Switching Tongues, Sharing Hearts: Identifying the Relationship between Empathy and Code-switching in Speech

Debasmita Bhattacharya[1*], *Eleanor Lin*[1*], *Run Chen*[1], *Julia Hirschberg*[1]

[1]Department of Computer Science, Columbia University, USA

debasmita.b@cs.columbia.edu, e.lin2@columbia.edu, runchen@cs.columbia.edu, julia@cs.columbia.edu

Abstract

Among the many multilingual speakers of the world, code-switching (CSW) is a common linguistic phenomenon. Prior sociolinguistic work has shown that factors such as expressing group identity and solidarity, performing affective function, and reflecting shared experiences are related to CSW prevalence in multilingual speech. We build on prior studies by asking: is the expression of empathy a motivation for CSW in speech? To begin to answer this question, we examine several multilingual speech corpora representing diverse language families and apply recent modeling advances in the study of empathetic monolingual speech. We find a generally stronger positive relationship of spoken CSW with the lexical correlates of empathy than with acoustic-prosodic ones, which holds across three language pairs. Our work is a first step toward establishing a motivation for CSW that has thus far mainly been studied qualitatively.

Index Terms: code-switching, empathy, speech analysis, computational paralinguistics

1. Introduction

At least half of the world is estimated to be multilingual [1]. Among multilinguals, code-switching (CSW), in which a speaker alternates between language varieties [2], is commonly observed [3, 4]. Prior work has shown that various psycho- and sociolinguistic factors influence the prevalence of CSW in conversation, including speaker competence, linguistic context, conversation topic, a speaker's emotional state, and listener identity (e.g. [5, 6, 7, 8]). While the influence of some of these has been demonstrated by statistical or computational analysis, other potential influences on CSW are yet to be investigated or confirmed through quantitative means. One such paralinguistic aspect of conversation is the demonstration of *empathy*, the ability to understand other people's feelings as if one were having them oneself [9] and respond accordingly. Given that expressing group identity and solidarity, performing affective function, and "reflect[ing] shared experiences" have all been attested as motivations for CSW, we might expect CSW to be related to empathy as well [10, 11]. In particular, [12] found that people with high cognitive empathy had a greater self-reported frequency of CSW with friends, and suggested that more empathetic individuals might code-switch more "to reflect the pattern of the interlocutor." However, self-reported frequency of CSW may not accurately reflect speakers' true frequencies of CSW. Thus, the claim that CSW is motivated by empathy has yet to be substantiated in a quantitative manner.

We take the first step toward filling this gap by studying the extent to which empathy is manifested in code-switched speech, in order to determine whether there is a relationship between the two communication phenomena. Given recent advances in measuring empathy in spoken language (e.g. [13]), we know it is now possible to measure whether or not speech displays empathetic style from the acoustic-prosodic and lexical correlates of empathy. Based on this, our work explores whether current metrics of empathy align with the incidence of CSW in speech across several language pairs. To address if there is indeed a relationship between empathy and spoken CSW, we examine conversational Spanish-English from the Bangor Miami corpus [14], Mandarin-English from the SEAME corpus [15], and Hindi-English from the MaSaC corpus [16]. We find evidence of a significant positive association between empathetic speech and code-switched speech across all three language-pairs in the corpora we investigate, and that this association is stronger for the lexical correlates of empathy than the acoustic-prosodic ones. Overall, the odds that a code-switched utterance is empathetic are higher than the odds for a monolingual utterance. Our results imply that expressing empathy may indeed be a motivation for multilingual speakers to code-switch, though further work is required to obtain a definitive *causal* answer. Our main contribution is the study of empathetic speech in the previously unexamined *multilingual, code-switched* setting.

