Overview

- Class description
- Introduction to MT
  - History
  - Major challenges
  - Types of systems
  - What the seminar will cover
- Corpora for Statistical MT
  - Parallel corpora
  - Comparable corpora
- Evaluation of MT Systems
  - Human Evaluation
  - Automatic Evaluation
Class description
Class Logistics

- Instructors: Chris Quirk and Kristina Toutanova
- Time: Mondays 4:30pm to 6:50pm LOW 202
- Office hours: ??
- Course website:
  - https://catalyst.uw.edu/workspace/kristout/20547/
  - Syllabus (with slides and pointers to readings): updated every week
  - Message board
  - CollectIt
Prerequisites

- Background in probability, statistics, and machine learning
  - LING 572 or similar
- NLP
  - Shallow processing LING 570
  - Syntax and parsing LING 571
- Ability to use commands on Unix, and program in a high-level programming language
Questions to students

- Why are you interested in machine translation?
- Do you speak a non-English language?
- Anything in particular you would like to get out from the seminar?

- Could you send us an email with your answers after class
Objectives

- To teach the foundations of state-of-the-art statistical MT systems so that you are able to
  - Understand current research papers on MT
  - Train and use an MT system
  - Have ideas of how you might improve the quality for a particular language-pair
- To give you the opportunity to experiment with MT systems and conduct a small research study
Assignments

- Four homework sets [60%] including
  - written problems
  - work with MT system software on Unix
  - implementation of simple components in a high-level programming language
  - Two short (10% each) and two longer (20% each)

- Term project [40%] (alone or group of 2 or 3 students)
  - Formulate a small-sized experiment (or a theoretical study), carry out the study, and present the results in a term paper and a final presentation
Schedule of Assignments

- Hw1 out 3/28 due 4/06 (short homework)
- Hw2 out 4/04 due 4/18
- Hw3 out 4/18 due 4/25 (short homework)
- Hw 4 out 4/25 due 5/16

- Final projects
  - Proposals due 5/02
  - Updates due 5/18 (not graded)
  - Project reports and presentations due last day of class
  - Talk to us or send us email with ideas/questions beforehand if you like
  - If you are ready to start on the project earlier you can propose earlier

- Preferences for the day of final project presentations?
  - 5/30 is a holiday
Other policies

- Late assignment policy
  - You have a total of 3 free late days which can be used for all assignments; after the late days are used up the penalty is a factor of 0.9 applied to your score for each 24 hours

- Recommendation
  - Complete the required reading before class
  - Ask questions or bring up discussion items during class
Assignments

- Four homework sets

- Goals of homework
  - Encourage thorough understanding of material
  - Provide opportunity to become familiar with MT system components, using existing packages (located at /NLP_TOOLS/mt_tools on Patas)
    - Moses phrase-based SMT system
    - GIZA++ word-alignment models
    - SRILM language model
    - Syntax-based SMT
Ideas for Final Project Topics

- Apply models to new languages and study the results
- Formulate a new word-alignment model or use new features in an existing model
  - Could make your own small set of annotated parallel sentence pairs for evaluation
- Implement a modification to the phrase-table of a phrasal MT system
- Implement a method for mining parallel documents or sentences from comparable corpora or from the web
- Implement a method for pre or post-processing (e.g. using syntax or morphology)
- Come up with a new method for phrase pair extraction from parallel corpora
- Test a new word-segmentation strategy for MT (e.g. for Chinese or morph. rich languages)
Alternative Final Project: A Survey

- If you would rather spend more time researching the relevant literature
- Could instead collect and write an overview of three or more related articles
- Also list ideas for future extensions
Introduction to MT
Translation: global problem and interesting research problem

- Non-English Internet content and user communities are increasing explosively

- Human translation costs are excessive: major languages range from 10-50 cents per word

Result: the vast majority of published material remains untranslated!
A brief history of MT
(Based on work by John Hutchins)

- Before the computer: In the mid 1930s, a French-Armenian Georges Artsrouni and a Russian Petr Troyanskii applied for patents for ‘translating machines’.

  When I look at an article in Russian, I say to myself: This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode.

