An Empirical Analysis of Task Relations in the Multi-Task Annotation of an Arabizi Corpus

Elisa Gugliotta Marco Dinarelli

Univ. Grenoble Alpes, CNRS, Grenoble INP*, LIG, 38000 Grenoble, France;

Institute of Engineering Univ. Grenoble Alpes; Groupe *getalp*

elisa.gugliotta@univ-grenoble-alpes.fr
marco.dinarelli@univ-grenoble-alpes.fr

Contents

Contents

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

  → Need of **building** resources

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

  → Need of **building** resources
  → Need of linguistic **annotated corpora** to study

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

→ Need of **building** resources
→ Need of linguistic **annotated corpora** to study
→ Need to process them **(semi-)automatically**

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

  → Need of **building** resources
  → Need of linguistic **annotated corpora** to study
  → Need to process them **(semi-)automatically**

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

  → Need of **building** resources
  → Need of linguistic **annotated corpora** to study
  → Need to process them **(semi-)automatically**

Computational Strategies

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

→ Need of **building** resources
→ Need of linguistic **annotated corpora** to study
→ Need to process them **(semi-)automatically**

Computational Strategies

Machine Learning Tools can speed up semi-automatic annotation of corpora

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

→ Need of **building** resources
→ Need of linguistic **annotated corpora** to study
→ Need to process them **(semi-)automatically**

Computational Strategies

Machine Learning Tools can speed up semi-automatic annotation of corpora
Multi-Task (MT) learning can enhance performance across various tasks to be learned, as compared to the individual tasks learned separately (Caruana, 1997).

Study Aims

Analysis of under-resourced languages

- Lack of **comprehensive**, **in-depth** and **up-to-date** descriptions

→ Need of **building** resources
→ Need of linguistic **annotated corpora** to study
→ Need to process them **(semi-)automatically**

Computational Strategies

Machine Learning Tools can speed up semi-automatic annotation of corpora
Multi-Task (MT) learning can enhance performance across various tasks to be learned, as compared to the individual tasks learned separately (Caruana, 1997).
Uncover Task Relations to improve annotation strategies and contribute to developing linguistic resources for under-resourced languages.

Contents

Tunisian Arabish Corpus (Gugliotta and Dinarelli, 2022)

Arabizi	**Class.**	**CODA***	**Token.**	**POS**	**Lemma**
Inchalah	Az.	ان شاء الله	ان شاء الله	INTERJ	ان شاء الله
cycle	Fr.	Fr.	Fr.	Fr.	Fr.
ejjay	Az.	الجاي	ال+جاي	DET+ADJ	جاي
wala	Az.	ولّا	ولّا	CONJ	ولّا
eli	Az.	اللي	اللي	REL_PRON	اللي
ba3dou	Az.	بعده	بعد+ه	ADV+ PRON_3MS	بعد

Table 1: Example of the annotation levels. "Az." means "Arabizi", "Fr." means "foreign". CODA* by Habash et al. (2018), POS inspired to Maamouri et al. (2004)

Tunisian Arabish Corpus (Gugliotta and Dinarelli, 2022)

Arabizi	Class.	CODA*	Token.	POS	Lemma
Inchalah	Az.	ان شاء الله	ان شاء الله	INTERJ	ان شاء الله
cycle	Fr.	Fr.	Fr.	Fr.	Fr.
ejjay	Az.	الجاي	ال+جاي	DET+ADJ	جاي
wala	Az.	ولّا	ولّا	CONJ	ولّا
eli	Az.	اللي	اللي	REL_PRON	اللي
ba3dou	Az.	بعده	بعد+ه	ADV+ PRON_3MS	بعد

Table 1: Example of the annotation levels. "Az." means "Arabizi", "Fr." means "foreign". CODA* by Habash et al. (2018), POS inspired to Maamouri et al. (2004)

Non-standardised spelling, code-mixing, script-mixing, etc...

