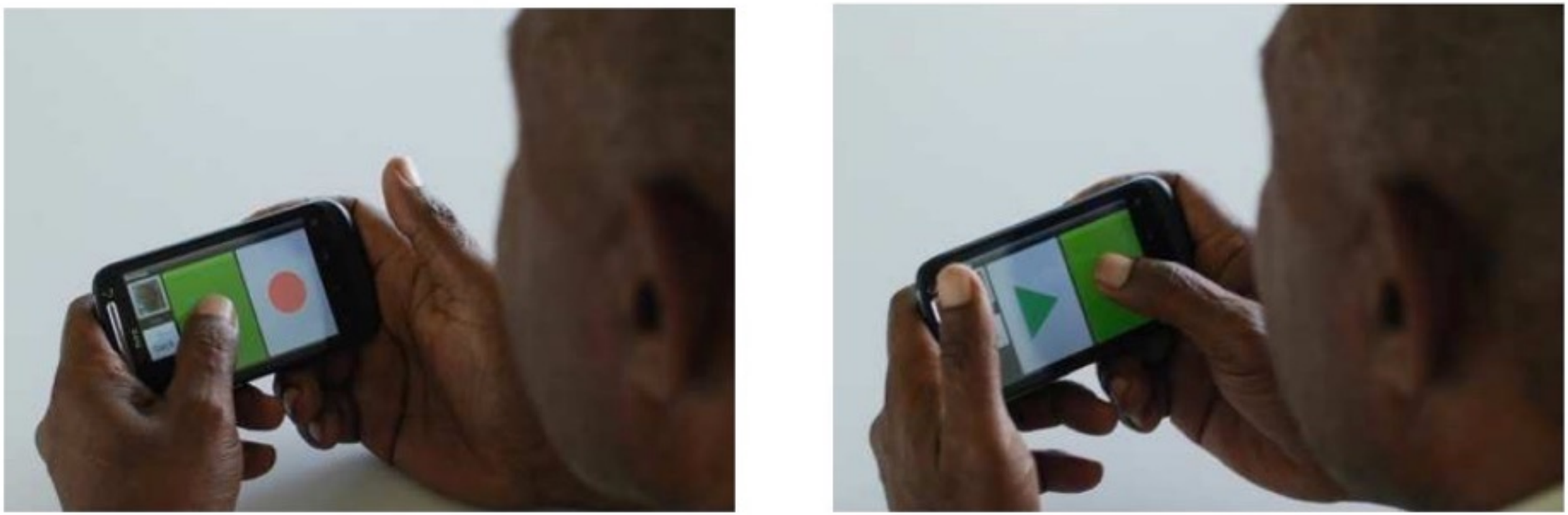
- Recording communicative events
- Transcribing and translating recordings
- Basic morphosyntactic analysis
- Eliciting paradigms
- Describing structure

Language documentation

eng	aso	bef	gah	ino	kbq	snp	yby	zuh
sun	ho	yege	ho	yake	zge	fo	homa	ho
water	noso	nagami	nagami	tina	tina	no	noma	nosa
fire	olo	logo	lo	ata	teve	soo	iizo	olo
earth	misumbo	mei	mikasi	mopa	mo'pa	mika	mika	mikesupa
tree	ya	yafa	za	yosa	zafa	yaa	yah	yah
mountain	golo	kosa	agoka	akoya	agona	obura	bora	gola
house	numuno	nohi	numuni	nona	nona	numuna	numuda	numuna
food	nosonite	nosena	nosa'neta	neya	ne'zane	aáwa'a	nodenesa	nosaneta
pig	ije	yaga	iza	afu	afu	savu	izah	iza
man	we	bo	ve	ve	ve'nene	wee	we	vemoha
woman	vene	amo	vena	a'ne	a're	wena	mena	vena
father	meneho'we	afonifu	ahono	afo nimo'e	nenfa	wemeteuo	ahone	meneho
mother	ijeneho	itonifu	izo'no	ita	anta'nimo	wena otevo	idone	izeneho

Figure 1: Comparative wordlist for the languages spoken near Goroka. Languages are identified by ISO 639-3 code. It is likely that, for some language pairs (e.g. aso-zuh, ino-kbq), many wordforms are related to one another by regular sound correspondences.

