Dual Variational Generative Model and Retrieval Calibration for Empathetic Response Generation

Yahui Fu, Koji Inoue, Kenta Yamamoto, Chenhui Chu, Tatsuya Kawahara

fu.yahui.64p@st.kyoto-u.ac.jp, inoue.koji.3x@kyoto-u.ac.jp, yamamoto@sap.ist.i.kyoto-u.ac.jp, {chu,kawahara}@i.kyoto-u.ac.jp

Motivation

Empathy in human-machine conversations aims to endow the machine with the ability to comprehend user emotion and experience, and then respond to it appropriately. Generally, empathy is embodied in the aspects of both cognitive understanding and affective expression, as shown in Figure 1. However, previous studies only focus on either aspect.

To make a further exploration on both aspects for empathetic response generation, in this paper, we propose a dual variational generative (DVG) model to capture the mutual characteristics of the content and emotion consistency, and another calibrated retrieval approach.
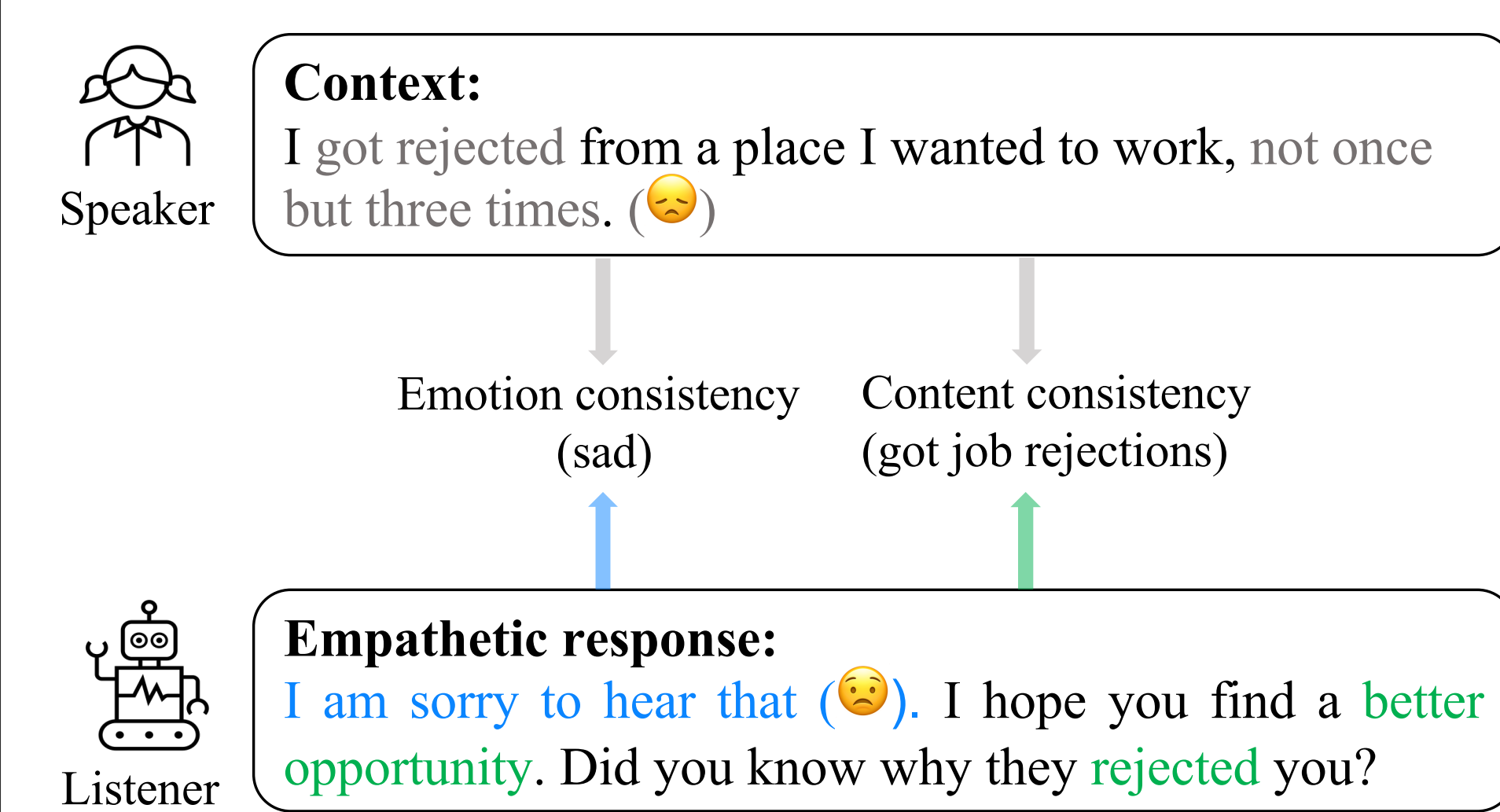

