Why should a dialogue system speak more than one language?

Jacqueline Brixey and David Traum

Abstract Many people can speak more than one language (bi- or multi-linguals). Often this manifests as speaking and understanding more than one language within the same dialogue (called code-switching). To support this kind of interaction with dialogue systems, we need bilingual dialogue systems, that can do the same. While many current systems can engage in dialogue in more than one language, the language processing components are distinct for each language and the systems are not capable of maintaining a consistent identity and dialogue state across languages. In this paper, we explore aspects of a taxonomy of bilingual dialogue systems, laying out the desirable behaviors that such systems should engage in and what might be required to achieve these behaviors, as well as characterizing where current systems fit in this taxonomy.

1 Introduction

Bilingualism is the ability to speak two languages. Code-switching is the switching between languages within a single dialogue. Switching between utterances, inter-sentential switching, is distinct from intra-sentential switching which is switching within a single utterance. While code-mixing is also a term that is used to describe code-switching, we will use the term code-switching throughout. [5] points out some contexts that promote code-switching: giving reported speech, change of interlocutor(s) in a group setting, parentheses or side-comments, reiterations, change of task, topic shift, language play, interjections or filler, and false start repair.

Bilingual dialogue systems are important to consider and implement because many people in the world are bilingual and use both languages within a single conversation (called code-switching). Programs that could respond to a user in several languages, and potentially within a single utterance, would allow more users to benefit and interact with dialogue systems comfortably, for a broader range of purposes.

USC Institute for Creative Technologies, e-mail: [last name]@ict.usc.edu

The current state of the art of bilingual dialogue systems includes some that can speak multiple languages, such as Amazon Alexa which can respond to queries in English or Spanish, among other languages. However, the state of the art does not fully support intra-sentential or intra-word code-switching. [11] interviewed users about their experiences interacting with existing bilingual dialogue systems. Users often found that the bilingual dialogue systems behaved more like two monolingual systems stitched together and lacked bi-cultural knowledge. Users also noted that the agents were inconsistent in their personas. For example, "I often switch my language setting because I work both in Korean and English. However, the agent has a very calm male voice in Korean, whereas a cold female voice in English. Although I'm the same person, I feel like I'm talking to two different systems separately." More recent research has started to address having a consistent voice across multiple languages [41].

Additionally, users expressed hope that the systems would code-switch in ways similar to their own, and were constantly analyzing the system for similarities. Some of the interviewees indicated that if the agent's multilingual and multicultural abilities were similar to their own, they would feel accepted and think their code-switching was "being acknowledged as an effective communication tool in the community rather than evidence of low language proficiency". However, most participants in the study found that because of constant failures and challenging experiences, the users were "convinced they were not designed to be the primary users of the conversational agents."

In this paper, we begin to lay out a taxonomy of bilingual dialogue systems, characterizing both where current systems fit in this taxonomy as well as areas where there may be a need for such systems in the future. There are multiple dimensions to this taxonomy, including the linguistic level at which code-switching happens (intra-word, intra-sentence, inter-sentence), as discussed above; how code-switching is supported within a system (the modules that need to be adjusted, the algorithms and data to support various approaches); and why the system should engage in code-switching, which has implications on what degree of performance is necessary to support the goals. In this paper, we focus mainly on the "why" dimension, though touching briefly on the others. The next section presents some of the main reasons for a system to engage in and understand code-switching with a human user, with examples of these motivations from the linguistic and computational linguistic literature. Section 3 briefly mentions some computational challenges that arise in trying to implement bilingual dialogue systems and summarizes aspects of the taxonomy concerning the reasons that existing systems engage in code-switching and revealing current gaps.

2 Selection of related work

The related work that was reviewed in this taxonomy was found through two methods. In the first method, we searched Google Scholar for terms such as "bilingual

dialogue systems", "bilingual dialog systems", "bilingual conversational systems", all of the previous terms using "multilingual" in place of "bilingual", and finally all of the previous terms replacing "bilingual" with "code-switching" and "code-mixing". We reviewed several survey papers on code-switching, such as [9]. We also reviewed the references of any publications that were found to locate additional papers not captured through searching Google Scholar.