Tunisian Arabish Corpus (Gugliotta and Dinarelli, 2022)

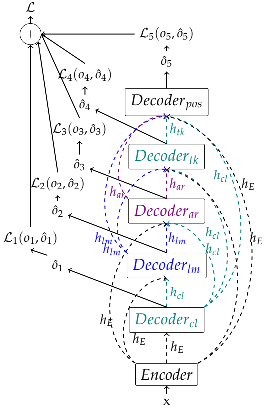
Arabizi	**Class.**	**CODA***	**Token.**	**POS**	**Lemma**
Inchalah	Az.	ان شاء الله	ان شاء الله	INTERJ	ان شاء الله
cycle	Fr.	Fr.	Fr.	Fr.	Fr.
ejjay	Az.	الجاي	الـ+جاي	DET+ADJ	جاي
wala	Az.	ولّا	ولّا	CONJ	ولّا
eli	Az.	اللي	اللي	REL_PRON	اللي
ba3dou	Az.	بعده	بعد+ه	ADV+ PRON_3MS	بعد

Table 1: Example of the annotation levels. "Az." means "Arabizi", "Fr." means "foreign". CODA* by Habash et al. (2018), POS inspired to Maamouri et al. (2004)

Non-standardised spelling, code-mixing, script-mixing, etc...
Impact on the performance of MT systems.

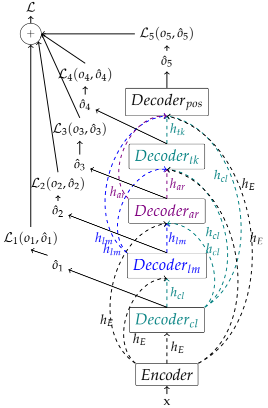
The Neural Multi-Task Architecture (Gugliotta et al., 2020)

- The Encoder to convert x into context-aware repr.;



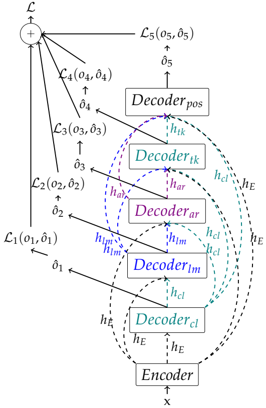
The Neural Multi-Task Architecture (Gugliotta et al., 2020)

- The Encoder to convert x into context-aware repr.;
- 5 Decoders (one for each annotation level);



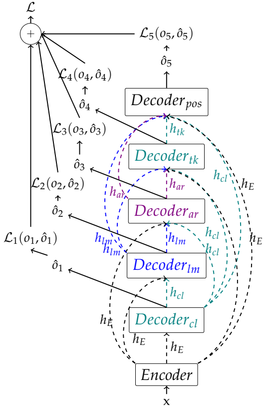
The Neural Multi-Task Architecture (Gugliotta et al., 2020)

- The Encoder to convert x into context-aware repr.;
- 5 Decoders (one for each annotation level);
- Attention mechanism pass previous hidden states ($h_{i-j_{<i}}$) to the modules;



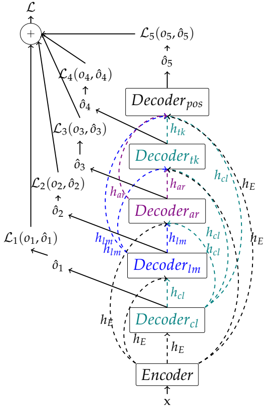
The Neural Multi-Task Architecture (Gugliotta et al., 2020)

- The Encoder to convert x into context-aware repr.;
- 5 Decoders (one for each annotation level);
- Attention mechanism pass previous hidden states ($h_{i-j_{<i}}$) to the modules;
- Single task loss computation (i.e. $\mathcal{L}_i(o_i, \hat{o}_i)$);



The Neural Multi-Task Architecture (Gugliotta et al., 2020)

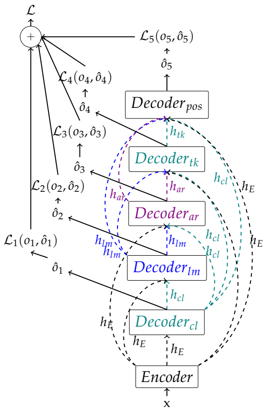
- The Encoder to convert x into context-aware repr.;
- 5 Decoders (one for each annotation level);
- Attention mechanism pass previous hidden states ($h_{i-j_{<i}}$) to the modules;
- Single task loss computation (i.e. $\mathcal{L}_i(o_i, \hat{o}_i)$);
- End-to-end learning of the whole architecture: $\mathcal{L} = \sum_{i=1}^{5} \mathcal{L}_i(o_i, \hat{o}_i)$.



