Chapter 6

Null subjects in heritage Greek, Russian and Turkish

Onur Özsoy[a,b], Vasiliki Rizou[b], Maria Martynova[b], Natalia Gagarina[a,b], Luka Szucsich[b] & Artemis Alexiadou[a,b]

[a]Leibniz-Centre General Linguistics [b]Humboldt-Universität zu Berlin

Greek, Russian, and Turkish represent three different types of null subject languages: Greek is characterized as a strict pro-drop language; Russian is claimed to be either a partial pro-drop, or a non-pro-drop language with abundant subject ellipsis; Turkish is classified as a pro-drop or topic pro-drop language in which overt subjects are required in unspecified contexts. This chapter reports results from corpus-based studies exploring the question of whether subjects in heritage varieties of these three languages in contact with English and German pattern alike regarding the realization of overt subjects. Moreover, the chapter examines whether heritage speakers of these three languages apply topic drop more widely in one of their majority languages, namely German. We consider different possibilities for our observations reaching from language contact effects, to language internal dynamics and most importantly a typological continuum within the different pro-drop types regarding the stability and dynamicity in their pro-drop systems.

1 Introduction

The main aim of this chapter is to examine the phenomenon of null subjects in Greek, Russian, and Turkish majority and heritage language varieties using a comparative approach. The study analyzes data from both monolingual and bilingual populations, where the latter are heritage speakers. Heritage speakers are typically individuals who use their heritage language(s), learned during early childhood in their families, in addition to the majority language(s) spoken in

the host community (cf. among many others Valdés 2005, Rothman 2009, Benmamoun et al. 2013, Guijarro-Fuentes & Schmitz 2015, Polinsky 2015, Montrul 2015). Heritage speakers are considered bilinguals belonging to a nativeness continuum as native speakers of their heritage language, but the proficiency in the languages they speak can vary throughout their lifetime (Benmamoun et al. 2013, Polinsky 2018, Wiese et al. 2022). In this chapter, we adopt this perspective and we do not compare the heritage data to the monolingual data as a measure of accuracy. Rather, we view the heritage and monolingual populations as belonging to distinct language varieties in their own right, in line with Rothman et al.'s (2023) perspective.

Additionally, we aim to gain an understanding of subject realization in the heritage languages Greek, Russian, and Turkish, which differ typologically, and investigate whether they are influenced by contact with the majority languages English and German. The latter are both characterized as non-pro-drop languages with German allowing for the omission of subjects in topic positions (Oppermann 2021, Schäfer 2021, Trutkowski 2016). Homeland varieties of Greek, Russian and Turkish all allow subject omission to varying degrees. Heritage Greek, Russian, and Turkish have been argued to exhibit similar tendencies regarding their pro-drop properties. Notably, overtly realized pronominal subjects have been found to be more frequent in heritage varieties, particularly those in contact with Germanic languages, than in monolingual ones. This has been discussed in various studies such as Argyri & Sorace (2007) and Tsimpli et al. (2004) for heritage Greek; Dubinina & Polinsky (2013), Gagarina (2012), and Ivanova-Sullivan (2014, 2015) for heritage Russian; and Haznedar (2010) and Koban Koç (2016) for heritage Turkish (but see Azar et al. 2020 for results showing that heritage and monolingual speakers of Turkish omit subjects to a similar degree).

While previous studies have mostly focused on the heritage language of these speakers, there are very few studies that investigate the majority language too. This chapter adds to these studies and discusses whether HSs of Greek, Russian, and Turkish in Germany apply topic drop more widely in their majority language German. An expansion of topic drop would represent a creative extension of German grammar which points to a reorganization at the pragmatics-syntax interface.

The investigation of subject realization in heritage Greek, Russian, and Turkish in the US and Germany, and of topic drop in majority German is a promising area of study since it relates to the interface between internal linguistic domains (core grammatical system) and external linguistic domains (information structure and discourse organization). To achieve the above-mentioned objectives, the study uses a multi-factorial analysis of semi-naturalistic and ecologically more

valid data from monolingual and heritage speakers, taking into account differences in formality and mode.

This study also extends previous research in directly addressing the effect on subject realization of animacy, mode, and formality – three factors which are meaningful to subject realization but are not often considered in this field. Some studies investigating languages unrelated to the languages of our study point to animacy playing an important role in subject realization. Animacy is an inherent feature of nominal referents (Corbett 1979, Comrie 1989). Numerous scholars have classified animacy on a hierarchy (Silverstein & Dixon 1986), placing it on a continuum with animate human discourse participants on the one end and inanimate and abstract referents on the other (cited from Corbett 2000: 90), as illustrated below:

(1)   speaker (first person pronouns) > addressee (second person pronouns) > third person > kin > human > animate > inanimate

Li et al. (2012) and Pu (1997) report for the radical pro-drop language Mandarin that animate subjects and objects are more prone to be null subjects than inanimate ones. Similarly, for some creole languages like Tamambo, animacy plays a significant role in subject deletion with animate subjects being more likely to be omitted (Meyerhoff 2009). Finally, it was reported that subject omission in the Austronesian languages Vera'a and Teop is affected by animacy (Vollmer 2019).

Another uncharted territory that this study explores regarding null subjects is their relation to different communicative situations and speakers' repertoires. Previous studies have shown that certain text types such as short messages on social media or on the phone facilitate subject omission and ellipsis more generally. For example, Frick (2017) shows that in German text messages, there are more null pronouns than realized ones. This is atypical for German generally, but it is typical for this specific text genre. We are not aware of a cross-modality and cross-formality study of subject omission, which makes the investigations in this chapter pioneering and exploratory.