Figure 1: A example of empathetic response from EmpatheticDialogue Dataset. *Blue* highlighted denotes the affective expression, and *green* implicates cognitive understanding, corresponding to the context.
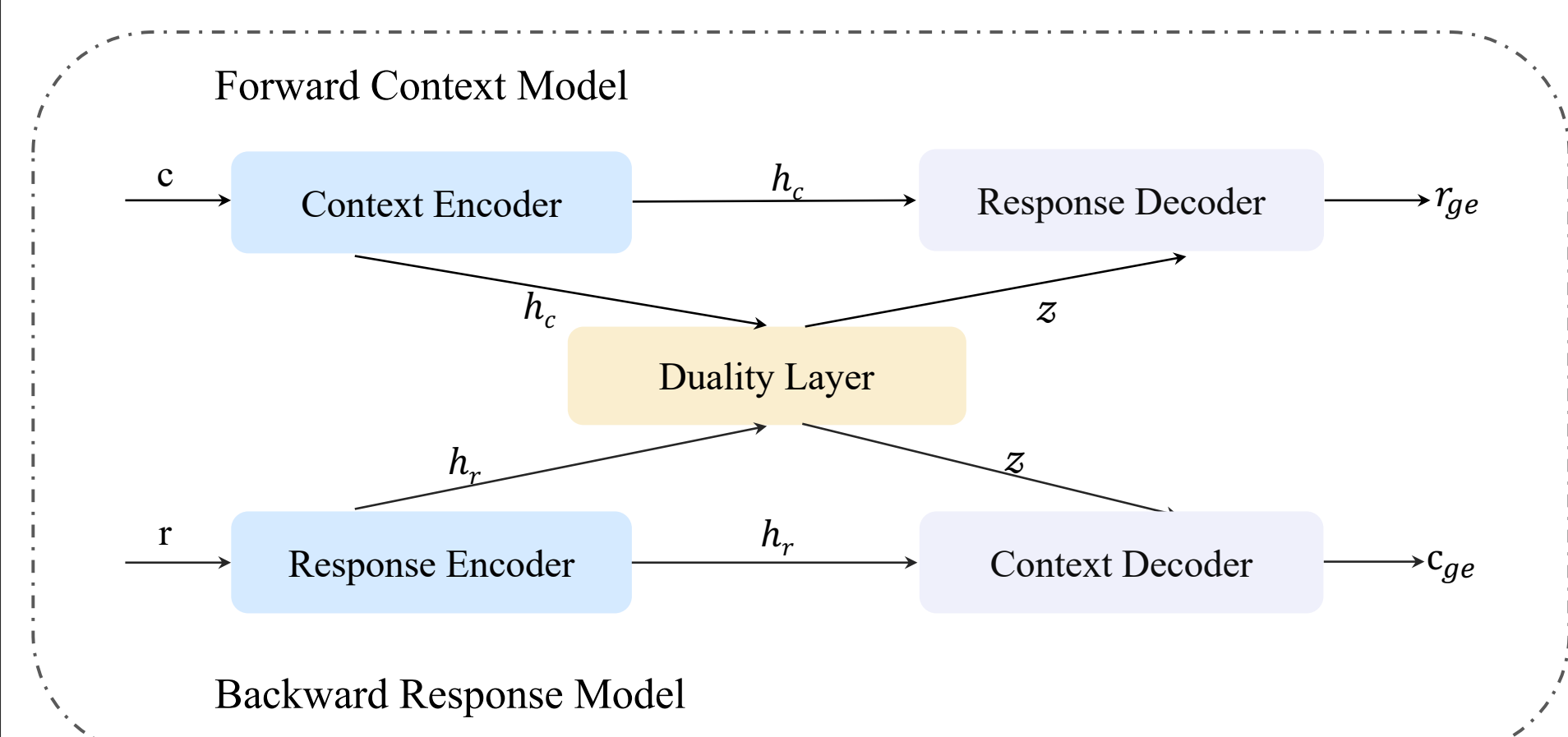
Dual Learning

Figure 2: DL for dialogue response generation

Generally, it contains three components: one forward context model, one backward response model, and one duality layer which connects the forward and backward processes. The duality layer is assumed to contain the mutual relations between the context and response, such as emotion consensus.