Language documentation

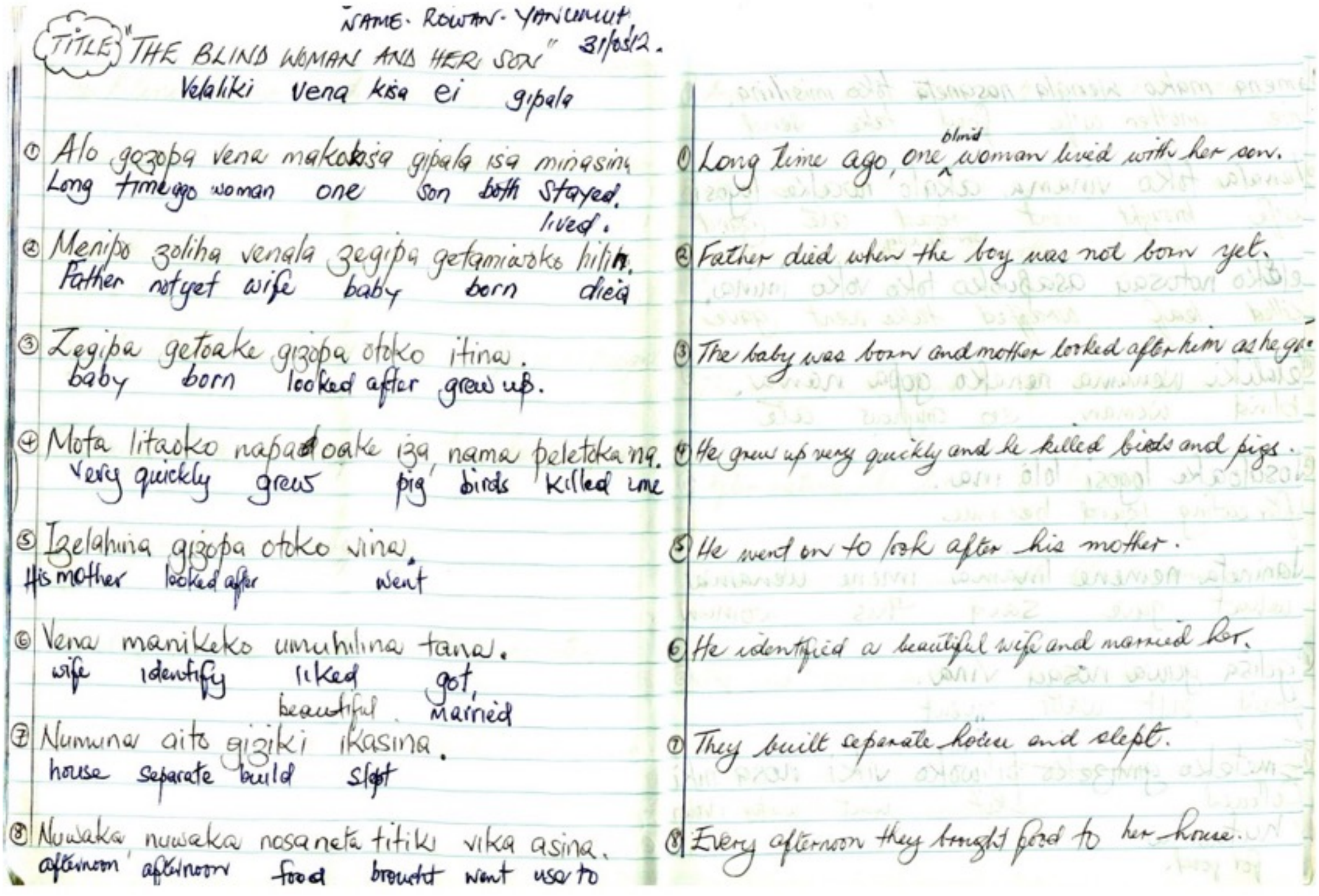


(a) Audio playback

(b) Respeaking and Interpretating

Figure 2: Mobile phone interface: (a) press and hold the play button to hear the original recording (b) press and hold the record button to record the respeaking or interpreting

Language documentation



NAME. ROWAN. YANCUMUT

TITLE "THE BLIND WOMAN AND HER SON" 31/05/12.

Velaliki vena kisa ei gipala

① Alo gezopa vena makakaisa gipala isa minasina
Long time ago woman one son both stayed. lived.

② Menipo zoliha venala zegipa getamiaska hilita.
Father not yet wife baby born died

③ Zegipa getoake gizopa otoko itina.
baby born looked after grew up.

④ Mota litaoko napaatoake iza, nama peletaka ng.
very quickly grew pig birds killed one

⑤ Izelahina gizopa otoko vina.
His mother looked after went

⑥ Vena manikeko umuhilina tana.
wife identify liked got, beautiful married

⑦ Numina aito giziki ikasina.
house separate build slept

⑧ Nuwaka nuwaka nosaneta titiki vika asina.
afternoon afternoon food brought went use to

① Long time ago, one blind woman lived with her son.

② Father died when the boy was not born yet.

③ The baby was born and mother looked after him as he gr

④ He grew up very quickly and he killed birds and pigs.

⑤ He went on to look after his mother.

⑥ He identified a beautiful wife and married her.

⑦ They built separate house and slept.

⑧ Every afternoon they brought food to her house.

Language documentation

(1) Velaliki veena kisa ei gipala (The blind woman and her son)

Alo gozopa vena makokisa gipala isa minasina.
long time ago woman one son both lived.
A long time ago, a blind woman lived with her son.

Menipo zoliha venala zegipa getamiwoko hilina.
father not yet wife baby born died.
The father died when the boy was not yet born.

Zegipa getoake gizopa otoko itina.
baby born looked after grew up.
The baby was born and the mother looked after him as he grew.

Mota litaoko napaoake iza, nama peletoka ana.
very quickly grew pig, birds killed came.
He grew up very quickly and he killed birds and pigs.

Izelahina gizopa otoko vina.
his mother looked after went.
He went on to look after his mother.

Week 10 plan

- Monday 14:30: final office hours
- Monday 16:00: decipherment lecture
- Tuesday 10:00: lab (coursework support)
- Wednesday 10:00: lab (coursework support)
- Thursday 16:00: history lecture
- After that: piazza