  Warren Weaver (1947)

- The pioneers (1947-1954): the first public MT demo was given in 1954 (by IBM and Georgetown University).
The decade of optimism (1954-1966):
- Foundations of rule-based approaches using varying levels of representation were established
- Different levels of representation explored
A brief history of MT (cont)

- ALPAC (Automatic Language Processing Advisory Committee) report in 1966: "there is no immediate or predictable prospect of useful machine translation."

- The aftermath of the ALPAC report (1966-1980): a virtual end to MT research

- The 1980s:
  - rule-based Interlingua models, data-driven example-based MT
  - first commercial MT systems

- The 1990s: research on statistical MT started but rule-based still more prominent
  - started by researchers at IBM, mostly word-based

- The 2000s: Statistical MT gathers full steam and rule-based MT systems are also improved by statistical methods
Where are we now?

- Huge potential/need due to the internet, globalization and international politics.
- Quick development time due to SMT, the availability of parallel data and computers.
- Translation is reasonable for language pairs with a large amount of resource.
- Starting to include more “minor” languages.
- Got stuck?
Things are Consistently Improving

Annual evaluation of Arabic-to-English MT systems

Exceeded commercial-grade translation here.
Bilingual Training Data

(Data stripped of formatting, in sentence-pair format, available from the Linguistic Data Consortium at UPenn).
Sample Learning Curves

Experiments by Philipp Koehn
What is MT good for?

- Rough translation (gisting): web data
  - Current MT system performance can be adequate for this task
- Computer-aided human translation
  - MT output saves time for human translators
- Fully automatic high-quality machine translation
  - Currently possible for limited domains only or when the source documents are written in controlled language
- Cross-lingual IR
  - Search documents in multiple languages
- Communication via email, chat
  - May be possible to communicate but could lead to serious problems
- Translation on hand-held devices
  - Already useful for limited tourist domain applications
Major challenges
Translation is hard even for humans

- Novels
- Word play, jokes, puns, hidden messages
- Concept gaps:
  - E.g when I speak to a Bulgarian colleague about research I often mix in English words due to lack of Bulgarian translation for some technical terms
- Other constraints: lyrics, dubbing, poem, ...
Major challenges

- Getting the right words:
  - Choosing the correct root form
  - Getting the correct inflected form
  - Inserting “spontaneous” words

- Putting the words in the correct order:
  - Word order: SVO vs. SOV, ...
  - Unique constructions
  - Divergence
Lexical choice

- Homonymy/Polysemy: bank, run

- Coding (Concept $\rightarrow$ lexeme mapping) differences:
  - More distinction in one language: e.g., kinship vocabulary.
  - Different division of conceptual space
Choosing the appropriate inflection

- Inflection: gender, number, case, tense, ...

- Ex:
  - Number: Ch-Eng: all the concrete nouns:
    ch_book ➔ book, books
  - Gender: Eng-Fr: all the adjectives
  - Case: Eng-Korean: all the arguments
  - Tense: Ch-Eng: all the verbs:
    ch_buy ➔ buy, bought, will buy
Inserting spontaneous words

- **Function words:**
  - **Determiners:** Ch-Eng:
    - ch_book ➔ a book, the book, the books, books
  - **Prepositions:** Ch-Eng:
    - ... ch_November ➔ ... in November
  - **Relative pronouns:** Ch-Eng:
    - ... ch_buy ch_book de ch_person ➔ the person who bought /book/
  - **Possessive pronouns:** Ch-Eng:
    - ch_he ch_raise ch_hand ➔ He raised his hand(s)
  - **Conjunction:** Eng-Ch:
    - Although S1, S2 ➔ ch_although S1, ch_but S2
- ...
Word order

- SVO, SOV, VSO, ...
- VP + PP $\rightarrow$ PP VP
- VP + AdvP $\rightarrow$ AdvP + VP
- Adj + N $\rightarrow$ N + Adj
- NP + PP $\rightarrow$ PP NP
- NP + S $\rightarrow$ S NP
- P + NP $\rightarrow$ NP + P
Translation divergences

- Source and target parse trees (dependency trees) are not identical.

- Example: I like Mary ➔ S: Marta me gusta a mi ('Mary pleases me')
MT quality varies greatly depending on language-pair divergence

French input
Nous savons très bien que les Traités actuels ne suffisent pas et qu’il sera nécessaire à l’avenir de développer une structure plus efficace et différente pour l’Union, une structure plus constitutionnelle qui indique clairement quelles sont les compétences des États membres et quelles sont les compétences de l’Union.

