

Although preliminary, our results are encouraging for recommendation systems and policy making. Data-driven approaches are especially necessary to assist policy makers in making well-informed decisions for immigrant communities. Future studies are welcome to analyze the communication patterns in other immigrant communities around the world to make comparisons and learn from the practices of each other.

7. ACKNOWLEDGMENTS

The first author was supported by the National Science Foundation Grant No. IIS-1247489. The second author was supported by Digital Humanities Research grant from Tilburg University (Netherlands). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding parties.

8. REFERENCES

- [1] J. Androutsopoulos. *The Multilingual Internet*, chapter Language choice and code-switching in German-based diasporic web forums, pages 340–361. Oxford University Press, 2007.
- [2] P. Auer. A conversation analytic approach to code-switching and transfer. *Codeswitching: Anthropological and sociolinguistic perspectives*, 48:187–213, 1988.
- [3] P. Auer. *Code-switching in conversation: Language, interaction and identity*. Routledge, 2013.
- [4] P. Auer and L. Wei. *Handbook of multilingualism and multilingual communication.*, chapter Introduction: Multilingualism as a problem? Monolingualism as a problem, pages 1–14. Berlin: Mouton de Gruyter, 2007.
- [5] Centraal Bureau voor de Statistiek. Bevolking, generatie, geslacht, leeftijd en herkomstgroepering. 2013., 2013.
- [6] R. Cotterell, A. Renduchintala, N. Saphra, and C. Callison-Burch. An algerian arabic-french code-switched corpus. In *LREC*, 2014.
- [7] A. S. Doğruöz and A. Backus. Postverbal elements in immigrant Turkish: Evidence of change? *International Journal of Bilingualism*, 11(2):185–220, 2007.
- [8] A. S. Doğruöz and A. Backus. Innovative constructions in Dutch Turkish: An assessment of ongoing contact-induced change. *Bilingualism: Language and Cognition*, 12(01):41–63, 2009.
- [9] A. S. Doğruöz and S. T. Gries. Spread of on-going changes in an immigrant language Turkish in the Netherlands. *Review of Cognitive Linguistics*, 10(2), 2012.
- [10] I. Eleta and J. Golbeck. Bridging languages in social networks: How multilingual users of twitter connect language communities? *Proceedings of the American Society for Information Science and Technology*, 49(1):1–4, 2012.
- [11] P. Gardner-Chloros and M. Edwards. Assumptions behind grammatical approaches to code-switching: When the blueprint is a red herring. *Transactions of the Philological Society*, 102(1):103–129, 2004.
- [12] S. A. Hale. Global connectivity and multilinguals in the twitter network. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 833–842. ACM, 2014.
- [13] V. Hinnenkamp. Deutsch, Doyc or Doitsch? Chatters as languagers—The case of a German–Turkish chat room. *International Journal of Multilingualism*, 5(3):253–275, 2008.
- [14] J. W. Y. Ho. Code-mixing: Linguistic form and socio-cultural meaning. *The International Journal of Language Society and Culture*, 21, 2007.
- [15] S. Kim, I. Weber, L. Wei, and A. Oh. Sociolinguistic analysis of twitter in multilingual societies. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media, HT*, volume 14, pages 243–248, 2014.
- [16] B. King and S. P. Abney. Labeling the languages of words in mixed-language documents using weakly supervised methods. In *HLT-NAACL*, pages 1110–1119, 2013.
- [17] T. Kolda and B. Bader. Tensor decompositions and applications. *SIAM review*, 51(3), 2009.
- [18] M. Lui, J. H. Lau, and T. Baldwin. Automatic detection and language identification of multilingual documents. *Transactions of the Association for Computational Linguistics*, 2:27–40, 2014.
- [19] C. Myers-Scotton. *Social motivations for codeswitching: Evidence from Africa*. Oxford University Press, 1995.
- [20] D. Nguyen and A. S. Doğruöz. Word level language identification in online multilingual communication. In *Proceedings of EMNLP 2013*, 2013.
- [21] E. E. Papalexakis, C. Faloutsos, and N. D. Sidiropoulos. Parcube: Sparse parallelizable tensor decompositions. In *Machine Learning and Knowledge Discovery in Databases*, pages 521–536. Springer, 2012.
- [22] S. Poplack. Sometimes i’ll start a sentence in spanish y termino en espanol: toward a typology of code-switching I. *Linguistics*, 18(7-8):581–618, 1980.
- [23] S. Poplack, D. Sankoff, and C. Miller. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26(1):47–104, 1988.
- [24] S. Romaine. Bilingualism (2nd edn). *Malden, MA: Blackwell Publishers*, 1995.
- [25] D. Tang, T. Chou, N. Drucker, A. Robertson, W. C. Smith, and J. T. Hancock. A tale of two languages: strategic self-disclosure via language selection on facebook. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pages 387–390. ACM, 2011.
- [26] S. G. Thomason and T. Kaufman. *Language contact*. Edinburgh University Press Edinburgh, 2001.
- [27] L. Tsaliki. Globalization and hybridity: the construction of greekness on the internet. *The Media of Diaspora*, Routledge, London, 2003.
- [28] L. Wei. *Codeswitching in conversation: Language, interaction and identity*, chapter The ‘why’ and ‘how’ questions in the analysis of conversational codeswitching, pages 156–176. Routledge, 1998.