However, there is related previous work from the domain of nominal reference which refers to demonstrative-marked noun phrases in Greek, Russian and Turkish mono- and bilingual speakers. Martynova et al. (2024) show that different communicative settings impact the use of definite noun phrases in different ways. For instance, Greek speakers use more double definite structures in informal settings, Turkish speakers use demonstratives more in spoken than in written mode, and Russian speakers use demonstratives more frequently in informal and oral communication settings. These findings imply that effects of communicative situations could be language-specific in the domain of null subjects too.

The chapter is structured as follows: Section 2 provides the theoretical background on the different types of null subject languages this chapter deals with. In this section, we also summarize results of previous research that discusses animacy, mode and formality as possible factors influencing when subjects are omitted. Section 3 introduces the research questions and the corresponding predictions for every language in our study derived from results in the literature. Section 4 describes our methodology, including the design, the participant sample, the corpus annotation scheme and the statistical analysis. Section 5 provides the results reported by the models and grouped according to the different factors accounted for in the analysis, namely typology (i.e., different pro-drop types), animacy, mode and formality. Finally, Section 6 presents a discussion of our results.

2 Null subjects

Languages have been grouped into different types regarding their ability to omit subjects. Leaving aside languages with so-called radical pro-drop or expletive subjects, most authors distinguish between languages readily allowing pro-drop (strict pro-drop) and non-pro-drop languages which by and large disallow the omission of subjects. The former are claimed to require rich verbal morphology (Perlmutter 1971, Holmberg 2005, Roberts & Holmberg 2010). A third class of languages are referred to as partial pro-drop; subject omission in these languages is more limited than in strict pro-drop languages, and may be restricted by (morpho)syntactic, information-structural, or lexical constraints (Holmberg 2005, Roberts & Holmberg 2010, Frascarelli 2018, Madariaga 2018). Lastly, a fourth class of discourse-prominent languages allow for the omission of subjects when they are topics; this process, however, follows different constraints in different languages (Öztürk 2001, 2008, Schäfer 2021). Greek, Russian, Turkish, English and German are classified as different types of null subject languages, as shown in Table 1.[1] In the following subsections, we will discuss the languages in more detail.

2.1 Null subjects in English

English is a language that does not allow null subjects. This means that subjects, both expletive and referential, are obligatorily overtly realized. Spoken English as well as diary English have been argued to contain null subjects, but it is a matter of controversy whether these registers involve truncated structures or are

---

[1]For a different view on language classification according to the pro-drop feature see Dryer (2013).

Table 1: Pro-drop across languages

Language	Pro-drop type
Greek	consistent/strict pro-drop
Russian	partial pro-drop (or non-pro-drop)
Turkish	topic pro-drop
English	non-pro-drop
German	topic drop

instances of topic drop; see Haegeman & Ihsane (2001) for discussion and references. A noticeable higher use of null subjects in informal conversational English has also been noted, particularly in the first person and in negated constructions such as *don't know* and *can't say* (Harvie 1998, Nagy et al. 2011). Additionally, English discourse fixed expressions in the first and second person singular, such as *gotcha!* and *wanna go?* permit null subjects (Haegeman 2010).

Research on heritage languages with strict pro-drop properties that come into contact with English has shown that their speakers tend to have more realized subjects compared to their monolingual counterparts, indicating transfer effects on subject realization. For instance, Paradis & Navarro (2003) report that bilingual Spanish-English children are more likely to use overt subjects in heritage Spanish than monolingual Spanish-speaking children. Austin et al. (2017) observe the reverse transfer effects in sequential bilingual Spanish-English children who are more proficient in their heritage language Spanish than in their majority language English. Second-generation heritage speakers of Polish in Toronto also tend to use more realized subjects in Polish compared to first-generation and monolingual speakers of Polish in subject continuity contexts (Chociej 2011). A study on heritage Welsh by Boon (2014) found that heritage speakers omit subject pronouns only with the copula verb 'to be' more often than the baseline, but not with other verbs, which she attributes to "incomplete acquisition" of the pro-drop parameter.

2.2 Null subjects in German

German is classified as a topic drop language in certain registers and under certain grammatical constraints (Fries 1988, Trutkowski 2016, Schäfer 2021). Subject drop in German is possible only in the left periphery (pre-field) and is claimed to show restrictions concerning agreement, with first person being especially prone

to subject drop resulting in V1 sentences (Schmitz et al. 2012, Trutkowski 2016, Schäfer 2021). Example 2 shows a sentence with prototypical German topic drop. The subject pronoun *ich* 'I' can be omitted because it is discourse-prevalent and the finite auxiliary verb marks number and person even though in this example first and third person singular would both be permitted. The reading for the first person singular is the default interpretation without further context, but in a context with a given third person singular topic or referent, the interpretation of the empty position would shift to that.

Based on a corpus study and two acceptability rating studies, Schäfer (2021) shows that topic drop in German is influenced by a mixture of factors such as the predictability of the omitted constituent in context, verb probability and distinct verbal inflection.