Statistical machine translation
We know very well that the current treaties are not enough and that in the future it will be necessary to develop a different and more effective structure for the union, a constitutional structure which clearly indicates what are the responsibilities of the member states and what are the competences of the union.

Human translation
We know all too well that the present Treaties are inadequate and that the Union will need a better and different structure in future, a more constitutional structure which clearly distinguishes the powers of the Member States and those of the Union.

“Easy” language pair: French-English translation with 40 million words training data. [SMT Ch1 Fig 1.8]
MT quality varies greatly depending on language-pair divergence

Chinese input
伦敦每日快报指出,两台记载黛安娜王妃一九九七年巴黎死亡车祸调查资料的手提电脑,被从前大都会警察总长的办公室里偷走。

Statistical machine translation
The London Daily Express pointed out that the death of Princess Diana in 1997 Paris car accident investigation information portable computers, the former city police chief in the offices of stolen.

Human translation
London’s Daily Express noted that two laptops with inquiry data on the 1997 Paris car accident that caused the death of Princess Diana were stolen from the office of a former metropolitan police commissioner.

Hard language pair: Chinese-English translation with 200 million words training data. [SMT Ch1]
Major approaches
Types of MT systems

- **Source of information**
  - Rule based: People write rules to specify translations of words, phrases
  - Data-driven: Use learning techniques to derive translation “rules” from data sources (e.g., parallel corpora)

- **Level of representation**
  - Interlingua
  - Semantic forms
  - Syntax trees
  - Phrases
  - Words
Brief description of rule-based systems: transfer-based systems

- Most successful rule-based systems are transfer-based

- Analysis, transfer, generation:
  1. Parse the source sentence
  2. Transform the parse tree with transfer rules
  3. Translate source words
  4. Get the target sentence from the tree

- Resources required:
  - Source parser
  - A translation lexicon
  - A set of transfer rules
Transfer-based systems (continued)

- Parsing: linguistically motivated grammar or formal grammar?
- Transfer:
  - context-free rules? A path on a dependency tree?
  - Apply at most one rule at each level?
  - How are rules created?
- Translating words: word-to-word translation?
- Generation: using LM or other additional knowledge?
- How to create the needed resources automatically?
First data-driven systems: example-based MT

- Basic idea: translate a sentence by using the closest match in parallel data.
  - First proposed by Nagao (1981).
  - Question: how to define sentence similarity and how to combine translation parts?

- Types of EBMT:
  - Lexical (shallow)
  - Morphological / POS analysis
  - Parse-tree based (deep)

- Types of data required by EBMT systems:
  - Parallel text
  - Bilingual dictionary
  - Thesaurus for computing semantic similarity
  - Syntactic parser, dependency parser, etc.
Advantages of data-driven translation

- We can model the genres of documents that we would like to model
  - Learn contextually appropriate translations for technical data, chat data, etc.

- Very flexible system
  - Given corpus $\mathbf{C} = \{(x_1, y_1), (x_2, y_2), \ldots\}$ of sentence pairs
  - $\text{Translate}(\mathbf{C}, x) = y$ is a function of the training data and the input sentence
  - To build a new system (or optimize our old one) we just change the data
Statistical MT: word-based

- Translate using a per-word process

Mary did not slap the green witch

Mary not slap slap the green witch

Mary not slap slap NULL the green witch

Maria no daba una bofetada a la bruja verde

Maria no daba una bofetada a la bruja verde

[One of Kevin Knight’s common examples]
Translate more than one word at a time

“Phrase” is not a linguistic constituent, just a contiguous sequence of words

morgen fliege ich nach Kanada zur Konferenz

 tomorrow I will fly to the conference in Canada
Use syntactic analyses to guide or influence the translation process

Phrase-structure grammar or dependency analyses can be used for the source and/or target languages
Beyond syntax?