We only considered publications in which a dialog system specifically was described, even if not fully implemented. We excluded corpora-only publications. In total, we were able to identify fourteen publications that met these qualifications.

3 Reasons for bilingual interactions

There are many distinct reasons why a system might need to engage in bilingual dialogue. These include user preference, the task required for code-switching, the system being an unbalanced bilingual, meeting socio-psychological goals, and accommodating the user's language proficiency. We discuss each in the following subsections, including examples of previously developed systems that support these reasons, where available.

3.1 User preference

Users may code-switch based on personal preference, whether they possess adequate proficiency or wish to practice a language they are learning. Balanced bilingualism systems support such users, as observed in instances like the Croatian-Slovene weather forecasting system [23]; the trilingual Cantonese, Putonghua, and English system [24]; a Swedish-English news discussion system [16]; and a Japanese-English robot with a knowledge base from Wikipedia [42]. In these cases, users initiated code-switching explicitly or through language switches.

Most large language models, such as Chat-GPT [8], are trained on multiple languages and can respond to these languages individually. These models can respond to a user's request for code-switching, as demonstrated in the bottom example of Figure 1. However the models do not always respond to a user's query when it includes code-switching, as shown in the top example of the same figure.

Two systems, discussed in [4] and [3], accommodated users speaking a second language but interacting in a different one. The former involved avatars speaking English, Arabic, or a mix, with users choosing avatars, while the latter proposed an agent connecting users to a module for network operating system learning. Although not built, it appeared that both systems could potentially handle code-switching, although intra-sentential code-switching was not explicitly addressed.

Exploring user responses to code-switching alterations, [6] presented text-based conversations in English and Hindi, varying code-switching usage between partic-

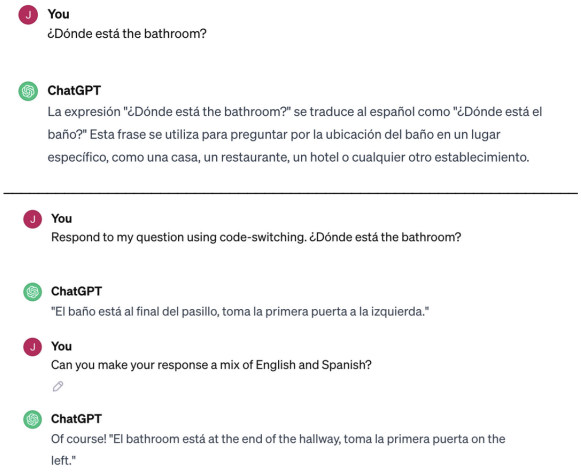

Fig. 1 In the top example, the user directly asks a question using code-switching. The question asked is "Where is the bathroom?", to which the system gives a corrected form in Spanish. In the bottom example, the user requests for the Chat-GPT system to use code-switching. The system provides the code-switched response on the second attempt.

ipants. Participant attitudes toward code-switching significantly influenced perceptions of its presence and quality. They suggested assessing user attitudes before incorporating code-switching. [7] introduced "nudging" by including small phrases from the other language, finding that bilingual speakers preferred code-switching agents for higher conversational ability and human likeness. Nudging proved a useful method to gauge user attitudes towards code-switching.

3.2 Task required

For some languages, specific vocabulary is borrowed from another language. In these instances, the language requires to include these borrowings. For example, the word "email" is typically used globally for electronic mail, even if a given language has created a specific noun phrase, such as *correo electrónico* in Spanish. Hindi is an example language for code-switching and borrowing, with numerous borrowings from English, as exemplified respectively in the following utterances.

- From [13]: very cute, bachpan main bhi face pe attitude hain, very very cute
  (*very cute, even in childhood there is that attitude on his face, very very cute*
- From [12]: Maim ais University ka internship kar raha hoon.
  (*I am doing an internship at this university.*

One reason a user might expect a bilingual dialogue system would be due to the conversational task requiring bilingualism. Dialogue systems support numerous

types of conversations and tasks. Most existing code-switching dialogue systems are task-oriented, such as for making reservations or retrieving information [33].