The Neural Multi-Task Architecture (Gugliotta et al., 2020)

- The Encoder to convert x into context-aware repr.;
- 5 Decoders (one for each annotation level);
- Attention mechanism pass previous hidden states ($h_{i-j_{<i}}$) to the modules;
- Single task loss computation (i.e. $\mathcal{L}_i(o_i, \hat{o}_i)$);
- End-to-end learning of the whole architecture: $\mathcal{L} = \sum_{i=1}^{5} \mathcal{L}_i(o_i, \hat{o}_i)$.

→ Are auxiliary tasks beneficial or do they produce negative transfer? (Changpinyo et al., 2018; Ruder, 2017)



Contents

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;
- Comparison between results in terms of Accuracy;

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;
- Comparison between results in terms of Accuracy;
- Incremental procedure by adding further levels one by one, based on the previous best results.

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;
- Comparison between results in terms of Accuracy;
- Incremental procedure by adding further levels one by one, based on the previous best results.

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;
- Comparison between results in terms of Accuracy;
- Incremental procedure by adding further levels one by one, based on the previous best results.

Multi-Task Experiments

- Iterative procedure by testing all possible combinations of all levels;

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;
- Comparison between results in terms of Accuracy;
- Incremental procedure by adding further levels one by one, based on the previous best results.

Multi-Task Experiments

- Iterative procedure by testing all possible combinations of all levels;
- Comparison between results in terms of Accuracy;

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;
- Comparison between results in terms of Accuracy;
- Incremental procedure by adding further levels one by one, based on the previous best results.

Multi-Task Experiments

- Iterative procedure by testing all possible combinations of all levels;
- Comparison between results in terms of Accuracy;

Experimental Procedure

Single-Task Experiments

- Iterative procedure by testing all possible combinations of two levels;
- Comparison between results in terms of Accuracy;
- Incremental procedure by adding further levels one by one, based on the previous best results.

Multi-Task Experiments

- Iterative procedure by testing all possible combinations of all levels;
- Comparison between results in terms of Accuracy;

Comparison between ST and MT strategies

(Hybrid) Single-Task Experiments

Tasks	Arabizi input (class.)	CODA* input
Token.	**80.0% (93.0%)**	**95.4%**
POS	73.8% (92.5%)	54.5%
Lemma	75.5% (92.8%)	**89.5%**
Translit.	**79.0%** (92.8%)	67.2%

Table 2: Starting ST Experiments

(Hybrid) Single-Task Experiments

Tasks	Arabizi input (class.)	CODA* input
Token.	**80.0% (93.0%)**	**95.4%**
POS	73.8% (92.5%)	54.5%
Lemma	75.5% (92.8%)	**89.5%**
Translit.	**79.0%** (92.8%)	67.2%

Table 2: Starting ST Experiments

Input	Tasks	Accuracy
Token.	POS	86.2%
Token.	**Lemma**	**92.4%**
Token.	**Lemma - POS**	**92.8% - 87.6%**
Token.	POS - Lemma	87.3% - 92.6%

Table 3: Intermediate Experiments

(Hybrid) Single-Task Experiments

Tasks	Arabizi input (class.)	CODA* input
Token.	**80.0% (93.0%)**	**95.4%**
POS	73.8% (92.5%)	54.5%
Lemma	75.5% (92.8%)	**89.5%**
Translit.	**79.0%** (92.8%)	67.2%

Table 2: Starting ST Experiments

Input	Tasks	Accuracy
CODA*	Lemma - POS	89.2% - 84.2%
CODA*	**POS - Lemma**	**85.9% - 90.5%**
CODA*	Token. - POS	95.3% - 85.2%
CODA*	POS - Token.	85.6% - 95.2%

Table 4: Additional Experiments

Input	Tasks	Accuracy
Token.	POS	86.2%
Token.	**Lemma**	**92.4%**
Token.	**Lemma - POS**	**92.8% - 87.6%**
Token.	POS - Lemma	87.3% - 92.6%