2. Related work

Prior work closely related to empathy has produced engagement and rapport in textual dialogue systems and multimodal avatars, focusing on enhancing linguistic features such as backchannels, turn-taking, expression of emotions such as happiness, and gestures and facial expressions, which indirectly increased perceived empathy of chatbots and robots [17, 18, 19, 20, 21]. Although there has been less work on quantifying empathy specifically in speech, some studies have found that empathetic speech is characterized by lower pitch and intensity [22, 23], longer pauses and utterance-final syllables [24], and lower jitter and speaking rate [13]. In addition to acoustic-prosodic features, [13] also found that relevant lexical features of empathetic speech include Linguistic Inquiry and Word Count (LIWC) emotion categories, specificity levels, and readability scores. In our work, we adopt the definition of (compassionate) empathy used by [13], which involves a listener attempting to understand and assist another person experiencing negative emotions.

While some have studied spoken empathy in languages other than English, e.g. Italian [22], Cantonese [25], and Japanese [26], all prior studies have been restricted to monolingual domains, partly due to data scarcity. To the best of our knowledge, our work is the first to quantitatively study empathy in *multilingual* speech where CSW occurs. Although much has been done to explore other paralinguistic aspects of speech influencing CSW – e.g. emotion [27], formality [28], and speaker gender [29] – little has been done to study the role of empathy in code-switched speech. Thus, our work attempts to fill gaps in both areas of speech research by addressing the following research questions: **RQ1**: Is there a relationship between CSW prevalence in speech and the lexical and/or acoustic-prosodic correlates of empathy? **RQ2**: Does the answer to RQ1 generalize across language pairs involving different language families?

3. Corpora

We examine three code-switched corpora: the Bangor Miami (BM) Spanish-English corpus of informal dyadic and multiparty conversations, the SEAME Mandarin-English corpus of informal conversations and interviews, and the MaSaC corpus of Hindi-English situational comedy (television)[1] speech [14, 15, 16]. All three comprise a mix of monolingual and code-switched speech, with transcripts of the same. BM is English-heavy, SEAME is roughly balanced between Mandarin and English words (ratio of 1.54:1), and MaSaC is primarily Hindi. BM provides word-level language identification (LID) labels, while this information can be inferred from the Simplified Chinese and English orthographies used in SEAME. For MaSaC, we obtain LID labels with Microsoft's open-source LID tool.[2] We summarize additional corpus statistics in Table 1.

4. Method

To detect empathetic utterances in code-switched speech, we use the top-performing deep learning approaches for empathy detection from [13]: a RoBERTa base model (125M parameters) and a multimodal model (130M parameters). Both models are fine-tuned to classify utterances as empathetic or non-empathetic on 830 English utterances from the corpus in [13]. The data is evenly split between classes, with 20% used for validation. We input utterance transcripts to RoBERTa[3] and fine-tune on a T4 GPU for 20 epochs,[4] selecting the model with the best validation accuracy (val. acc.: 0.771; val. F1: 0.798). The multimodal model combines RoBERTa pooled outputs of the utterance transcripts *and* the INTERSPEECH 2009 Emotion Challenge acoustic-prosodic feature set (extracted from corresponding speech with openSMILE[5]). We fine-tune this model for 60 epochs on a T4 GPU,[6] then select the best model based on validation accuracy (val acc.: 0.612; val F1: 0.619).

To enable the application of these empathy detection models to code-switched speech, we translate all utterances which are not monolingual English to English.[7] We translate monolingual Mandarin and code-switched Mandarin-English utterances in SEAME to English using M2M100[8] (which can handle mixed scripts/languages and performs well on Chinese-to-English translation, based on our spot-checks of 1% of the data), setting the source language to Mandarin. Since Hindi segments in MaSaC are Romanized in transcription, but M2M100 was trained on Hindi transcribed in Devanagari script, M2M100 performs poorly at translating the MaSaC transcripts. We instead translate monolingual Hindi and code-switched Hindi-English MaSaC utterance transcripts to English using the Google Translate web interface, checking each for faithfulness, coherence, and retention of originally English segments. BM includes human translations of all Spanish and code-switched utterances to English, so we use these translations in our experiments. We apply both empathy detection models to each utterance in each code-switched corpus, classifying it as empathetic or non-empathetic. For monolingual English utterances, the original transcript is input to the model. For non–monolingual English utterances, the English translations are input. To quantify CSW, we compute utterance-level M- and I-indices [31, 32]. Both range from 0 (monolingual) to 1 (code-switched utterance evenly mixed between languages) and measure the extent of multilingualism within code-switched utterances. The I-index additionally encodes switching frequency. We assume that our multimodal model is language-independent with respect to acoustic-prosodic features, and thus can provide reliable gold labels of empathy across the code-switched corpora. Based on work on monolingual empathy in languages other than English (see Section 2), we believe this is a fair assumption, as many of the acoustic-prosodic correlates of empathy found in English have also been found in other languages.