The system in [32] was designed to build a corpus of code-switched human-system data. The system would initiate the code-switching in English, Spanish, or Hindi to encourage the user to also code-switch between those languages. Code-switching was intra-sentential to capture this variety of language use.

3.3 System is an unbalanced bilingual

A system may resort to code-switching as a strategy to compensate for its limitations in one or both languages. Typically, dialogue systems are crafted based on established norms for human-to-human communication. The development pipeline involves collecting data from human interlocutors exhibiting specific behaviors, creating the necessary computational components, and subsequently testing the system with human users. While it's recognized that expectations for a computer interlocutor may vary in some respects, they often align with human expectations. Therefore, the initial step in constructing a dialogue system involves obtaining a data set of human conversations, which serves as the foundation for modeling the understanding and generation components of the system.

An inherent challenge arises from the fact that out of the 7,000 languages worldwide [18], numerous languages lack adequate computational resources to support the implementation of a dialogue system. An example of an unbalanced bilingual system addressing this challenge is Tauira [35]. In Tauira, the system prompts users to provide the correct usage of an unfamiliar word in a second language so that the system can expand its English vocabulary and offer translations in Māori.

3.4 Socio-psychological

Code-switching offers bilingual speakers a method to increase the flexibility of expression and allows them to "index the nuances of social relationships by exploiting the socio-psychological associations of the languages employed" [26]. Language choice might simply follow the expected social cues of the language community, but it can also communicate aspects about the speaker, such as their identity and group membership(s). The choice of language might additionally be used to build a relationship with the interlocutor.

3.4.1 Cues of language community

While code-switching can be speaker dependent [40], [15] suggests that factors such as age can influence language choice [14]. [5] points out that a speaker can distance

themselves from the language community by not abiding by norms. [28] found that Czech-English speakers switch to English for high-information content words in prominent prosodic positions when speaking Czech. For Hindi-English bilingual users of Twitter, Hindi is preferred for the expression of negative sentiment [34] and for swearing [1]. [29] investigated how a user will accommodate the dialogue system's code-switching choices. The system in [2] explicitly asks users to rate if the dialogue system code-switched as a human member of the language community would. The dialogue system in [32] had dialogue utterances that were modeled on natural human code-switching to encourage participants to produce additional natural code-switching examples to form a corpus.

3.4.2 Identity establishment

Social identity is the sense of membership within a social group. Membership could be defined by ethnicity, gender, social class, and/or nationality [25], but also language can be a salient marker [30]. Speakers may choose to speak in the language that they feel would best symbolize their identity [27]. Language learners can transform their identity to include the second language as part of their overall identity as they gain knowledge in that language [25]. Code-switching could then be an attempt to negotiate a different balance of rights, obligations, and identity representation within the conversation [27, 30]. [29, 2, 32] all demonstrated their identities as speakers of both languages through code-switching. All of the systems intended to collect human-system dialogue data, which was achieved by the system demonstrating its knowledge of both languages.

3.4.3 Facework and Rapport

Face, as defined by [17], is the image individuals hold of themselves, shaped by socially approved attributes like their profession. It intertwines with identity, acting as the public representation of one's self [39, 22]. [17] notes that an individual's face is reinforced when the image they project aligns internally and is supported by their conversational partners. The balance of face is crucial, as threats or support to a partner's face simultaneously impact one's own [22], leading to the use of politeness strategies to safeguard everyone's face [17].

Rapport management, grounded in face sensitivities, perceived social rights and obligations, and interaction goals [38], plays a pivotal role. Language choice becomes a tool, signaling solidarity or deference, thereby altering social distance. By supporting the interlocutor's face and diminishing social distance, the foundation for rapport is created. Conversely, language choices indicating power differences, anger, or resistance can widen social gaps [30]. When a speaker's face is threatened, perhaps due to a language barrier, emotions like frustration, vulnerability, shame, and anger may emerge [39], while positive recognition fosters confidence. Face threats can reshape one's sense of language identity [17].

To our knowledge, no existing dialogue systems have incorporated code-switching as a deliberate strategy for constructing positive or negative facework and rapport with users.

3.5 User proficiency

While a prevailing stereotype about bilinguals is a flawless command of both languages, this represents the exception rather than the rule [20]. In reality, most bilinguals exhibit a dominant language–unbalanced bilingualism. Various factors contribute to this imbalance in proficiency [36], such as use only with family or in professional settings.