(2) *pro* kann die Verabredung leider   nicht einhalten
     e   can   the appointment unfortunately not   keep
     'Unfortunately, (I) cannot keep the appointment.' (Schäfer 2021: 116)

Several studies have investigated subject realization in bilingual populations with a null subject heritage language and with German as a majority language (Schmitz et al. 2012, Flores & Rinke 2020, Brehmer et al. 2023). In contrast to what has been found for English as a majority language, these studies suggest that heritage speakers in contact with German do not show transfer effects in their majority language. Brehmer et al. (2023) found in their study that German-Polish bilingual children and their Polish monolingual peers behaved similarly in their narratives and dialogues regarding the realization of subject pronouns.

2.3 Null subjects in Greek

Greek is a consistent pro-drop language showing rich verbal morphology, which means that null subjects can be used in all grammatical feature combinations and syntactic environments (Philippaki-Warburton 1987, Alexiadou & Anagnostopoulou 1998). In the generative tradition, this has been associated with overt movement of the verb to the T(ense) category licensing null subjects (Alexiadou & Anagnostopoulou 1998). In Greek, the default strategy is to use null subjects, except in certain situations such as topicalization, focusing, and contrastive contexts where overt subject realization is required (Tsimpli et al. 2004).

Research on null subjects in Greek bilinguals has focused on both situations where both languages allow null subjects as well as situations where the additional language does not allow null subjects. The findings show that bilingual

adults generally exhibit similar results to monolingual groups. For example, a study conducted by Giannakou & Sitaridou (2022) on adult Greek heritage speakers in Argentina found that these participants produced overt and null subjects in Greek similarly to the monolingual group of Greek speakers. There was also no overuse of overt subjects. The only significant difference was that heritage speakers used null subjects more often in topic shift contexts. Daskalaki et al. (2019) focused on the realization of subjects in heritage Greek children residing in Canada and the US who spoke English as their majority language. The results showed patterns of subject realization to be comparable to monolingual Greek-speaking children. According to Andreou (2020), English-Greek, German-Greek and Greek-Albanian bilingual children exhibited overspecification patterns by using full noun phrases when null subjects were the more appropriate choice. These patterns were modulated either by low language experience in the case of English-Greek and German-Greek speakers with Greek as their heritage language or by language dominance in Greek in the case of Greek-Albanian children with Greek as their majority language. Turning to adult Greek-Italian heritage speakers, Di Domenico & Baroncini (2019) claim that the null subject in Greek is the most employed anaphoric device across speaker groups regardless of Greek being the heritage or the majority language. In contrast, L2 Greek speakers with Spanish as their L1 show diverging patterns in subject realization in Greek depending on the manipulated context (contrastive vs. non-contrastive) and their proficiency in Greek, as reported by Margaza & Gavarró (2020). Finally Torregrossa et al. (2021) tested whether Greek and Italian monolingually-raised speakers differ in the use and interpretation of null and overt subject pronouns known as anaphora resolution and found that Greek null subjects are more flexible referring to subject and object antecedent than Italian null subjects. In sum, heritage speakers' performance is similar to monolinguals, but with some inconsistencies.

2.4 Null subjects in Russian

Since Russian exhibits frequent realization of subjects, but may drop them, it is sometimes classified as a partial pro-drop language or as a non-pro-drop language with frequent subject ellipsis (Fehrmann & Junghanns 2008, Shushurin 2018, Budennaya 2019, 2020). Realization and omission of subject pronouns are determined by various factors, especially information structural categories, the distance between an antecedent and the null subject, and lexical choices such as the type of connector in subordination contexts (Fougeron & Breillard 2004, Bizzarri 2015, Madariaga 2018, 2022, Pekelis 2018, Budennaya 2020). These factors

often interact with each other, which makes it difficult to determine the weight of each factor, especially since pro-drop is optional in many cases.

For instance, it has been claimed that being part of an A(boutness)-Topic chain facilitates omission of the subject with additional requirements concerning the distance between the null subject and the antecedent (Madariaga 2018, 2022). Madariaga (2022) claims that the subject in the second sentence in (3) may be omitted, since it continues an A-Topic chain, whereas the subject in the third sentence does not, thus has to be overtly realized:

(3) My_i idëm na ozero. pro_i Nadeemsja tam vstretit' $Ivana_j$. *$(On)_j$
     we.NOM go.1.PL to lake (we) hope.1.PL there see.INF Ivan.ACC he.NOM
     nam obeščal peredat' ključi.
     we.DAT promised.M pass.INF keys.ACC
     'We are going to the lake. We hope to see Ivan there. He has promised us
     to pass the keys.'

For subordinate contexts, Pekelis (2018) shows that certain connectors like *kak* 'how', *čto* 'that', *potomu čto* 'because', and *esli* 'if', sometimes with and without certain correlates in the main clause, may facilitate or prohibit null subjects. For example, the conditional complementizer *esli* 'if' without a correlate *to* 'then' disallows null subjects as shown in (4) (Shushurin 2018: 155).

(4) [Esli on_i dejstvitel'no idet po toj ulice] *(to) pro_i skoro uvidit
      if he really walks along that street then (he) soon will.see
     stanciju.
     station
     'If he is really walking that street then he'll see the station soon.'

Research on heritage Russian speakers yields inconsistent findings concerning null vs. overt pronominal subjects. Some studies have shown that overt pronominal subjects are more common in heritage Russian varieties in the US than in monolingual ones, which is claimed to indicate transfer from English (Isurin 2011, Dubinina & Polinsky 2013, Ivanova-Sullivan 2014). In contrast, Nagy et al. (2011), examining first vs. second/third generation Russian heritage speakers in Toronto, found no generational differences in the frequency of null subjects, leading the authors to conclude that English had not caused any changes in this heritage Russian population.