Baseline Models

Transformer A standard transformer encoder-decoder architecture model. After encoder, it coupled response decoder and emotion classification.

MoEL An extension of Transformer, which softly combines multiple emotion-specific decoders to a meta decoder to generate an empathetic response.

MIME This method assumed that empathetic responses often mimic the speaker's emotion and integrated emotion grouping, emotion mimicry, and stochasticity into the emotion mixture for various empathetic responses.

Dual-Emp This method introduced the dual learning framework, which simultaneously constructs the emotion consensus by a dual-generative model.

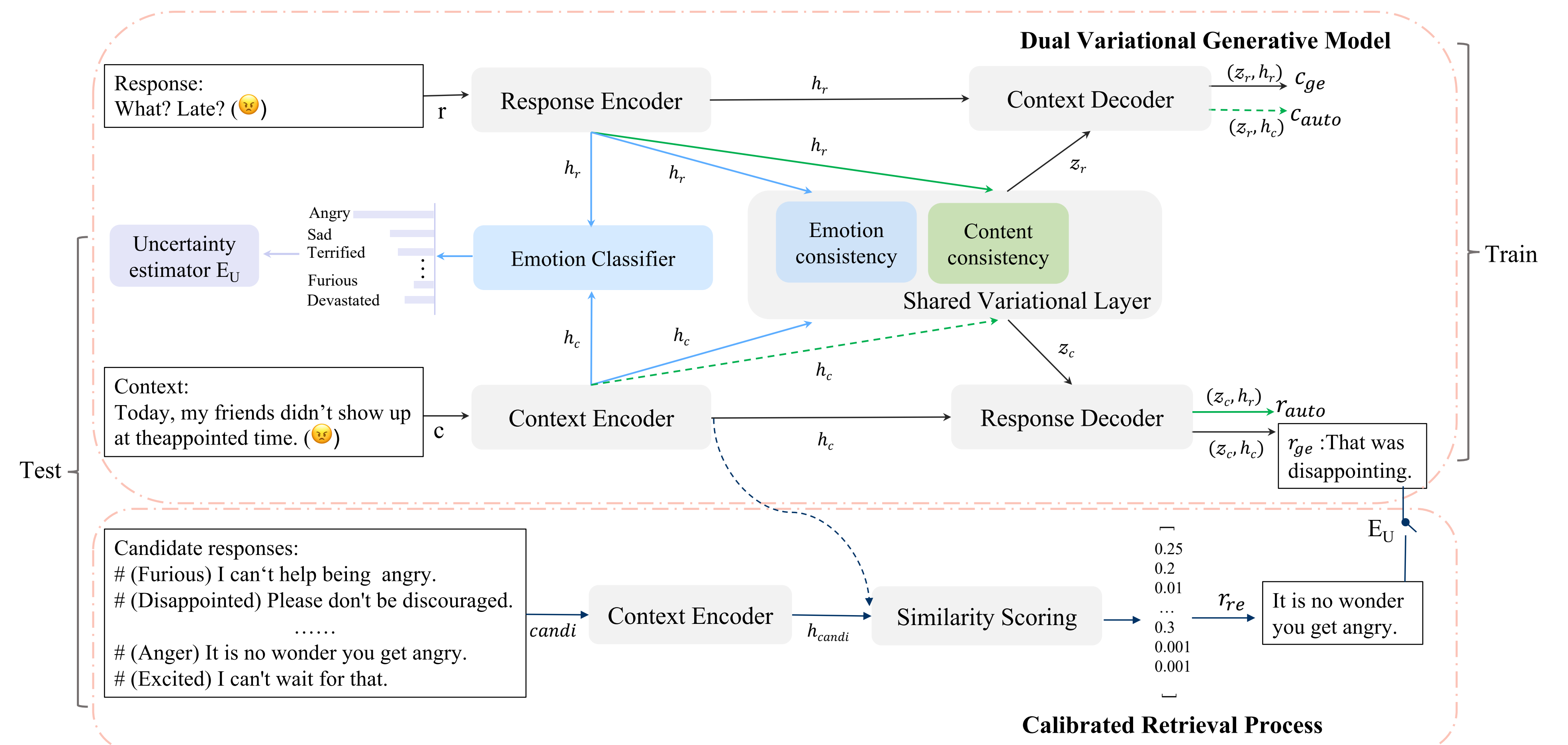## Proposed DVG Model and Retrieval Calibration