The systems discussed in Section 3.1 support user preferences, but these preferences may involve underlying proficiency issues. [2] found that more proficient bilinguals tend to favor code-switching at the clause level rather than resorting to word insertions. While dialogue systems could potentially leverage code-switching in domains like computer-assisted language learning (CALL), it's noteworthy that, to the best of our knowledge, there are no existing CALL dialogue systems with support for code-switching.

4 Computational considerations

There are several computational challenges to consider when implementing a bilingual dialogue system. Below are some of the current tasks:

- **Code-switched Data:** Code-switching data poses unique challenges, impacting language identification, parsing, and intra-word mixing [9]. Example sparsity and diversity demands specialized tools and approaches.
- **Language Modeling:** Code-switch language modeling is hindered by the lack of data and challenges in text normalization [9]. Predicting language switches and handling intra-sentential and intra-word code-switching are ongoing challenges.
- **Language Identification:** Achieving high accuracy requires distinguishing between inter-sentential and intra-sentential code-switching [9].
- **Text Normalization:** An often essential step when working with languages with distinct writing systems [9].
- **Speech Recognition:** Splitting code-switched speech for monolingual recognition loses semantic information. Integrated systems with multilingual models offer advantages, but effectiveness varies across languages [10]. Meta transfer learning outperforms paired monolingual recognizers [43].
- **Generation of code-switched data:** Difficulty in obtaining code-switched data prompts synthetic generation efforts [31]. Diverse approaches include pre-trained encoders, dependency parsers, and machine translation [21], [19].

5 Summary of the research space and future directions

Table 1 summarizes the code-switching dialogue systems research space as outlined by the dimensions of the taxonomy. What can be noted is the absence of dialogue systems that produce or process intra-word code-switching, few systems address the system becoming bilingual as a reason to code-switch, and few systems have been developed to address socio-psychological reasons.

	Type of code-switching processed/generated			
Reasons for code-switching	**Interturn**	**Intraturn, intersentential,**	**Intraturn, intrasentential,**	**Intra-word**
User preference	[22], [23], [15], [41], [4], [3]	[6], [7], [8]	[6], [7], [8]	
Task required	[31]	[32]	[32]	
System becoming bilingual	[34]			
User proficiency				
Socio-psychological - language community cue				
Socio-psychological - demonstrate an identity		[2], [29], [32]	[2], [29], [32]	
Socio-psychological - facework				
Socio-psychological - rapport				

Table 1 Summary of existing code-switching dialogue systems research space

While there are many challenges to the NLP aspect of code-switching systems, [37] notes, "a code-switching intelligent agent has to be more than just the sum of parts that can handle code-switching. To build effective systems that can code-switch, we will also have to leverage the work done in sociolinguistics to understand how, when, and why to code-switch." In sum, novel dialogue systems and policies need to be developed to build systems that accurately reflect how, when, and why to switch between languages. Future directions could expand on the many reasons for code-switching, including those without previous research, and implement systems that represent more of the reasons for code-switching.

Acknowledgments

We thank the reviewers for their helpful feedback. The work depicted here was sponsored by the U.S. Army. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