2.5 Null subjects in Turkish

Turkish is characterized as a discourse prominent language allowing topic pro-drop, meaning that subjects are not always necessary in certain highly specified contexts as in (5), or when there is continuity of the topic (Öztürk 2001, 2008). In these cases, the use of overt pronouns is considered to be pragmatically marked (Enç 1986). The most widely accepted approaches to analyzing null subject in Turkish have followed the formal analysis that null subjects can be identified through the verbal agreement morphology, which specifies person and number (Enç 1986). However, Öztürk (2001) argues that pro-drop in Turkish is guided entirely by pragmatic constraints and, as such, Turkish should be viewed as a topic drop language, rather than a typical pro-drop language. For the purposes of the current study, we do not take a stance on the specific formal analysis of null subjects in Turkish, but it is important to note that the omission of subjects in Turkish is common. We propose the term "topic pro-drop" to unify existing approaches and signify that Turkish is not a typical pro-drop language like Greek or Italian.

(5) (Ben) su   iç-iyor-um.
     pro (I) water drink-PROG-1SG
     'I am drinking water.'

Studies on Turkish heritage speakers have produced inconclusive results with regard to subject realization in heritage Turkish varieties. Some studies, such as those conducted by Haznedar (2010) and Koban Koç (2016) on heritage Turkish in contact with majority English, found that pronominal subjects are more frequently realized in heritage varieties compared to the monolingual standard. However, other studies, such as those by Azar et al. (2020) and Dikilitaş (2017), found contrasting results. Azar et al. (2020) observed a slightly higher use of overt pronouns in the repertoire of heritage speakers in contact with Dutch, but there were no significant differences between the groups. They argued that the groups align in their frequency of realization of overt subjects and referents overall. Similarly, a longitudinal case study by Dikilitaş (2017) of a Turkish-English bilingual child revealed no evidence of cross-linguistic influence from majority English, as the use of null and overt subjects did not deviate from those of monolingual Turkish peers. These divergent findings may be attributed to several factors, including sociolinguistic background variables that guide the rate of overt subject use, differences in the communities in which heritage speakers acquire the language, and differences in methodologies, including statistical tests and sample sizes.

3 Research questions and hypotheses

Our review of null subjects in the three languages under investigation demonstrates that there are clear differences in the type of pro-drop languages at hand. While Greek requires null subjects in most contexts which are licensed by agreement features, this is more limited in Russian and Turkish. We also outlined that the two majority contact languages in this study, German and English, are both non-pro-drop languages which only allow null subjects in very restricted contexts and to different degrees. The participants in this study exhibit interesting bilingual language combinations. These combinations, along with the unique cross-linguistic comparisons we can draw from this set of languages, prompt several lines of research. These research lines also capture grammatical and extra-linguistic factors, such as animacy and different communicative situations, which we have introduced. Drawing from the current body of research, we formulate the subsequent research questions and corresponding hypotheses:

RQ1: Is the way in which subjects are expressed in heritage Greek, Russian, and Turkish similar to the way in which monolingual speakers express subjects?

H1: We anticipate that heritage speakers will show a higher use of overtly realized subjects.

There are several plausible explanations for a higher use of overt subjects by heritage speakers. One reason might be the effect of cross-linguistic influence from the dominant languages German and English. Bilinguals experience interaction between their language systems, according to the Non-Autonomous Version of the Separate Development Hypothesis (as described by Hulk & Müller 2000, Müller & Hulk 2001, Tracy 2000, 2014). Throughout a speaker's life, transfer from the dominant language may be used to fill in the gaps caused by reduced input in the heritage language. Heritage speakers may adopt subject realization strategies from their dominant language for their heritage language (Gawlitzek-Maiwald & Tracy 1996, Döpke 2000). At the same time, research has shown that heritage speakers tend to be more explicit, which could be another factor that leads them to use overt subjects more frequently than monolinguals (Pashkova et al. 2020).

RQ2: Are there any similarities or differences in the patterns of subject omission observed among various groups of heritage speakers and, if so, what factors may account for these patterns?

H2.1: It can be assumed that these different pro-drop types allow for varying levels of language-internal change or cross-linguistic influence from the majority languages. Therefore, we might anticipate higher use of overtly realized subjects from heritage Greek to heritage Russian to heritage Turkish. However, the design only includes two majority languages that are both non-pro-drop languages. This makes it unsuitable to test for a true effect of cross-linguistic influence as this would require a comparison with a pro-drop majority language. So this sub-hypothesis is only falsifiable outside of the framework of this study.

Apart from cross-linguistic influence, the typological differences might also indicate how stable or dynamic the pro-drop system of a language is internally, regardless of the majority language influence. A strict pro-drop system as in Greek might lead to a more stable use of null subjects in the heritage grammar too. Following the same logic, pro-drop systems that are guided more by pragmatic and situational factors such as those of Russian and Turkish might be open to more dynamicity in the use of null subjects in the heritage language.