5. Results

We perform chi-squared tests on each code-switched corpus to determine if CSW (labeled using word-level LIDs in each corpus) and empathy (labeled using models discussed in Section 4) are related. We then calculate odds ratios on each corpus to quantify the strength of association between CSW and empathy. We also calculate Pearson correlation coefficients between values of the M- and I-indices and the probability score produced by our empathy detection models on each utterance to determine if there is a linear relationship between amount of CSW per utterance and probability that an utterance is empathetic.

5.1. Spoken CSW aligns with lexical correlates of empathy

Spanish-English CSW in BM. We begin by using fine-tuned RoBERTa to detect empathetic utterances from speech transcripts and find a significant relationship between CSW and empathy in the BM corpus ($\chi^2(1, N = 45611) = 29.06$, $p < 0.001$). Over the entire BM corpus, the odds that a Spanish-English code-switched utterance is empathetic are **1.30** times the odds for an empathetic monolingual utterance (Figure 1). We also find a significant relationship between CSW and empathy in the dyadic subset of BM conversations ($\chi^2(1, N = 39033) = 29.13$, $p < 0.001$). In this subset, the odds that a code-switched utterance is empathetic are **1.32** times those for a monolingual utterance (Figure 1). In contrast, we find a very weak negative correlation between the amount of CSW per utterance and the probability output by fine-tuned RoBERTa that an utterance is empathetic. The M-index and $p(empathetic)$ output by RoBERTa are very weakly negatively correlated ($r(44683) = -0.08$, $p < 0.001$). The I-index and $p(empathetic)$ are also very weakly negatively correlated ($r(44683) = -0.03$, $p < 0.001$). So, neither the overall amount of CSW nor the frequency with which languages alternate are correlated with the likelihood of an empathetic utterance. This pattern also holds on the dyadic subset of BM, where

[1] This genre might inflate expression of empathy compared to non-acted speech, but this effect is likely offset by MaSaC's shorter utterances, leaving less scope for affective expression compared to the longer utterances in the other corpora.

[2] https://github.com/microsoft/LID-tool

[3] https://huggingface.co/FacebookAI/roberta-base

[4] Batch size:16, learning rate:2e-5, weight decay:0.01, and ϵ:1e-8 for AdamW optimizer.

[5] https://github.com/audeering/opensmile

[6] Batch size:8, learning rate:2e-5, and ϵ:1e-8 for AdamW optimizer.

[7] We are inspired by the general approach of [30], who argue for leveraging the "mature technology" of machine translation and existing larger, high-performance English language models over building smaller native language models. From a practical standpoint, this applies to our work as there is limited empathetic multilingual data to train and evaluate new models for each of the language pairs we investigate.

[8] https://huggingface.co/facebook/m2m100_418M

Table 1: *Summary of corpus statistics for each data set.*

Corpus	Total hours of speech	Number of dialogues	Number of transcribed words	Number of speakers
Bangor Miami	35	56	242,475	84
SEAME	192	256	1,074,032	156
MaSaC	13	1,190	39,369	5

