

Supplementary Material for What Code-Switching Strategies are Effective in Dialogue Systems?

Emily Ahn^{1*} Cecilia Jimenez² Yulia Tsvetkov³ Alan W Black³

¹University of Washington ²University of Pittsburgh ³Carnegie Mellon University
eahn@uw.edu, CCJ16@pitt.edu, {ytsvetko, awb}@cs.cmu.edu

1 Crowdsourcing Information

We presented most information and instructions in Spanish to the crowdworkers. The preliminary instructions encouraged users to “use English, Spanish, or a mixture of the two”. The title of the crowdsourced task was also called “*Charlemos en Spanglish!*”, meaning “Let’s chat in Spanglish!” Our modified bilingual interface is shown in Figure 1.

¿Quién es nuestro amigo en común?

Usted y otro usuario tienen exactamente un amigo en común. Usted sabe unos atributos para cada uno de sus amigos (como nombre, escuela, etc.). Tu meta es encontrar el amigo en común usando estos atributos!

Instrucciones

- Por favor use **oraciones naturales** lo más posible.
 - **Escribe:** tres de mis amigos trabajan en el banco
 - **No Escribas:** 3 banco
- Evite simplemente enumerar los atributos (nombre, compañía, etc.) de sus amigos.
- Mira la **lista de amigos** a la derecha.
- Usa la **ventana de chat** para descubrir más de los amigos de su pareja.

Tus amigos / Your friends

# ▲▼	Especialidad Mayor ▲▼	Hobby Hobby ▲▼	Tiempo Time ▲▼
Select	educación education	actuar acting	la noche night
Select	administración management	el ganchillo crocheting	la mañana morning
Select	diseño gráfico graphic design	correr running	la noche night

Time / Tiempo: 7:10

[06/22/18 17:30:02] <You entered the room.>
[06/22/18 17:30:03] <Your partner has joined the room.>

Enter your message here

Figure 1: Crowdsourced users would see an interface similar to this when chatting with our dialogue system via text. Unique features include the bilingual table of friend attributes.

2 Post-Task Survey Questions

After chatting with our agent, the user would fill out this post-task survey, given in English.

*This work was done while the first author was a student at Carnegie Mellon University.

Part 1: This Task

To what extent do you agree with these statements? Mark answers on a 5-point Likert scale: “Strongly disagree” (1), “Disagree” (2), “Neutral” (3), “Agree” (4), “Strongly agree” (5).

1. I understood the task perfectly.
2. My task partner was very cooperative.
3. My task partner was very human-like.
4. My task partner perfectly understood what I was typing.
5. I am very likely to chat like I did in this task when messaging with my bilingual friends.
6. My task partner texts like someone I know.
7. If technology (like Siri or Alexa) progressed to allow me to use Spanglish, I could see myself using both languages with it.

Part 2: Linguistic Background

1. At what age did you begin to learn Spanish?
2. At what age did you begin to learn English?
3. What is your current age?
4. What is your gender?
5. Rate your overall Spanish ability on a scale from 1-5
6. Rate your overall English ability on a scale from 1-5
7. Which country/countries do you or your family originate from?
8. What percentage of the time do you write online (i.e. text, social media, etc.) in Spanish?
9. What percentage of the time do you write online (i.e. text, social media, etc.) in English?
10. What percentage of the time do you write online (i.e. text, social media, etc.) in mixed Spanish and English?

3 List of Discourse Markers for the Casual Formality Setting

1. ah
2. ay

3. *este*
4. *hey*
5. *nah*
6. *oh*
7. *ok*
8. *oye*
9. *pues*
10. *so*
11. *yeah*
12. *.. y tú?*