H2.2: Alternatively, in the context of heritage languages, we could argue that there may be converging developments regardless of the pro-drop type of the languages. Since the (non)realization of pronominal subjects also involves the interface with external linguistic domains (information structure and discourse organization), we may expect uniform transfer effects from majority languages following the Interface Hypothesis (Sorace 2011). We expect to find more overtly realized subjects in heritage varieties for several reasons. First, there is great variability in linguistic phenomena at the syntax-discourse interface (Tsimpli et al. 2004, Sorace 2011). Additionally, as the morphological repertoire weakens in heritage languages (Polinsky 2018) and this might affect agreement marking which licenses pro-drop, we expect a higher use of overtly realized subjects across all heritage varieties of Greek, Russian and Turkish. Furthermore, some heritage speakers with limited language exposure may not be able to adequately mark verb agreement, leading them to avoid null subjects and use more overt subjects instead (Isurin 2011, Daskalaki et al. 2019). Therefore, individual variance within each group needs to be considered to avoid false positives resulting from the analysis of aggregated frequencies (Baayen et al. 2008, Winter 2011, Brezina & Meyerhoff 2014, Winter 2019).

RQ3: Do Greek, Russian and Turkish heritage speakers' expressions of null subjects in majority German align with monolingual German speakers' productions?

H3: German belongs to the non-pro-drop languages and typologically differs from the typology of the three heritage varieties, namely Greek, Russian, and Turkish, and thus we expect cross-linguistic influence with extended use of null subjects in majority German productions. Additionally, the heritage speakers might also use the reverse strategy and produce more standard-like and formal expressions of subject use in German to display that they are proficient users of their majority language German. This would lead to less topic drop and more overt subjects.

RQ4: Do intra-linguistic factors, such as animacy, or extra-linguistic factors, such as formality (formal vs. informal) and mode (written vs. spoken), have an impact on subject realization in both monolingual and heritage varieties?

H4: Based on the literature discussed earlier, we expect null subjects to be preferred for animate subjects. However, in light of a lack of prior research on the impact of the external factors formality and mode on null subjects for Greek and Turkish and conflicting results for Russian, we approach these factors in an exploratory way.

4 Methodology

Following the *Language Situations* approach (Wiese 2020) described in Section 4.2 below, we collected ecologically-valid semi-spontaneous data through narration tasks to identify emerging trends in the language systems. The sections below provide detailed information on the composition of speakers in the study, corpus annotation and queries as well as on the statistical analysis applied.

4.1 Participants

Table 2 presents the overall information about the number of participants grouped by country of elicitation and the number of tokens in the different sub-corpora. Tokens are used in the way they are defined by the TreeTagger tokenization script (Schmid 1997). All participants were recruited from urban areas to minimize the role of dialect. Urban areas facilitate regional dialect levelling

and were therefore preferred (Britain et al. 2010). Monolingual[2] participants were recruited from Athens, Greece; St. Petersburg, Russia; and İzmir and Eskişehir, Turkey. Bilingual participants were recruited from the United States and Germany. Heritage speakers in the US lived in the greater Washington DC area (including Virginia and Maryland), Chicago, and the greater New York City area (including New Jersey). Heritage speakers in Germany were recruited from the Berlin and Brandenburg urban area.

Table 2: Participant and corpus metadata based on the corpus version 1.0

Country of elicitation	Group	N	Tokens
Greece	monolinguals	64	27,931
Russia	monolinguals	66	25,930
Turkey	monolinguals	64	20,947
	heritage Greek	48	19,782
Germany	heritage Russian	61	32,882
	heritage Turkish	64	23,722
	heritage Greek	63	18,302
USA	heritage Russian	60	29,214
	heritage Turkish	58	18,502

Candidates were invited to take part in the study if they were a) either mono- or bilingual speakers, b) born and raised in the country of the majority language or, in case of bilingual speakers, moved there before the age of 48 months,[3] c) bilinguals should use their heritage language with family and friends regularly, d) not diagnosed with hearing or speech disorders, e) exposed to the majority language (English or German) before the age of five.

In addition, for the written production in Greek and Russian, participants were given the option to use Latin script if they were not able to write in Greek or Cyrillic scripts. Prior to the study, participants were informed about their rights, data protection, and the procedure, and asked to sign a consent form in the

---

[2]We use monolingual as a shorthand to refer to monolingually-raised participants who have not acquired another language other than their first language in childhood (and later potentially learned other languages in school settings) and, crucially, do not use another language than their first language in everyday life.

[3]In exceptional cases participants who moved to the hosting country before the age of six years were admitted to the study.

majority language of the country where the elicitation took place. In cases where the participant was a minor, one of their parents or legal guardians was asked to sign the consent form.

Several considerations were taken to ensure demographic comparability between the participant groups. Two age groups of participants were included: adolescents (aged 14–18) and adults (aged 22–35). Adolescent participants were either currently attending school or had recently graduated. The two different age groups are not relevant for the present study as no effects of age on null subjects were expected. An additional criterion excluded bilingual candidates with extensive formal education in the heritage language, such as those who attended bilingual primary or secondary schools.

From a sociolinguistic and demographic standpoint, Greek, Russian, and Turkish migrant communities in the US and Germany exhibit similar characteristics in their language-related behaviors and experiences within the host countries. These communities often establish tight-knit groups, acting as crucial centers for cultural and linguistic preservation, actively contributing to language maintenance initiatives. Nevertheless, despite endeavors to preserve and pass on linguistic heritage to the younger generations, the heritage language of second, third and fourth generation migrants is observed to undergo substantial transformations (Benmamoun et al. 2013, Polinsky 2018).