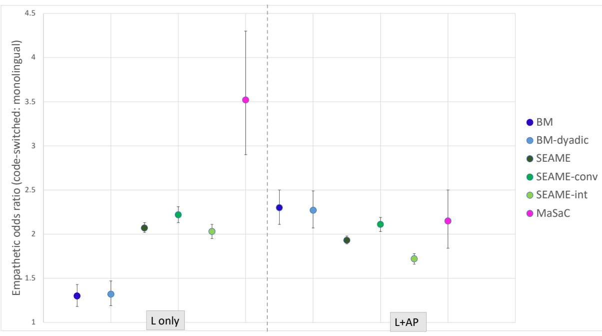

Figure 1: *Odds (with 95% CI) that a code-switched utterance is empathetic relative to the odds for a monolingual utterance in each corpus, based on only lexical (L only) and both lexical and acoustic-prosodic (L+AP) correlates of empathy.*

the correlation between the M-index and $p(empathetic)$ output by RoBERTa is very weakly negative ($r(39031) = -0.07$, $p < 0.001$), as is that between the I-index and $p(empathetic)$ ($r(39031) = -0.03$, $p < 0.001$). Overall, we find a significant, but non-linear, association between empathetic and code-switched Spanish-English speech in the BM corpus.

Mandarin-English CSW in SEAME. We also find a significant relationship between CSW and empathy in the SEAME corpus ($\chi^2(1, N = 106080) = 2678.12$, $p < 0.001$). This positive relationship holds for both the interview subset ($\chi^2(1, N = 59173) = 1300.38$, $p < 0.001$) and the conversation subset of SEAME ($\chi^2(1, N = 46907) = 1486.67$, $p < 0.001$). The odds ratio for the entire corpus is **2.07**, while that for the conversation subset is **2.22**, and that for the interview subset is **2.03** (Figure 1). So, Mandarin-English code-switched utterances in the SEAME corpus have higher odds of being empathetic than monolingual utterances, regardless of whether these utterances are spoken in conversational or interview settings. These values align with our expectation that there is greater scope for expressing empathy in conversational settings than in interviews. However, we find only a very weak positive correlation between amount of CSW per utterance and probability that an utterance is empathetic. The I-index and $p(empathetic)$ output by RoBERTa are very weakly positively correlated ($r(106078) = 0.01$, $p < 0.001$), as are the M-index and $p(empathetic)$ ($r(106078) = 0.05$, $p < 0.001$). These patterns replicate in the conversation subset of the corpus, where we find a weak positive correlation between both I-index and $p(empathetic)$ output by RoBERTa and M-index and $p(empathetic)$ ($r(46905) = 0.03$, $p < 0.001$ and $r(46905) = 0.09$, $p < 0.001$, respectively). We also find similar patterns on the interview subset of SEAME, where there is no correlation between I-index and $p(empathetic)$ output by RoBERTa ($r(59171) = 0.00$, $p = 0.72$) and a weak positive correlation between M-index and $p(empathetic)$ ($r(59171) = 0.02$, $p < 0.001$). As with Spanish-English CSW, we find a significant association between empathetic and code-switched Mandarin-English speech in the SEAME corpus, but limited evidence of this association being linear.

Hindi-English CSW in MaSaC. Finally, we find a significant relationship between CSW and empathy in the MaSaC corpus ($\chi^2(1, N = 6475) = 181.11$, $p < 0.001$). The odds that a Hindi-English code-switched utterance is empathetic are **3.52** times the odds for monolingual utterances (Figure 1). Similar to SEAME, we find a weak positive correlation between $p(empathetic)$ output by RoBERTa and M-index ($r(6351) = 0.13$, $p < 0.001$), and $p(empathetic)$ and I-index ($r(6351) = 0.07$, $p < 0.001$). These results are generally consistent with those we found on the other code-switched corpora.

In all three language pairs, we see a significant, positive, non-linear association between CSW and empathy encoded by its **lexical** correlates in speech. Inspection of code-switched utterances labeled as empathetic by our model reveals that it picks up on similar lexical features to those [13] found in empathetic language. For example, the code-switched BM utterance "*pero ahora mismo tú me estabas diciendo que te gustaba y te parecía* nice" ["but right now you were telling me that you like it and you think it's nice"] references the addressee's cognitive processes and feelings, which [13] also found to be referenced more frequently in empathetic speech.