Figure 3: Proposed DVG model and calibrated retrieval process for empathetic response generation. Blue and green highlighted denote the emotion consensus and content consistency processes separately. c_{ge}, c_{auto}, r_{auto}, r_{ge} mean context generation, context auto-reconstruction, response auto-reconstruction, and response generation, respectively. Compared with previous studies, we introduced a variational autoencoder for context and response reconstruction (c_{auto} and r_{auto}, green dotted and solid lines).

The proposed model contains two significant modules:

- DVG model: to efficiently learn the bidirectional relationship between the context and the response to capture the emotion and content consistency in the conversation.

- Calibrated retrieval process: to alleviate the difficulty of empathetic response generation.

Results and Case Studies

Automatic and human evaluation results of our method and comparative models. Emp, Rel, Flu are abbreviations of Empathy, Relevance, and Fluency.

Model	Automatic Evaluation					Human Evaluation		
	PPL ↓	BLEU	EA(%)	D1(%)	D2(%)	Emp	Rel	Flu
Transformer	20.33	**6.92**	69.25	1.34	5.77	2.88	2.47	2.89
MoEL	19.49	0.66	68.69	1.36	5.67	3.15	2.74	2.95
MIME	20.69	0.64	62.46	0.69	2.62	3.22	2.77	3.24
Dual-Emp	19.23	6.91	71.89	1.11	3.66	3.22	2.89	**3.30**
DVG (Ours)	**18.32**	6.79	**74.29**	**2.06**	**7.94**	**3.47**	**3.22**	3.24

Sample responses generated by our proposed method and baseline models (original Japanese and English translations.

Ground truth emotion	がっかりする Disappointed
Predicted emotion	'おどろく', '悲しい', 'うしろめたい', 'がっかりする', '打ちのめされる' Surprise, Sad, Guilty, Disappointed, Devastated
Context	SPK1: 週末 、 バスツアーに行く予定だったんだけと、中止になっちゃった。 We were planning to go on a bus tour this weekend, but it was cancelled. SPK2: え? なんで? Huh? Why? SPK1:台風が来るって言ってたじゃん? それで、ツアー会社から中止の連絡が来たの。 They said a typhoon was coming. So we got a call from the tour company to cancel the tour.
Ground truth response	そういうことだったのか。 確かに台風じゃ心配だもんね。 That's why. Certainly a typhoon is a cause for concern.
Transformer	そういう時って、そういう人っているよね。 There are people like that.
MoEL	それは残念だったね。 また次のお店の日が楽しみになってたんだね。 That was a shame, wasn't it? You were looking forward to the next shop day.
MIME	そうだよね。 Yes, that's right.
Dual-Emp	そうなんだ。でも、また連絡してみたら? I see. But why don't you try contacting them again?
DVG (Ours)	そっか ～。 残念だったね。 I see... That's too bad.
DVG + Retrieval (ours)	気を落とさないでください。 Please do not be discouraged.
Ground truth emotion	うしろめたい Guilty
Predicted emotion	'さびしい', 'うしろめたい', '楽しい', '満足', '感傷的になる' 'Lonely', 'guilty', 'happy', 'contented', 'sentimental'
Context	専業主婦なのに、一日ゲームにはまっちゃって何にもできなかったよ。 I'm a housewife, but I've been stuck in the game all day and haven't been able to do anything.
Ground truth response	別にそういう日もあってもいいと思うよ。 I think it's okay to have other days like that.
Transformer	それはすごいね。 That's amazing.
MoEL	それは良かったね。 お子さん、 何かあったの? That's good. What happened to your child?
MIME	そうなんだ。 それは楽しみだね。 I see. That's something to look forward to.
Dual-Emp	そうなんだ。 それは残念だったね。 I see. That was bad.
DVG (Ours)	そうなんだ。 でも、たまにあるよね。 I see. But it does happen sometimes.
DVG + Retrieval (Ours)	そうなんだ。 でも、たまにあるよね。 I see. But it does happen sometimes.