4.2 RUEG corpus

The approach employed in this research is a modification of the setup introduced by Wiese (2020), known as the *Language situations* paradigm. This paradigm facilitates the generation of semi-spontaneous data, offering comparable naturalistic information in both oral and written forms, as well as in formal and informal situations. In the elicitation process, participants were presented with a short video depicting a fictional minor car accident, and their task was to narrate the incident as if they had witnessed it, addressing either a close friend or a police officer. Consequently, participants engaged in four distinct communication scenarios during a single session. Heritage speakers took part in two sessions separated by at least three days: one in the majority language and one in their heritage language. Monolingually-raised participants only participated in a single session conducted in the majority language of the respective country. To examine the impact of formality and mode on narrations, we simulated formal-spoken, formal-written, informal spoken, and informal-written settings. The formal part of the elicitation occurred in an office setting, where the elicitor and participant faced each other. The elicitor wore formal attire, and used standardized language

and honorifics. Spoken narration involved leaving a voice message on the police department's answering machine, while the written task required typing a witness report on a "police laptop". For the informal part, another elicitor, stylized as talkative and casually dressed, engaged participants in casual conversation before instructing them to narrate the video's contents via a voice message on WhatsApp to a close friend. The written task involved sending a text message about the accident on WhatsApp to the same friend. The entire session was audiorecorded and the data were pseudonymized. Elicitation orders for the communicative situations were balanced. Once transcribed and annotated, the data were released as the RUEG Corpus (Lüdeling et al. 2024), developed within the Research Unit *Emerging Grammars in Language Contact Situations* (RUEG) funded by the German Research Foundation (https://hu.berlin/rueg). The multilevel annotated RUEG corpus is accessible through the ANNIS interface (Krause & Zeldes 2014), comprising audio for spoken data and visualization options across six sub-corpora for English, German, Greek, Russian, Turkish, and Kurmanji. Additional sub-corpora feature special annotations such as aspect and tense. The present study utilized data from the RUEG corpus (Lüdeling et al. 2024) including narrations from a total of 548 speakers. This encompassed both monolingually-raised and heritage speakers of Greek, Russian, and Turkish.

4.3 Corpus annotation

The RUEG subcorpora for Greek, Russian and Turkish were manually annotated in EXMARaLDA (Schmidt & Wörner 2014) with respect to subject realizations on the three following levels:

- *denotation and animacy of the referents* derived from the context (the full list can be found at https://osf.io/25tw6)

- *syntactic realization* (overt vs. null subjects)

- *expectedness* of subject (i.e., whether an overt subject is expected in the given context using the tags *yes, no, not sure*)[4]

---

[4]The judgments regarding the *expectedness* of subjects were done by at least two native or near-native speakers of each language and cases where the annotators did not agree were further discussed.

4.3.1 Animacy of referents

Subjects were identified in relation to a finite verb corresponding to the broad syntactic notion of subjecthood (McCloskey 1997). The category of subject includes noun phrases such as lexical nouns, pronominal expressions, and demonstratives.

Regarding the denotation of the subjects, we considered human referents like *man*, *woman*, *driver*, and animals like *dog* and personal pronouns referring to them as animate subjects while physical objects like *ball*, *cars*, *groceries* were considered as inanimate subjects. Besides the individual referents, we found some conjoint referents in our data as well. We classified conjoints referents that included at least one animate referent as animate (Adamson & Anagnostopoulou 2024) as in the example below:

(6) para   s   koljaskoj
    couple with stroller
    'A/the couple with a/the stroller.'

4.3.2 Overtness of subjects

The following explains how the layers "syntactic realization" and "expectedness" were operationalized to assess the expression of subject reference in this study. The concept of expectedness was introduced in order to make the present analysis cross-linguistically comparable. Expectedness captures the factors that license and facilitate null subjects which are highly language-dependent as presented in Section 2. By creating the concept of expectedness, we turn the variety of grammatical factors into a binary choice (Yes, an overt subject was expected according to the grammatical rules of the language; or No, an overt subject was not expected according to those extensive criteria). Some details from the guidelines are listed below. The elaborate annotation guidelines that explain the additional annotation layers for this study in depth are available in a PDF-file at https://osf.io/pvmx5.

Our working concept for expectedness differs from the widely used notion of accessibility (Chafe 1987, Gundel et al. 1993, Ariel 2001, Arnold 2010). As Allen et al. (2015) point out, accessibility, generally, is seen as a cognitive proxy to describe how likely a speaker is to produce something on the spectrum between a full form like a noun phrase and a null form like subject drop. Accessibility is one of the factors that we take into account when we code for expectedness which takes into account many additional factors such as topicality and language-specific

grammatical factors. Topicality adds to a threeway interaction with accessibility and expectedness. Additionally, as our descriptions of pro-drop in Greek, Russian and Turkish illustrate, there seem to be stricter grammatical constrains between languages that determine whether a subject may be omitted or overt. This is true regardless of accessibility as Turkish speakers can, for example, realize subjects even when they are accessible, but this is not possible in Greek.

4.3.3 Language specific aspects and exclusions

In all languages of investigation, we excluded dative subjects, subjects in imperatives, rhetorical questions, interrogative as well as exclamative sentences, addressings, and impersonal constructions from the annotation.

For Greek, fixed expressions like *ksero go* 'I don't know' were also excluded as the subject is always expected post-verbally and this could influence the results for overt subjects. 21 such cases were observed in the Greek data.

For the annotation of null subjects in Russian, various factors mentioned in Section 2 were taken into account. Particularly, *communicative situation, discourse, distance between antecedent and subject, lexical markers, information structural categories* were considered. For instance, in an informal-spoken communicative situation, as a tendency, we would expect less overt subject realization in general. More specific situations like contrast or topic change were marked as expecting subjects. Besides, subjects within discourse markers like *(ty) ponimaeš'* '(you) understand' were annotated as not expected ones.