5.2. Spoken CSW somewhat aligns with acoustic-prosodic correlates of empathy

Spanish-English CSW in BM. Applying our multimodal model to detect empathetic utterances from speech transcripts *and* openSMILE features, we again find a significant relationship between CSW and empathy in the BM corpus ($\chi^2(1, N = 45607) = 379.00$, $p < 0.001$). The odds that a Spanish-English code-switched utterance is empathetic are **2.30** times the odds for monolingual utterances (Figure 1). As with the lexical experiments, we find that this positive relationship is also present in the dyadic subset of BM ($\chi^2(1, N = 39029) = 318.30$, $p < 0.001$), where the corresponding odds ratio is **2.27**. However, we now find a very weak positive correlation between the amount of CSW per utterance and the probability output by the multimodal model that an utterance is empathetic. Both the correlations between I-index and $p(empathetic)$ output by the multimodal model and the M-index and $p(empathetic)$ are very weakly positive ($r(44680) = 0.03$, $p < 0.001$ and $r(44680) = 0.03$, $p < 0.001$, respectively). This pattern holds true for the dyadic subset of the BM corpus as well, where the I-index and $p(empathetic)$ output by the multimodal model are very weakly positively correlated ($r(38249) = 0.03$, $p < 0.001$), as are the M-index and $p(empathetic)$ ($r(38249) = 0.03$, $p < 0.001$). Overall, the inclusion of acoustic-prosodic features in our analysis of BM reveals an even stronger association between empathy and CSW than we found in our lexical-only experiments. The odds that an utterance is empathetic, given that it is code-switched, are higher when we incorporate both the lexical and acoustic-prosodic correlates of empathy. This suggests that in addition to a significant non-linear relationship between the lexical correlates of empathy and CSW, there may be a similar relationship between the acoustic-prosodic correlates of empathy and CSW.[9]

[9]The difference in odds ratios may also stem from the performance gap between our RoBERTa and multimodal models. Future work

Mandarin-English CSW in SEAME. We next apply the multimodal model on the SEAME corpus, finding a significant relationship between empathy and CSW in Mandarin-English speech ($\chi^2(1, N = 105972) = 2706.48$, $p < 0.001$). As with the lexical experiments, this significant relationship holds for both the interview subset ($\chi^2(1, N = 59066) = 974.63$, $p < 0.001$) and the conversation subset ($\chi^2(1, N = 46906) = 1529.41$, $p < 0.001$). The odds that a code-switched utterance in SEAME is empathetic are **1.93** times the odds for monolingual utterances, while the odds ratio for code-switched interview utterances is **1.72**, and the odds ratio for code-switched conversational utterances is **2.11** (Figure 1). As with our lexical experiments, we find a very weak positive correlation between the amount of CSW per utterance and the probability output by the multimodal model that an utterance is empathetic. The I-index and $p(empathetic)$ output by the multimodal model are very weakly positively correlated ($r(104790) = 0.05$, $p < 0.001$). Similarly, the M-index and $p(empathetic)$ output by the multimodal model are very weakly positively correlated ($r(104790) = 0.07$, $p < 0.001$). This pattern of weak positive correlation also holds for both the conversation and interview subsets, for correlations between $p(empathetic)$ and both I-index ($r(46471) = 0.06$, $p < 0.001$ and $r(58317) = 0.04$, $p < 0.001$, respectively) and M-index ($r(46471) = 0.09$, $p < 0.001$ and $r(58317) = 0.06$, $p < 0.001$, respectively). Although these results mirror those of our lexical experiments, including acoustic-prosodic features in addition to lexical ones slightly decreases the strength of the non-linear association between empathy and Mandarin-English CSW, as shown by the decrease in odds ratios across all subsets of the SEAME corpus. This is in contrast to our findings on BM, suggesting that, although there is a relationship between the acoustic-prosodic correlates of empathy and CSW in SEAME, this relationship may be less strongly positive than that between the lexical correlates of empathy and CSW.