Table 3: Intermediate Experiments

(Hybrid) Single-Task Experiments

Tasks	Arabizi input (class.)	CODA* input
Token.	**80.0% (93.0%)**	**95.4%**
POS	73.8% (92.5%)	54.5%
Lemma	75.5% (92.8%)	**89.5%**
Translit.	**79.0%** (92.8%)	67.2%

Table 2: Starting ST Experiments

Input	Tasks	Accuracy
CODA*	Lemma - POS	89.2% - 84.2%
CODA*	**POS - Lemma**	**85.9% - 90.5%**
CODA*	Token. - POS	95.3% - 85.2%
CODA*	POS - Token.	85.6% - 95.2%

Table 4: Additional Experiments

Input	Tasks	Accuracy
Token.	POS	86.2%
Token.	**Lemma**	**92.4%**
Token.	**Lemma - POS**	**92.8% - 87.6%**
Token.	POS - Lemma	87.3% - 92.6%

Table 3: Intermediate Experiments

- Tokenization seems to be the *easiest* task, it helps classification (T2).

(Hybrid) Single-Task Experiments

Tasks	Arabizi input (class.)	CODA* input
Token.	**80.0% (93.0%)**	**95.4%**
POS	73.8% (92.5%)	54.5%
Lemma	75.5% (92.8%)	**89.5%**
Translit.	**79.0%** (92.8%)	67.2%

Table 2: Starting ST Experiments

Input	Tasks	Accuracy
CODA*	Lemma - POS	89.2% - 84.2%
CODA*	**POS - Lemma**	**85.9% - 90.5%**
CODA*	Token. - POS	95.3% - 85.2%
CODA*	POS - Token.	85.6% - 95.2%

Table 4: Additional Experiments

Input	Tasks	Accuracy
Token.	POS	86.2%
Token.	**Lemma**	**92.4%**
Token.	**Lemma - POS**	**92.8% - 87.6%**
Token.	POS - Lemma	87.3% - 92.6%

Table 3: Intermediate Experiments

- Tokenization seems to be the *easiest* task, it helps classification (T2).
- POS benefits from Lemma and *vice versa* (T3).

(Hybrid) Single-Task Experiments

Tasks	Arabizi input (class.)	CODA* input
Token.	**80.0% (93.0%)**	**95.4%**
POS	73.8% (92.5%)	54.5%
Lemma	75.5% (92.8%)	**89.5%**
Translit.	**79.0%** (92.8%)	67.2%

Table 2: Starting ST Experiments

Input	Tasks	Accuracy
CODA*	Lemma - POS	89.2% - 84.2%
CODA*	**POS - Lemma**	**85.9% - 90.5%**
CODA*	Token. - POS	95.3% - 85.2%
CODA*	POS - Token.	85.6% - 95.2%

Table 4: Additional Experiments

Input	Tasks	Accuracy
Token.	POS	86.2%
Token.	**Lemma**	**92.4%**
Token.	**Lemma - POS**	**92.8% - 87.6%**
Token.	POS - Lemma	87.3% - 92.6%

Table 3: Intermediate Experiments

- Tokenization seems to be the *easiest* task, it helps classification (T2).
- POS benefits from Lemma and *vice versa* (T3).
- Lemma helps the POS task as much as/more than Token. (T4).

MT Experiments Based on ST strategy

	Accuracies on tasks			
Exp. ID	**Token.**	**Lemma**	**POS**	**Arabizi**
I	95.4%	-	-	-
II	95.3%	89.8%	-	-
III	**96%**	**90.7%**	**86.2%**	-
IV	94.4%	88.9%	84.5%	67.8%

Table 5: CODA* input

	Accuracies on tasks				
Exp. ID	**Class.**	**Token.**	**Lemma**	**POS**	**CODA***
I	86.2%	-	-	-	-
II	93%	**80**%	-	-	-
III	95%	**80**%	**78.2**%	-	-
IV	94.1%	78.9%	77.5%	77.8%	-
V	**94.2**%	78.9%	77.3%	**78.6**%	**79.5**%