- Semantics:
  - Gives finer insights into the meaning of a sentence
  - However, few (freely available) semantic parsers
  - As analysis deepens, accuracy tends to suffer

- Interlingua:
  - Enticing from a theoretical standpoint
  - Can bypass the $n^2$ problem with multilingual corpora
  - Difficult to define string <-> interlingua mappings
  - Languages often underspecified vs. one another

- Lots of fun open problems!
What the course will cover

- MT system evaluation
- Word-level SMT models
  - Word alignment models
- Phrase-based SMT models
  - Models & Decoding
- Tree-based SMT models
  - Models using hierarchical phrases
  - Syntax tree-based models
- Other techniques (still subject to change)
  - Factored translation models
  - Mining parallel data
  - Discriminative training
  - Feature-rich models
What techniques will be used

- Complex structured statistical models of translation correspondence
- Supervised and unsupervised statistical models
- Known search algorithms from AI and new ones proposed for MT
- Probabilistic grammars
- Monolingual language analysis techniques
  - Word segmentation and morphological analysis
  - Part-of-speech tagging
  - Syntactic parsing
  - Semantic parsing?
  - Discourse?
Corpora for MT
Parallel Corpora

- Document pairs which are direct translations
  - LDC releases large corpora for Chinese-English, Arabic-English, and French-English (Canadian Hansard)
  - Europarl contains multi-parallel corpora in 21 languages (parliamentary proceedings)
  - JRC-Acquis (Acquis Communicare) contains multi-parallel corpora in 22 languages (EU legal documents)
  - Parallel documents are available on the web (multi-lingual websites)
    - Can be mined automatically but can be tricky [e.g. Resnik 99]
  - If you need a corpus for a specific language pair, we may be able to find one elsewhere (let us know).
Comparable corpora

- Documents in two languages that are about a similar event, topic, or concept, but are not direct translations
  - News stories in multiple languages, covering the same events
  - Documents that contain some translated material but also include divergent content (e.g. Wikipedia articles in different languages connected with links)
Sentence alignment in parallel corpora

- Need sentence boundary detection
  - Good but not perfect tools exist (e.g. 99% accuracy)
- Sentence alignment: many models proposed
  - Gale and Church 1993 propose a simple fast algorithm
- Assumes sentences are in monotone correspondence, either 1-to-1, 0-to-1, 1-to-2, or 2-to-2
- Assigns scores to each possible alignment of documents which decomposes into matching scores for each individual sentence group alignment
- The scores depend on the type of correspondence (1-to-1 most likely) and the similarity of lengths of aligned sentences
- Globally best alignment can be found using a dynamic program
- Details on p.56 in the textbook
Using comparable corpora

- Harder to find corresponding sentences
- Munteanu and Marcu [05] use news data and build a feature-based pairwise model for parallel sentence extraction
- Smith et al. [10] use Wikipedia and define a CRF sentence alignment model, using document structure and additional features
- Munteanu and Marcu [06], Quirk et al. [07] also extract and use parallel sentence fragments
- A large body of work on extracting translation lexicons from comparable corpora
- We will cover some of this work in more detail later
Evaluation of MT Systems
Evaluating MT Systems

- Evaluation is very hard
  - No single correct translation
  - Distinctions between correct and incorrect translations are vague

Ten translations for a Chinese sentence. An example from the NIST 2001 evaluation set. [borrowed from SMT Ch 8]
How to evaluate

- Manual subjective evaluation by humans
- Automatic evaluation using measures with respect to one or more reference translations
- Task-Oriented Evaluation
  - Cost of post-editing
  - Content understanding tests
Manual evaluation of MT output

- Best case: bilingual speakers evaluate translation given source sentence
- Also works: target language speakers compare system translation to reference translation
- Goal: assign a score to a system output (usually 1 to 5)
- Helpful to evaluate two dimensions separately
  - Fluency
    - Is the output good fluent English? Is it grammatically correct, is word usage idiomatic?
  - Adequacy
    - Does the output convey the same meaning as the source? Are important pieces lost or distorted? Is extra information added?
Manual evaluation of MT output

<table>
<thead>
<tr>
<th>Adequacy</th>
<th>Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>flawless English</td>
</tr>
<tr>
<td>4</td>
<td>good English</td>
</tr>
<tr>
<td>3</td>
<td>non-native English</td>
</tr>
<tr>
<td>2</td>
<td>disfluent English</td>
</tr>
<tr>
<td>1</td>
<td>none</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The cat sat on the mat</th>
</tr>
</thead>
<tbody>
<tr>
<td>The cat sat on the floor</td>
</tr>
<tr>
<td>On the mat sat the cat</td>
</tr>
<tr>
<td>The the cat sat on the straw mat</td>
</tr>
<tr>
<td>The dog sat on the mat</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Adequacy</td>
</tr>
<tr>
<td>?</td>
</tr>
<tr>
<td>?</td>
</tr>
<tr>
<td>?</td>
</tr>
<tr>
<td>?</td>
</tr>
</tbody>
</table>
# Manual evaluation of MT output