Hindi-English CSW in MaSaC. We finally apply our multimodal model on the MaSaC corpus and find a significant relationship between Hindi-English CSW and empathy ($\chi^2(1, N = 6475) = 102.85$, $p < 0.001$). The odds that a code-switched utterance is empathetic are **2.15** times the odds for monolingual utterances (Figure 1). However, unlike the other corpora, there is no correlation between I-index and $p(empathetic)$ ($r(6474) = 0.01$, $p = 0.603$), nor between M-index and $p(empathetic)$ ($r(6474) = 0.00$, $p = 0.942$). As with SEAME, the MaSaC results indicate the relationship between acoustic-prosodic correlates of empathy and CSW in speech may not be as strongly positive as that between the lexical correlates of empathy and CSW, as this odds ratio is lower than that from the previous lexical-only experiments. Further interpretation is difficult as the effect of including acoustic-prosodic features of empathetic speech, as well as lexical ones, may be more complex than a simple additive relationship.

Our results confirm a non-linear relationship between empathy and amount of CSW across the three language pairs. There is a relationship between the **acoustic-prosodic** correlates of empathy and CSW, but this is weaker than that with the lexical correlates for Mandarin- and Hindi-English CSW. Inspection of code-switched empathetic utterances again reveals aspects of empathetic language found in [13]; the MaSaC utterance "of course *mujhe pata hai lekin ek baar tum kaho na* you know 24th February *ke baare mein jab tum kehti ho na to aur accha lagta hai*" ["Of course I <u>know</u>... But, once you say it... You know about 24th February, when you say it, it <u>feels</u> better..."] references <u>cognitive processes</u> and has a jitter of 2.1%, which is lower than average for the corpus and thus likely to characterize empathy. While the association between the lexical correlates of empathy and CSW is strongest in Hindi-English, then Mandarin-English, and finally Spanish-English, this ranking is reversed for the strength of association between the acoustic-prosodic correlates and CSW. This may reflect differences in relative contributions of lexical and acoustic-prosodic features to expressing empathy in the three language pairs.

should develop better multilingual, multimodal empathy models to better understand the role of acoustic-prosodic features.

6. Conclusion

We examine the relationship between empathy and CSW in speech. We find (1) a positive non-linear association between CSW and the lexical correlates of empathy, and a weaker but still positive non-linear association between CSW and the acoustic-prosodic correlates of empathy; and (2) patterns from lexical features hold across language pairs while those from acoustic-prosodic features are slightly less consistent. We conclude that current metrics of empathy in speech generally align with the incidence of CSW, with the odds that code-switched speech is empathetic about twice the odds for monolingual speech, though the relationship between acoustic-prosodic empathetic features and CSW may be more subtle than expected. While on the lexical side, the relationship between empathy and CSW generalizes across language pairs and corpus subsets, this is less clear on the acoustic-prosodic side. Further work should determine whether a set of acoustic-prosodic features exists that clearly encodes a positive relationship between empathy and CSW. We expect this challenging task to become more tractable with improvements in the measurement and modeling of both monolingual and multilingual empathetic speech.

Our findings serve as a first step toward validating prior qualitative work, e.g. [10, 11], and answering the question of whether the expression of empathy is a motivation for CSW in speech. We hope to motivate future work to further validate whether there is a causal relationship between empathy and CSW. We also hope that this work and its future extensions will contribute to innovation in interactive voice technology through the inclusion of appropriate paralinguistic features in voice assistant responses to code-switched input. This could lead to the specific benefits associated with enhancing empathy-producing aspects of voice assistant speech, as people have been shown to not only prefer interacting with empathetic dialogue systems, but also have greater trust in such systems, even when they are aware they are not speaking to another human [33]. Further, we hope that our findings inform future work on conversational agents which can code-switch naturally, an important consideration given that multilingual users prefer such agents [34].

7. Limitations

While we recognize that using an English-trained model to detect empathy in other languages is not ideal, we believe that this is still the best available current option, given the low-resource setting and lack of alternative models for this task. We also acknowledge that exploring alternative methods to using a classifier and/or translation as a proxy for ground truth would be a fruitful future direction. Our work is a first step toward understanding empathy in a new multilingual setting given a current lack of appropriately labeled, accessible multilingual data.