Table 6: Arabizi input

MT Experiments Based on ST strategy

	Accuracies on tasks			
Exp. ID	**Token.**	**Lemma**	**POS**	**Arabizi**
I	95.4%	-	-	-
II	95.3%	89.8%	-	-
III	**96%**	**90.7%**	**86.2%**	-
IV	94.4%	88.9%	84.5%	67.8%

Table 5: CODA* input

	Accuracies on tasks				
Exp. ID	**Class.**	**Token.**	**Lemma**	**POS**	**CODA***
I	86.2%	-	-	-	-
II	93%	**80**%	-	-	-
III	95%	**80**%	**78.2**%	-	-
IV	94.1%	78.9%	77.5%	77.8%	-
V	**94.2**%	78.9%	77.3%	**78.6**%	**79.5**%

Table 6: Arabizi input

- Tasks are auxiliary tasks til the Exp III (T5).

MT Experiments Based on ST strategy

	Accuracies on tasks			
Exp. ID	**Token.**	**Lemma**	**POS**	**Arabizi**
I	95.4%	-	-	-
II	95.3%	89.8%	-	-
III	**96%**	**90.7%**	**86.2%**	-
IV	94.4%	88.9%	84.5%	67.8%

Table 5: CODA* input

	Accuracies on tasks				
Exp. ID	**Class.**	**Token.**	**Lemma**	**POS**	**CODA***
I	86.2%	-	-	-	-
II	93%	**80**%	-	-	-
III	95%	**80**%	**78.2**%	-	-
IV	94.1%	78.9%	77.5%	77.8%	-
V	**94.2**%	78.9%	77.3%	**78.6**%	**79.5**%

Table 6: Arabizi input

- Tasks are auxiliary tasks til the Exp III (T5).
- Visible negative transfer effect of the "Arabizi" task (Exp. IV - T5).

MT Experiments Based on ST strategy

	Accuracies on tasks			
Exp. ID	**Token.**	**Lemma**	**POS**	**Arabizi**
I	95.4%	-	-	-
II	95.3%	89.8%	-	-
III	**96%**	**90.7%**	**86.2%**	-
IV	94.4%	88.9%	84.5%	67.8%

Table 5: CODA* input

	Accuracies on tasks				
Exp. ID	**Class.**	**Token.**	**Lemma**	**POS**	**CODA***
I	86.2%	-	-	-	-
II	93%	**80%**	-	-	-
III	95%	**80%**	**78.2%**	-	-
IV	94.1%	78.9%	77.5%	77.8%	-
V	**94.2%**	78.9%	77.3%	**78.6%**	**79.5%**

Table 6: Arabizi input

- Tasks are auxiliary tasks til the Exp III (T5).
- Visible negative transfer effect of the "Arabizi" task (Exp. IV - T5).
- The most *difficult* task seems to be the POS one (Exp. IV - T6).

Multi-Task Experiments

	Accuracies on tasks				
Exp. ID	**Class.**	**CODA***	**Lemma**	**Token.**	**POS**
I	97.3	82.6	82.3(5)	82.3(3)	71.4(4)
II	**99**	**84.2**	82.8(4)	**83.5**(3)	**83.1**(5)
III	92.9	78.5	54.2(4)	75.9(5)	78(3)
IV	94.3	78.3	76.4(5)	77.9(4)	78.1(3)
V	97.9	**84.3**	**83.6**(3)	82.3(4)	82.3(5)
VI	98.8	83.5	82.4(3)	82.3(5)	82.3(4)

Table 7: Arabizi input

Multi-Task Experiments

	Accuracies on tasks				
Exp. ID	**Class.**	**CODA***	**Lemma**	**Token.**	**POS**
I	97.3	82.6	82.3(5)	82.3(3)	71.4(4)
II	**99**	**84.2**	82.8(4)	**83.5**(3)	**83.1**(5)
III	92.9	78.5	54.2(4)	75.9(5)	78(3)
IV	94.3	78.3	76.4(5)	77.9(4)	78.1(3)
V	97.9	**84.3**	**83.6**(3)	82.3(4)	82.3(5)
VI	98.8	83.5	82.4(3)	82.3(5)	82.3(4)

Table 7: Arabizi input

Table 6: Arabizi input

Exp. ID	**Class.**	**Token.**	**Lemma**	**POS**	**CODA***
V	**94.2**%	78.9%	77.3%	**78.6**%	**79.5**%