References

1. Agarwal, P., Sharma, A., Grover, J., Sikka, M., Rudra, K., Choudhury, M.: I may talk in English but gaali toh hindi mein hi denge: A study of English-Hindi code-switching and swearing pattern on social networks. In: 2017 9th international conference on communication systems and networks (comsnets), pp. 554–557. IEEE (2017)
2. Ahn, E., Jimenez, C., Tsvetkov, Y., Black, A.W.: What code-switching strategies are effective in dialog systems? In: Proceedings of the Society for Computation in Linguistics 2020, pp. 213–222 (2020)
3. Al-Shawakfa, E., Evens, M.: Bilingual dialogs with a network operating system. In: Proceedings of Midwest Artificial Intelligence and Cognitive Science Conference, vol. 99, pp. 7–12 (1999)
4. Ali, D.A., Ahmad, M., Al Hassan, H., Dozsa, P., Hu, M., Varias, J., Habash, N.: A bilingual interactive human avatar dialogue system. In: Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue, pp. 241–244 (2018)
5. Auer, P.: The pragmatics of code-switching: a sequential approach. In: L. Milroy, P. Muysken (eds.) One speaker, two languages, chap. 6, pp. 115–135. University of Cambridge (1995)
6. Bawa, A., Choudhury, M., Bali, K.: User perception of code-switching dialog systems. In: 15th International Conference on NLP, vol. 171 (2018)
7. Bawa, A., Khadpe, P., Joshi, P., Bali, K., Choudhury, M.: Do multilingual users prefer chatbots that code-mix? Let's nudge and find out! Proceedings of the ACM on HCI **4**(CSCW1), 1–23 (2020)
8. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al.: Language models are few-shot learners. Advances in neural information processing systems **33**, 1877–1901 (2020)
9. Çetinoğlu, Ö., Schulz, S., Vu, N.T.: Challenges of computational processing of code-switching. In: Proceedings of the 2nd Workshop on Computational Approaches to Code Switching, pp. 1–11 (2016)
10. Chapuis, E., Colombo, P., Labeau, M., Clavel, C.: Code-switched inspired losses for generic spoken dialog representations. In: Proceedings of the 2021 Conference on Empirical Methods in NLP (2021)
11. Choi, Y.J., Lee, M., Lee, S.: Toward a multilingual conversational agent: Challenges and expectations of code-mixing multilingual users. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1–17 (2023)
12. Dey, A., Fung, P.: A Hindi-English code-switching corpus. In: LREC, pp. 2410–2413 (2014)
13. Dąbrowska, M.: Functions of code-switching in Polish and Hindi facebook users' posts. Studia Linguistica Universitatis Iagellonicae Cracoviensis (130), 63–84 (2013)
14. Doğruöz, A.S., Sitaram, S., Bullock, B., Toribio, A.J.: A survey of code-switching: Linguistic and social perspectives for language technologies. In: Proceedings of the 59th Annual Meeting of the ACL and the 11th International Joint Conference on NLP, pp. 1654–1666 (2021)
15. Gardner-Chloros, P., Edwards, M.: Assumptions behind grammatical approaches to code-switching: when the blueprint is a red herring. Transactions of the Philological Society **102**(1), 103–129 (2004)
16. Gawronska, B., House, D.: Information extraction and text generation of news reports for a Swedish-English bilingual spoken dialogue system. In: 5th International Conference on Spoken Language Processing (1998)
17. Goffman, E.: On face-work. In: Interaction ritual: Essays on face-to-face interaction., chap. 2, pp. 5–46. Aldine (1967)
18. Gordon Jr, R.G.: Ethnologue, languages of the world. http://www. ethnologue. com/ (2005)
19. Gregorius, B., Okadome, T.: Generating code-switched text from monolingual text with dependency tree. In: Proceedings of the The 20th Annual Workshop of the Australasian Language Technology Association, pp. 90–97. Australasian Language Technology Association, Adelaide, Australia (2022)