8. Acknowledgments

*DB and EL are equal contributors to this work and are designated as co-first authors. This work was supported in part by the National Science Foundation under Grant IIS 2327564.

9. References

[1] F. Grosjean and P. Li, *The Psycholinguistics of Bilingualism*. Wiley-Blackwell, 2013.

[2] S. Poplack, "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching," *Linguistics*, vol. 18, pp. 581–618, 01 1980.

[3] J. M. Meisel, "Code-switching in young bilingual children: The acquisition of grammatical constraints," *Studies in Second Language Acquisition*, vol. 16, no. 4, pp. 413–439, 1994. [Online]. Available: http://www.jstor.org/stable/44487780

[4] P. Muysken, *Bilingual speech: a typology of code-mixing*. Cambridge University Press, 2000.

[5] A. Bell, "Language style as audience design," *Language in Society*, vol. 13, no. 2, pp. 145–204, 1984. [Online]. Available: http://www.jstor.org/stable/4167516

[6] M. Broersma, "Triggered codeswitching between cognate languages," *Bilingualism: Language and Cognition*, vol. 12, no. 4, pp. 447–462, 2009.

[7] S. Dornic, *The Bilingual's Performance: Language Dominance, Stress, and Individual Differences*. Boston, MA: Springer US, 1978, pp. 259–271. [Online]. Available: https://doi.org/10.1007/978-1-4615-9077-4_23

[8] P. Gardner-Chloros, *Code-switching*. Cambridge University Press, 2009.

[9] R. Baumeister and K. Vohs, *Encyclopedia of Social Psychology*. SAGE Publications, 2007, no. 1. [Online]. Available: https://books.google.com.hk/books?id=Mqk5DQAAQBAJ

[10] C. Hoffmann, *Introduction to Bilingualism*, ser. Longman Linguistics Library. Taylor & Francis, 2014. [Online]. Available: https://books.google.com.hk/books?id=NFugBAAAQBAJ

[11] J. Holmes, *An Introduction to Sociolinguistics*, ser. Learning about language. Pearson Longman, 2008. [Online]. Available: https://books.google.com.hk/books?id=UOQnUDuEKcEC

[12] J.-M. Dewaele and W. Li, "Intra-and inter-individual variation in self-reported code-switching patterns of adult multilinguals," *International Journal of Multilingualism*, vol. 11, no. 2, pp. 225–246, 2014.

[13] R. Chen, H. Chen, A. Kulkarni, E. Lin, L. Pang, D. Tadimeti, J. Shin, and J. Hirschberg, "Detecting empathy in speech," submitted to Interspeech 2024.

[14] M. Deuchar, "Miami corpus: Preliminary documentation - bangortalk," 2011. [Online]. Available: http://bangortalk.org.uk/docs/Miami_doc.pdf

[15] D.-C. Lyu, T.-P. Tan, E. Chng, and H. Li, "Mandarin–English code-switching speech corpus in South-East Asia: SEAME," in *Language Resources and Evaluation*, vol. 49, 01 2010, pp. 1986–1989.

[16] M. Bedi, S. Kumar, M. Akhtar, and T. Chakraborty, "Multi-modal sarcasm detection and humor classification in code-mixed conversations," *IEEE Transactions on Affective Computing*, 5 2023.

[17] L. Huang, L.-P. Morency, and J. Gratch, "Virtual rapport 2.0," in *Lecture Notes in Computer Science*, 09 2011, pp. 68–79.

[18] R. Niewiadomski, E. Bevacqua, M. Mancini, and C. Pelachaud, "Greta: An interactive expressive eca system," in *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, vol. 2, 01 2009, pp. 1399–1400.

[19] M. Ochs, C. Pelachaud, and D. Sadek, "An empathic virtual dialog agent to improve human-machine interaction," in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*, Richland, SC, 2008, p. 89–96.

