

Teen talk isn't simple: Register-dependent complexity in Turkish adolescent speech

S. Özmemiş¹, G. Thiberge¹, et al.

¹ LLF (Université Paris Cité/CNRS), Paris

Background. Adolescents are often described as drivers of language change [1] and as particularly proficient in informal peer interaction [2]. However, success in institutional contexts (e.g., school, legal settings) requires mastery of more formal registers. Despite this, little is known about how adolescents adapt their linguistic production across registers even though this question is of crucial importance both in real-life and for psycholinguistic experiments. Moreover, demographic variables such as socio-economic background are known to affect language development and attainment [3,4], yet their precise role in shaping register-dependent variation remains unclear. Drawing on a corpus of spoken productions collected through a semi-controlled production task (RUEG: Research Unit Emerging Grammars [5]), we are developing a systematic approach to data to disentangle how multiple dimensions of linguistic complexity vary across formal and informal contexts in adolescent speech, and whether these patterns are related to SES.

Methods In a first step of a bigger multilingual project, we analyze the Turkish section of the RUEG corpus, which contains naturalistic data from 34 monolingual Turkish-speaking adolescents. In the spoken production task, participants described a traffic accident video in two conditions: (i) a formal voice message to the police (witness report), and (ii) an informal voice message to a friend (WhatsApp). Through a partly-automatized chain of reannotation and analysis over the original data from the corpus, we operationalize complexity across 5 domains: 1) Lexical complexity through Ure's lexical density [6] and normalized type/token ratio; 2) Morphological complexity through an adapted inflectional diversity index (MCI [7]; 3) Syntactic complexity through a normalized score based on the number of complex sentences per t-unit; 4) Discourse complexity through coherence relations [8] and pronoun/NP ratio [9]; 5) Information density through Content Unit/Unit Ratio [10].

Results Analyses are ongoing. The adolescents showed variation in their productions, which range from 20 to 225 words. A single speaker example (page 3) shows the normalized syntactic complexity score is higher for the formal (1.85) than for the informal (0.42) production, although more information is conveyed (i.e. higher information density) in the informal compared to the formal production. This example, and other preliminary observations, suggest that adolescents systematically adapt their linguistic production across registers, highlighting their ability to flexibly allocate linguistic resources depending on communicative context. Ongoing analyses will assess the robustness of these effects across speakers and complexity domains, and characterize whether and how SES, but also other speaker-specific variables such as age and gender, predict individual differences in register adaptation.

References

- [1] Eckert, Penelope. 1997. Age as a sociolinguistic variable. In Florian Coulmas (ed.), *The handbook of sociolinguistics*, 151–167. Oxford: Blackwell. DOI: 10.1002/9781405166256.ch9
- [2] Eckert, P. 2003. Language and adolescent peer groups. *Journal of language and social psychology* 22.1 : 112-118
- [3] Pace, A. Luo, R. Hirsh-Pasek, K., Michnick Golinkoff, R. 2017. Identifying Pathways Between Socioeconomic Status and Language Development. *Annual Review Linguistics*. 3:285-308. DOI: 10.1146/annurev-linguistics-011516-034226.
- [4] Stamatis, I., & Stavrakaki, S. 2026. Socioeconomic status and language development in children: a scoping review including cross-cultural and methodological viewpoints. *International Journal of Multilingualism*, 1–29. DOI: 10.1080/14790718.2026.2648812
- [5] Lüdeling, A., Alexiadou, A., Allen, S., Bunk, O., Gagarina, N., Grigoriadou, S., Hartz, R. G., Iefremenko, K., Jahns, E., Katsika, K. Keller, M. Klotz, M., Krause, T., Labrenz, A., Martynova, M., Özsoy, O. Pashkova, T., Pohle, M., Purkarthofer, J., Rizou, V., Schroeder, C., Shadrova, A. Szucsich, L., Tracy, R., Tsehay, W., Wiese, H., Zerbian, S., Zuban, Y., Zürn, N. 2024. RUEG Corpus (Version 1.0) [Data set]. Zenodo. DOI: [10.5281/zenodo.11234583](https://doi.org/10.5281/zenodo.11234583)
- [6] Ure J. 1971. Lexical density and register differentiation in Perren G., Trim J. L. M. (eds.), *Applications of linguistics*, 443–452, Cambridge University Press.
- [7] Brezina & Pallotti 2016. Morphological complexity in written ESL texts. *Second Language Research*.
- [8] Sanders, T., Spooren, W., & Noordman, L. 1993. Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*, 4, 93–133.
- [9] Petti, U., Baker, S., Korhonen, A., & Robin, J. 2023. The Generalizability of Longitudinal Changes in Speech Before Alzheimer's Disease Diagnosis. *Journal of Alzheimer's disease: JAD*, 92(2), 547–564. DOI: 10.3233/JAD-220847
- [10] Ravid, D., & Berman, R. A. 2003. On modeling information density in narrative texts. *Determination of information and tenor in texts: Multidisciplinary approaches to discourse*, 17-36.

speaker TUmo56FT (**formal**)

TR: memur bey burada bir kaza yaşandı. e kadın ve köpeği yolda yürürken karşıdan gelen bir adam vardı. köpek ona doğru koşmaya başladı. e yoldan geçen araba da bunu görüp frene bastı. arkadaki gelen araba bu olayı görmediği için hızlı bir şekilde geldi. ve bir kaza yaşandı. arabaya arkadan vurdu

RENDERED IN ENGLISH: *officer there was an accident here. while a woman and her dog were walking down the road there was a man approaching from the opposite direction. The dog started running towards him. the passing car saw this and braked. because the car coming behind didn't see this it came speeding up. and an accident happened. (It) hit the car from behind.*

Raw Score (13) / T-units (7) = 1.85 Normalized Syntax Complexity Score

speaker TUmo56FT (**informal**)

TR: yolda yürürken bir kaza oldu. gerçekten çok korktum. yani bana bir şey olacak diye çok korktum açıkcası. e iki aile vardı. karşıdaki ailenin de bir köpeği var. senin köpeğine benziyor. e diğer bir ailenin de çocuğu var. sanırım o da babası elinde topla oynuyor. topu karşıya yola kaçırdı. oradan da bir araba geldi. e o topu gören köpek de yola atlayınca araba da ani fren yaptı. onun arkasındaki de o arabaya çarptı. öyle işte yani çok korktum bu kadar.

RENDERED IN ENGLISH: *while walking down the street there was an accident. i was really scared. I mean I was terrified that something would happen to me. there were two families. and the family across (the street) has a dog. it looks like your dog. and the other family has a child. i think he is his father playing with the ball. he dropped the ball from his hands it rolled into the road. a car came along. when the dog that saw the ball jumped into the road the car braked suddenly. the car behind it then hit the car. so that's all, I was really scared.*

Raw Score (6) / T-units (14) = 0.42 Normalized Syntax Complexity Score