20. Grosjean, F.: Bilingualism: A short introduction. In: F. Grosjean, P. Li (eds.) The psycholinguistics of bilingualism, chap. 1, pp. 1–25. Wiley-Blackwell Chichester (2013)
21. Gupta, D., Ekbal, A., Bhattacharyya, P.: A semi-supervised approach to generate the code-mixed text using pre-trained encoder and transfer learning. In: Findings of the ACL: EMNLP 2020, pp. 2267–2280 (2020)
22. Holtgraves, T.: Face, politeness and interpersonal variables: implications for language production and comprehension. In: M. Haugh, F. Bargiela-Chiappini (eds.) Face, communication and social interaction, chap. 10, pp. 192–207. Equinox London (2009)
23. Martincic-Ipsic, S., Zibert, J., Ipsic, I., Mihelic, F.: A bilingual spoken dialog system for Slovenian and Croatian weather forecast. In: The IEEE Region 8 EUROCON 2003. Computer as a Tool, vol. 2, pp. 140–143. IEEE (2003)
24. Meng, H.M., Chan, S.F., Wong, Y.F., Fung, T.Y., Tsui, W.C., Lo, T.H., Chan, C.C., Chen, K., Wang, L., Wu, T.Y., et al.: Isis: A multilingual spoken dialog system developed with corba and kqml agents. Citeseer
25. Mitchell, R., Myles, F., Marsden, E.: Second Language Learning Theories. Routledge (2013)
26. Myers-Scotton, C.: Duelling languages: Grammatical structure in codeswitching. Oxford University Press (1997)
27. Myers-Scotton, C.: Codes and consequences: Choosing linguistic varieties. Oxford University Press (1998)
28. Myslín, M., Levy, R.: Code-switching and predictability of meaning in discourse. Language pp. 871–905 (2015)
29. Parekh, T., Ahn, E., Tsvetkov, Y., Black, A.W.: Understanding linguistic accommodation in code-switched human-machine dialogues. In: Proceedings of the 24th Conference on Computational Natural Language Learning, pp. 565–577 (2020)
30. Pavlenko, A., Blackledge, A.: Introduction: New theoretical approaches to the study of negotiation of identities in multilingual contexts. In: A. Pavlenko, A. Blackledge (eds.) Negotiation of Identities in Multilingual Contexts, pp. 1–33. Multilingual Matters LTD (2004)
31. Pratapa, A., Bhat, G., Choudhury, M., Sitaram, S., Dandapat, S., Bali, K.: Language modeling for code-mixing: The role of linguistic theory based synthetic data. In: Proceedings of the 56th Annual Meeting of the ACL, pp. 1543–1553 (2018)
32. Ramanarayanan, V., Suendermann-Oeft, D.: Jee haan, i'd like both, por favor: Elicitation of a code-switched corpus of Hindi-English and Spanish-English human-machine dialog. In: INTERSPEECH, pp. 47–51 (2017)
33. Razumovskaia, E., Glavaš, G., Majewska, O., Korhonen, A., Vulic, I.: Crossing the conversational chasm: A primer on multilingual task-oriented dialogue systems. arXiv preprint arXiv:2104.08570 (2021)
34. Rudra, K., Rijhwani, S., Begum, R., Bali, K., Choudhury, M., Ganguly, N.: Understanding language preference for expression of opinion and sentiment: What do Hindi-English speakers do on twitter? In: Proceedings of the 2016 conference on empirical methods in NLP, pp. 1131–1141 (2016)
35. van Schagen, M., Knott, A.: Tauira: A tool for acquiring unknown words in a dialogue context. In: Proceedings of the Australasian Language Technology Workshop 2004, pp. 131–138 (2004)
36. Schwartz, A.I., Kroll, J.F.: Bilingual lexical activation in sentence context. Journal of memory and language **55**(2), 197–212 (2006)
37. Sitaram, S., Chandu, K.R., Rallabandi, S.K., Black, A.W.: A survey of code-switched speech and language processing. arXiv preprint arXiv:1904.00784 (2019)
38. Spencer-Oatey, H.: Face, identity and interactional goals. In: M. Haugh, F. Bargiela-Chiappini (eds.) Face, communication and social interaction, pp. 137–154. Equinox London (2009)
39. Ting-Toomey, S.: Facework collision in intercultural communication. In: M. Haugh, F. Bargiela-Chiappini (eds.) Face, communication and social interaction, chap. 12, pp. 227–249. Equinox London (2009)
40. Vu, N.T., Adel, H., Schultz, T.: An investigation of code-switching attitude dependent language modeling. In: Statistical Language and Speech Processing: 1st International Conference, SLSP, pp. 297–308. Springer (2013)

41. Wilcock, G.: Bilingual Japanese/English robot dialogues with knowledge graphs and conversational AI. In: The 37th Annual Conference of the Japanese Society for AI (2023)
42. Wilcock, G., Jokinen, K., Yamamoto, S.: What topic do you want to hear about?: A bilingual talking robot using english and japanese wikipedias. In: COLING (2016)
43. Winata, G.I., Cahyawijaya, S., Lin, Z., Liu, Z., Xu, P., Fung, P.: Meta-transfer learning for code-switched speech recognition. In: Proceedings of ACL, pp. 3770–3776 (2020)