Multi-Task Experiments

	Accuracies on tasks				
Exp. ID	**Class.**	**CODA***	**Lemma**	**Token.**	**POS**
I	97.3	82.6	82.3(5)	82.3(3)	71.4(4)
II	**99**	**84.2**	82.8(4)	**83.5**(3)	**83.1**(5)
III	92.9	78.5	54.2(4)	75.9(5)	78(3)
IV	94.3	78.3	76.4(5)	77.9(4)	78.1(3)
V	97.9	**84.3**	**83.6**(3)	82.3(4)	82.3(5)
VI	98.8	83.5	82.4(3)	82.3(5)	82.3(4)

Table 7: Arabizi input

Table 6: Arabizi input

Exp. ID	Class.	Token.	Lemma	POS	CODA*
V	**94.2**%	78.9%	77.3%	**78.6**%	**79.5**%

Multi-Task Experiments

	Accuracies on tasks				
Exp. ID	**Class.**	**CODA***	**Lemma**	**Token.**	**POS**
I	97.3	82.6	82.3(5)	82.3(3)	71.4(4)
II	**99**	**84.2**	82.8(4)	**83.5**(3)	**83.1**(5)
III	92.9	78.5	54.2(4)	75.9(5)	78(3)
IV	94.3	78.3	76.4(5)	77.9(4)	78.1(3)
V	97.9	**84.3**	**83.6**(3)	82.3(4)	82.3(5)
VI	98.8	83.5	82.4(3)	82.3(5)	82.3(4)

Table 7: Arabizi input

- POS was defined as the most difficult task in ST experiments. (78% in T6 - 83% in T7, where POS is the *last* task).

Multi-Task Experiments

	Accuracies on tasks				
Exp. ID	**Class.**	**CODA***	**Lemma**	**Token.**	**POS**
I	97.3	82.6	82.3(5)	82.3(3)	71.4(4)
II	**99**	**84.2**	82.8(4)	**83.5**(3)	**83.1**(5)
III	92.9	78.5	54.2(4)	75.9(5)	78(3)
IV	94.3	78.3	76.4(5)	77.9(4)	78.1(3)
V	97.9	**84.3**	**83.6**(3)	82.3(4)	82.3(5)
VI	98.8	83.5	82.4(3)	82.3(5)	82.3(4)

Table 7: Arabizi input

- POS was defined as the most difficult task in ST experiments. (78% in T6 - 83% in T7, where POS is the *last* task).
- Lemma in in 4th position helps Token. and POS => *cushioning effect*

Multi-Task Experiments

	Accuracies on tasks				
Exp. ID	**Class.**	**Lemma**	**Token.**	**POS**	**Arabizi**
I	97.2	88.8(5)	94.5(3)	83.6(4)	**68.8**(2)
II	98.1	89.3(4)	95.3(3)	83.4(5)	68.3(2)
III	98.1	89.1(4)	95.2(5)	83.4(3)	68.5(2)
IV	97.4	88.6(5)	94.7(4)	83.3(3)	68.4(2)
V	97.8	88.9(3)	95.2(4)	84.3(5)	68.7(2)
VI	97.5	89.2(3)	94.4(5)	83.4(4)	68.3(2)
VII	97.5	89.3(2)	95(4)	83.6(5)	68.7(3)
VIII	**98.3**	89.2(2)	**95.4**(5)	**84.8**(4)	68.6(3)

Table 8: Other MT experiments to predict Arabizi

Multi-Task Experiments

	Accuracies on tasks				
Exp. ID	**Class.**	**Lemma**	**Token.**	**POS**	**Arabizi**
I	97.2	88.8(5)	94.5(3)	83.6(4)	**68.8**(2)
II	98.1	89.3(4)	95.3(3)	83.4(5)	68.3(2)
III	98.1	89.1(4)	95.2(5)	83.4(3)	68.5(2)
IV	97.4	88.6(5)	94.7(4)	83.3(3)	68.4(2)
V	97.8	88.9(3)	95.2(4)	84.3(5)	68.7(2)
VI	97.5	89.2(3)	94.4(5)	83.4(4)	68.3(2)
VII	97.5	89.3(2)	95(4)	83.6(5)	68.7(3)
VIII	**98.3**	89.2(2)	**95.4**(5)	**84.8**(4)	68.6(3)