Since Turkish is a topic pro-drop language, as specified earlier, the available number and person marking on the finite verb allow the subject to be omitted. In many instances, speakers also accept realized subjects in non-requiring contexts and positions as grammatical even though these utterances might be perceived as slightly redundant when speakers judge them. However, our data show that these kinds of overt subject realizations are common in different mono- and bilingual speaker's repertoires. Additionally, Turkish requires subjects to be overtly realized to mark a new or later reoccurring topic, including contrast contexts which also can require overt subjects. Example (7) demonstrates a contrastive sentence where the subjects of the main and the subordinate clause are both overtly realized:

(7) ben evi   temizlerken sen boş   boş   film   izliyorsun.
     I   house clean   you empty empty movie watch
     'I'm cleaning the house and you're idly watching a movie.'

One exception in Turkish, however, is the first person pronoun *ben* in narratives. When a finite verb is marked for 1SG, it is unambiguously clear that it refers to the narrator. In this case, it is often not necessary to overtly realize the pronoun, except when it is in a contrastive or emphatic use.

Additionally, 232 data points that were tagged as syntactically realized or not realized had to be excluded because their *expectedness* status was not clear. This means that the formal grammars of the language do not allow a clear prediction as to whether a subject would be expected or not in a given context. Compared to the 31,539 data points that we accounted for, the number of excluded cases is marginal, i.e., <1%. The exact number of data exclusion based on this case is as follows:

Table 3: Excluded data points grouped by language and syntactic condition.

Language	N overt subjects	N null subjects
Greek	3	1
Russian	115	51
Turkish	32	30

4.4 Statistical Analysis

To draw cross-linguistic (Greek, Russian and Turkish heritage speaker groups) and also across-variation (monolingual vs. heritage speakers) comparisons, we ran mixed-effects models in R using the lme4 package (Bates et al. 2015). As independent variables, we included *language* (Greek, Russian, Turkish), *country* (Germany, Greece, Russia, Turkey, US), *expectedness*, *animacy*, *formality* (formal vs. informal), and *mode* (spoken vs. written). Further, to capture the individual speaker variability, we specified random effects by participant. In R notation, the model for each language looked like follows:

```
glmer(realization ~ country + language + expectedness + mode +
    formality + animacy + (1|participant), data = ProdropModel,
    family = binomial, control = glmerControl(calc.derivs=FALSE))
```

In words, this code uses the `lme4` package function `glmer` to create a generalized linear mixed model of the binomial distribution family. The formula models Subject realization with its levels Null and Overt as a function of language group

(heritage USA, heritage Germany, monolingual majority homeland), expectedness according to our annotation scheme (levels yes and no), mode with its levels spoken and written, formality with its levels formal and informal, and animacy with its levels animate and inanimate.

The normalization of data points is conducted according to the number of finite verbs. Table 4 reports the number of finite verb tokens per group.

Table 4: Number of finite verbs per group.

Country of elicitation	Group	N finite verbs
Greece	monolinguals	4,954
Russia	monolinguals	3,965
Turkey	monolinguals	4,609
	heritage Greek	3,494
Germany	heritage Russian	4,624
	heritage Turkish	4,986
	heritage Greek	3,471
USA	heritage Russian	4,342
	heritage Turkish	4,257

5 Results

This section provides the results of the statistical analysis on subject realization and subject drop in different conditions in Greek, Russian, and Turkish. To describe the results from different perspectives, we organize this section in the following subsections: descriptive results (Section 5.1) and intra- and extralinguistic factors (Section 5.2) including formality, mode and animacy. Before we address those perspectives one by one, we present general results and model outcomes in the following.

First, we show how the heritage and monolinguals speakers of Greek, Russian and Turkish use subjects along the two variables which both have two levels, namely *realization* (null vs. overt) and *expectedness* (YES = expected, NO = unexpected).

Figure 1 illustrates all three languages and all participant groups for a better comparison across the same scale. The figure represents a faceted plot with a grid for each of the nine speaker groups sorted by different countries of elicitation
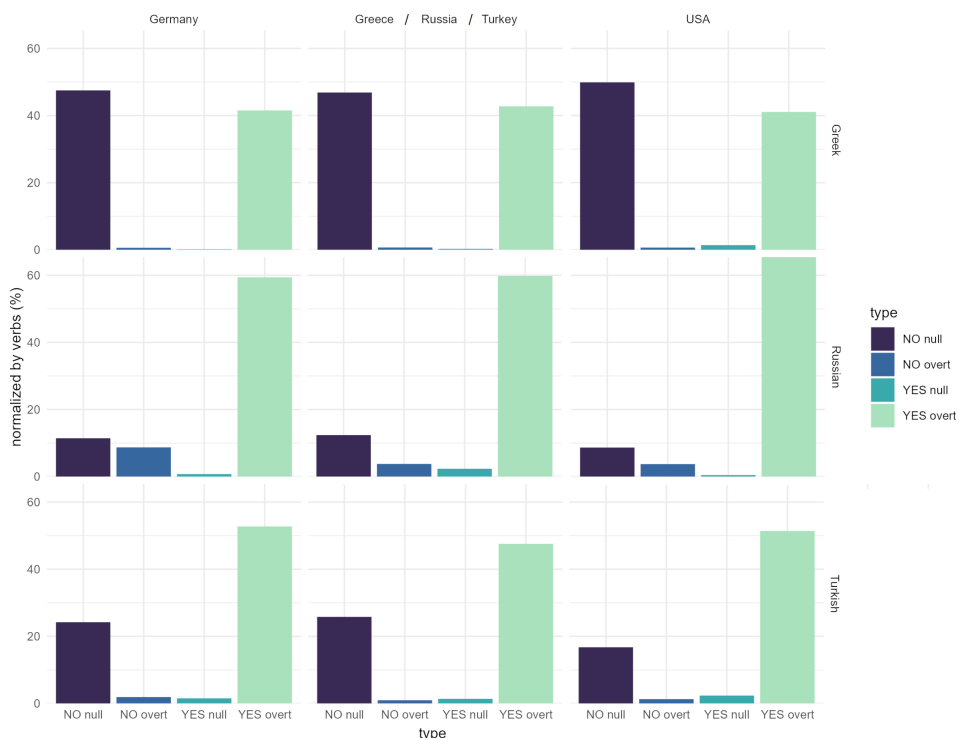
Figure 1: Normalized number of occurrences for four different combinations/categorizations.