[20] R. Zhao, T. Sinha, A. W. Black, and J. Cassell, "Socially-aware virtual agents: Automatically assessing dyadic rapport from temporal patterns of behavior," in *International Conference on Intelligent Virtual Agents*, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:13786126

[21] A. Ghandeharioun, D. McDuff, M. Czerwinski, and K. Rowan, "Emma: An emotion-aware wellbeing chatbot," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2019, pp. 1–7.

[22] F. Alam, M. Danieli, and G. Riccardi, "Annotating and modeling empathy in spoken conversations," *Computer Speech & Language*, vol. 50, pp. 40–61, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S088523081730133X

[23] B. Xiao, D. Bone, M. V. Segbroeck, Z. E. Imel, D. C. Atkins, P. G. Georgiou, and S. S. Narayanan, "Modeling therapist empathy through prosody in drug addiction counseling," in *Proc. Interspeech 2014*, 2014, pp. 213–217.

[24] J. H. N. Lee, D. Tao, H. Chui, T. Lee, S. Luk, N. W. T. Lee, and K. Fung, "Durational Patterning at Discourse Boundaries in Relation to Therapist Empathy in Psychotherapy," in *Proc. Interspeech 2022*, 2022, pp. 5248–5252.

[25] D. Tao, T. Lee, H. Chui, and S. Luk, "Characterizing Therapist's Speaking Style in Relation to Empathy in Psychotherapy," in *Proc. Interspeech 2022*, 2022, pp. 2003–2007.

[26] Y. Saito, Y. Nishimura, S. Takamichi, K. Tachibana, and H. Saruwatari, "STUDIES: Corpus of Japanese Empathetic Dialogue Speech Towards Friendly Voice Agent," in *Proc. Interspeech 2022*, 2022, pp. 5155–5159.

[27] A. V. A. Ferreira, "Code-switching and emotions display in Spanish/Galician bilingual conversation," *Text & Talk*, vol. 37, pp. 47 – 69, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:199060969

[28] D. Bhattacharya, J. Chi, J. Hirschberg, and P. Bell, "Capturing Formality in Speech Across Domains and Languages," in *Proc. INTERSPEECH 2023*, 2023, pp. 1030–1034.

[29] K. A. FINNIS, "Variation within a Greek-Cypriot community of practice in London: Code-switching, gender, and identity," *Language in Society*, vol. 43, no. 3, pp. 287–310, 2014. [Online]. Available: http://www.jstor.org/stable/43903872

[30] T. Isbister, F. Carlsson, and M. Sahlgren, "Should we stop training more monolingual models, and simply use machine translation instead?" in *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*. Reykjavik, Iceland (Online): Linköping University Electronic Press, Sweden, May 31–2 Jun. 2021, pp. 385–390.

[31] G. A. Guzman, J. Serigos, B. E. Bullock, and A. J. Toribio, "Simple tools for exploring variation in code-switching for linguists," in *Proceedings of the Second Workshop on Computational Approaches to Code Switching*, M. Diab, P. Fung, M. Ghoneim, J. Hirschberg, and T. Solorio, Eds. Austin, Texas: Association for Computational Linguistics, Nov. 2016, pp. 12–20. [Online]. Available: https://aclanthology.org/W16-5802

[32] R. Barnett, E. Codó, E. Eppler, M. Forcadell, P. Gardner-Chloros, R. van Hout, M. Moyer, M. C. Torras, M. T. Turell, M. Sebba, M. Starren, and S. Wensing, "Automatic analysis," *International Journal of Bilingualism*, vol. 4, no. 2, pp. 198–208, 2000. [Online]. Available: https://doi.org/10.1177/13670069000040020107

[33] G. M. Lucas, J. Boberg, D. Traum, R. Artstein, J. Gratch, A. Gainer, E. Johnson, A. Leuski, and M. Nakano, "Getting to know each other: The role of social dialogue in recovery from errors in social robots," in *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2018, pp. 344–351.

[34] A. Bawa, P. Khadpe, P. Joshi, K. Bali, and M. Choudhury, "Do multilingual users prefer chat-bots that code-mix? let's nudge and find out!" *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW1, pp. 1–23, 2020.