Table 8: Other MT experiments to predict Arabizi

	Accuracies on tasks				
Exp. ID	**Class.**	**Lemma**	**Token.**	**POS**	**CODA***
I	94.1	76.3(4)	77.9(2)	77.9(3)	78.1
II	94.2	77.3(3)	**78.9**(2)	78.6(4)	**79.5**
III	94	77.2(3)	78.2(4)	78.5(2)	78.2
IV	93.8	76.3(4)	78.1(3)	78.1(2)	78
V	94	77.2(2)	78.4(3)	78.5(4)	78.5
VI	94.2	77.3(2)	78.7(4)	78.8(3)	78.7

Table 9: Other MT experiments to predict CODA*

Contents

General Conclusions

- It clearly emerges the existence of **task-relations** in ST experiments.

General Conclusions

- It clearly emerges the existence of **task-relations** in ST experiments.
- **Linguistic intuition** would say that morphological tasks necessarily support other morphological tasks. This is not always the case (Lemma supports POS).

General Conclusions

- It clearly emerges the existence of **task-relations** in ST experiments.
- **Linguistic intuition** would say that morphological tasks necessarily support other morphological tasks. This is not always the case (Lemma supports POS).
- It seems better to isolate *difficult* tasks (POS), leaving it as the last task and preceding it with a *simple* and *related* task (Lemma), exploiting its **cushioning effect to the negative transfer**.

General Conclusions

- It clearly emerges the existence of **task-relations** in ST experiments.
- **Linguistic intuition** would say that morphological tasks necessarily support other morphological tasks. This is not always the case (Lemma supports POS).
- It seems better to isolate *difficult* tasks (POS), leaving it as the last task and preceding it with a *simple* and *related* task (Lemma), exploiting its **cushioning effect to the negative transfer**.

General Conclusions

- It clearly emerges the existence of **task-relations** in ST experiments.
- **Linguistic intuition** would say that morphological tasks necessarily support other morphological tasks. This is not always the case (Lemma supports POS).
- It seems better to isolate *difficult* tasks (POS), leaving it as the last task and preceding it with a *simple* and *related* task (Lemma), exploiting its **cushioning effect to the negative transfer**.

- **ST learning logic** has been shown to be an **uncertain strategy** compared to an MT strategy.

General Conclusions

- It clearly emerges the existence of **task-relations** in ST experiments.
- **Linguistic intuition** would say that morphological tasks necessarily support other morphological tasks. This is not always the case (Lemma supports POS).
- It seems better to isolate *difficult* tasks (POS), leaving it as the last task and preceding it with a *simple* and *related* task (Lemma), exploiting its **cushioning effect to the negative transfer**.

- **ST learning logic** has been shown to be an **uncertain strategy** compared to an MT strategy.
- Specific task ordering in an **MT robust system with attention mechanism** matters up to a certain point.

References I

Rich Caruana. Multitask learning. *Machine learning*, 28(1):41–75, 1997.

Soravit Changpinyo, Hexiang Hu, and Fei Sha. Multi-task learning for sequence tagging: An empirical study. *arXiv preprint arXiv:1808.04151*, 2018.

Elisa Gugliotta and Marco Dinarelli. Tarc: Tunisian arabish corpus first complete release. In *13th Conference on Language Resources and Evaluation (LREC 2022)*, 2022.

Elisa Gugliotta, Marco Dinarelli, and Olivier Kraif. Multi-task sequence prediction for Tunisian Arabizi multi-level annotation. *arXiv preprint arXiv:2011.05152*, 2020.

Nizar Habash et al. Unified guidelines and resources for Arabic dialect orthography. In *Proceedings of the Eleventh International Conference on LREC*, 2018.

Mohamed Maamouri, Ann Bies, Tim Buckwalter, and Wigdan Mekki. The Penn Arabic Treebank: Building a large-scale annotated Arabic corpus. In *NEMLAR conference on Arabic language resources and tools*, volume 27, pages 466–467. Cairo, 2004.

Sebastian Ruder. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*, 2017.

Thank you very much for your attention :)