<table>
<thead>
<tr>
<th>Adequacy</th>
<th>Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>all meaning</td>
</tr>
<tr>
<td>4</td>
<td>most meaning</td>
</tr>
<tr>
<td>3</td>
<td>much meaning</td>
</tr>
<tr>
<td>2</td>
<td>little meaning</td>
</tr>
<tr>
<td>1</td>
<td>none</td>
</tr>
<tr>
<td>5</td>
<td>flawless English</td>
</tr>
<tr>
<td>4</td>
<td>good English</td>
</tr>
<tr>
<td>3</td>
<td>non-native English</td>
</tr>
<tr>
<td>2</td>
<td>disfluent English</td>
</tr>
<tr>
<td>1</td>
<td>none</td>
</tr>
</tbody>
</table>

**The cat sat on the mat**  
reference  
Adequacy  
Fluency

- The cat sat on the floor  
  MT 1  
  4 or 3  
  5

- On the mat sat the cat  
  MT 2  
  5 or 4  
  3 or 4

- The the cat sat on the straw mat  
  MT 3  
  4 or 3  
  2 or 3

- The dog sat on the mat  
  MT 4  
  2 or 3  
  5

*Distinctions are vague and inter-annotator agreement is low. Comparing pairs of translations (which of two translations is better) is more reliable.*
Automatic Evaluation of MT output

- Much faster and cheaper than human evaluation
- Goal
  - Compare system translation to one or more reference translations
  - Given a test set of sentences, tell whether system A is better than system B
- After we have come up with a set of reference translations for a test set, we can run automatic evaluation with different system configurations at no cost
- Problem: how to compare system output to reference translations
Automatic Evaluation of MT output

- Compute similarity between system and reference translations
- Current widely used measures look at
  - Whether the system has the same words as the reference
  - Whether the words are in the same order
- Active research on more sophisticated measures
  - Taking into account synonyms
  - Modeling ordering better
  - Taking into account the syntactic structure of sentences and the varying importance of words
N-gram precision

- N-gram precision: the percent of n-grams in the system output that are correct.

- Clipping:
  - Sys: the the the the the the the the
  - Ref: the cat sat on the mat
  - Unigram precision: 2/6
  - Max_ref_count: the max number of times a ngram occurs in any single reference translation.

\[
\text{count\_clip} = \min(\text{count}, \text{max\_ref\_count})
\]
N-gram precision with multiple references

\[ P_n = \frac{\sum_{S \in Sys} \sum_{ngram \in S} \text{Count}_{clip}(ngram)}{\sum_{S \in Sys} \sum_{ngram \in S} \text{Count}(ngram)} \]

i.e. the percent of n-grams in the system output that are correct (after clipping).
Precision, recall, F-measure

- N-gram precision, recall, and f-measure
  - higher order (2+) n-grams measure word order as well
  - up to 4-gram usually sufficient

<table>
<thead>
<tr>
<th></th>
<th>reference</th>
<th>1-prec</th>
<th>1-rec</th>
<th>1-F1</th>
<th>2-prec</th>
<th>2-rec</th>
<th>2-F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>the cat sat on the mat</td>
<td>reference</td>
<td>5/6</td>
<td>5/6</td>
<td>.833</td>
<td>4/5</td>
<td>4/5</td>
<td>.8</td>
</tr>
<tr>
<td>on the mat sat the cat</td>
<td>MT 1</td>
<td>5/6</td>
<td>5/6</td>
<td>.833</td>
<td>4/5</td>
<td>4/5</td>
<td>.8</td>
</tr>
<tr>
<td>the the cat sat on the straw mat</td>
<td>MT 2</td>
<td>6/6</td>
<td>6/6</td>
<td>1</td>
<td>3/5</td>
<td>3/5</td>
<td>.6</td>
</tr>
<tr>
<td>the cat sat on the floor</td>
<td>MT 3</td>
<td>6/8</td>
<td>6/6</td>
<td>.857</td>
<td>4/7</td>
<td>4/5</td>
<td>.667</td>
</tr>
</tbody>
</table>

Reference 1-grams: the, cat, sat, on, the, mat
Reference 2-grams: the cat, cat sat, sat on, on the, the mat

- F-measure is a combination of precision and recall:
  $2 \times \text{precision} \times \text{recall} / (\text{precision} + \text{recall})$
BLEU

- Proposed by Papineni et. al. (2002)
- Most widely used in MT community.
- BLEU is a weighted average of n-gram precision \( p_n \) between system output and all references, multiplied by a brevity penalty (BP).

\[
BLEU = BP \times \prod_{n=1}^{N} p_n^{w_n}
\]

\[
= BP \times \sqrt[1-N]{p_1 \times p_2 \times \ldots \times p_N} \quad (\text{when} \quad w_n = \frac{1}{N})
\]
Brevity Penalty

- For each sent \( s_i \) in system output, find closest matching reference \( r_i \) (in terms of length).

\[
Let \quad c = \sum_{i} |s_i|, \quad r = \sum_{i} |r_i|
\]

\[
BP = \begin{cases} 
1 & \text{if } c > r \\
\exp^{-r/c} & \text{otherwise}
\end{cases}
\]

- Longer system output is already penalized by the n-gram precision measure.
- This version of the BP does not match the one in the book.
Word Error Rate (WER)

- Word Error Rate (WER) was traditionally used for MT research most frequently before the BLEU metric was proposed.
- It is the edit distance between the translation and reference sentences, divided by the length of the reference.
- The Edit distance (a.k.a. Levenshtein distance) is defined as the minimal cost of transforming \( \text{str1} \) into \( \text{str2} \), using three operations (substitution, insertion, deletion).
- Can compute it in time \( O(s^r) \) using a dynamic program.
Automatic evaluation: Is the metric good?

- We want an automatic metric to correlate with manual evaluation scores.
- Graph shows BLEU correlates well with manual evaluation of fluency 88.0% and adequacy 90.2%.
- However BLEU is not perfect:
  - Rule-based systems scored lower.
  - Human BLEU scores can be same as system.

[from G. Doddington, NIST]
Automatic Evaluation

- Other metrics
  - METEOR
    - Emphasizes recall more, uses knowledge about word morphology and synonyms
  - Automatically learned combinations of other metrics
  - Metrics using syntactic information and other features

- Active area of research
  - Can have a term project on creating a new evaluation measure
Assignments for next time

- Skim Chapters 2 and 3 or read carefully if you lack some prerequisite
- Homework 1 due 4/6 (posted later tonight)
- Can read SMT Ch 4.1 to 4.4.2 and optional readings from website to be prepared to understand lecture better
- Send us responses to questions from Slide 6 (questions to students)
- Think about how you would solve the puzzle on the next page
Think how to translate this to Arcturan: farok crrrok hihok yorok clok kantok ok-yurp

---

1a. ok-voon ororok sprok .

1b. at-voon bichat dat .

---

2a. ok-drubel ok-voon anok ploksprok .

2b. at-drubel at-voon pippat rrat dat .

---

3a. erok sprok izok hihok ghirok .

3b. totat dat arrat vat hilat .

---

4a. ok-voon anok drok brok jok .

4b. at-voon krat pippat sat lat .

---

5a. wiwok farok izok stok .

5b. totat jjat quat cat .

---

6a. lalok sprok izok jok stok .

6b. wat dat krat quat cat .

---

7a. lalok farok ororok lalok sprok izok enemok .

7b. wat jjat bichat vat dat vat eneat .

---

8a. lalok brok anok ploks brok nok .

8b. iat lat pippat rrat nnat .

---

9a. wiwok nok izok kantok ok-yurp .

9b. totat nnat quat oloat at-yurp .

---

10a. lalok mok nok yorok ghirok clok .

10b. wat nnat gat mat bat hilat .

---

11a. lalok nok crrrok hihok yorok zanzanok .

11b. wat nnat arrat mat zanzanat .

---

12a. lalok rarok nok izok hihok mok .

12b. wat nnat forat arrat vat gat .