from left to right (Germany, respective countries of the monolingual varieties, US) and three languages from top to bottom (Greek, Russian, Turkish). Each row shows one language, and each column reflects one country. For instance, the first grid of the first column represents heritage speakers of Greek in Germany, the second grid in the first column represents heritage speakers of Russian in Germany and the third row in the first column represents heritage speakers of Turkish in Germany. The second column represents monolingual speakers in Greece, Russia and Turkey, respectively. The third column shows heritage speakers of Greek, Russian and Turkish residing in the US, respectively.

Each subplot shows four conditions that are the result of combination of *realization* (null vs. overt) and *expectedness* (YES = expected, NO = unexpected) variables. We list and illustrate each of them below. From left to right, each bar represents the following conditions:

- *NO null*: overt subject is unexpected and it was null

- *NO overt*: overt subject is unexpected, but it was realized

- *YES null*: overt subject is expected, but it was null

- *YES overt*: overt subject is expected and it was realized

The tag "NO overt" applies when an overt subject was observed even though a null subject was expected. In pro-drop languages, this would typically be the case when a referent was just introduced and is just maintained in the discourse. An overt subject in this position would be non-canonical given that it is a pro-drop language. This would be the case if the sentence in Example (8) is a continuation of (10).

(8)   Köpek topu yakalıyor.
     dog   ball catches
     'The dog catches the ball.'

The combination "NO null" was annotated when a canonical case of null subject use was observed. In the given context of the example, this would be the case if the dog is maintained as a referent. We illustrate this in the canonical continuation in (9) as a contrast to non-canonical continuation in (8).

(9)   Topu yakalıyor.
     ball catches
     '(The dog) catches the ball.'

The combination "YES overt" is used to tag canonically overtly expressed subjects in a language. In pro-drop languages, for example, subjects are often expressed overtly when they are introduced for the first time in discourse such as in Example (10). In this case, the dog is introduced in the narrative. We will use this context for all tag combinations.

(10)   Ondan sonra bir köpek topa doğru   koşuyor.
      after   that   a   dog   ball towards runs
      'Then, a dog runs towards the ball.'

Finally, there is the less frequent combination "YES null" where a subject is null even though an overt form was expected. Given our example context, this would be the case if a new referent is introduced in the story but it is not overtly realized as a subject as in Example (11).

(11) Köpeğe kızıyor.
      dog   rant
      '(S)he talks angrily to the dog.'

The *NO null* and *YES overt* conditions represent the canonical way of subject realization, whereas *NO overt* and *YES null* conditions represent the non-canonical subject realization.

The first view on Figure 1 provides the impression that on the one hand, Greek, Russian and Turkish behave differently from each other, and on the other hand, heritage and monolingual speakers of these languages seem to behave similarly to each other.

To address these patterns in a more interpretable way, we present the results for the generalized linear mixed-effects regression models for each language. We fitted three logistic mixed-effects models to predict syntactic realization of subjects (overt vs. null) with *Country*, *Expectedness*, *Mode*, *Formality* and *Animacy* as fixed effects. The model included *Participant* as random effect.

5.1 Descriptive results grouped by language

This subsection presents the descriptive results from heritage Greek, Russian, Turkish and majority German.

5.1.1 Null subjects in Greek, Russian and Turkish

We first asked whether null subjects in the three languages work in the same way. Starting the exploration of our results based on the typological differences on pro-drop in the three languages under investigation we confirm the claim that Greek clearly belongs to strict pro-drop languages while Russian and Turkish do not. As seen in Figure 1, Greek preserves pro-drop in all varieties explored in this study while Russian and Turkish align in the sense that overt subjects are more prominent in all varieties. Russian speakers across all groups realize subjects to a similar extent in the condition where the subject realization is expected. Specifically, the amount of overt subjects is higher than the amount of null subjects. The same observation holds for Turkish speakers.

5.1.2 Topic drop in majority German

We then asked whether each group of heritage speakers produced topic drop in the same way in their majority language, German. An overview of the results from the German data is shown in Figure 2. The black dots that represent the

means almost perfectly mirror each other in all subplots. This points to no differences between the groups with regard to subject realization strategies in the majority language German. We confirmed in a Bayesian model that we report extensively in Özsoy & Burbelko (in preparation) that there are no meaningful differences between the groups. The only strong and important effect of topic drop that we find is based on mode. Specifically, we find that utterances in spoken language are much more likely to include null subjects than utterances in written language.
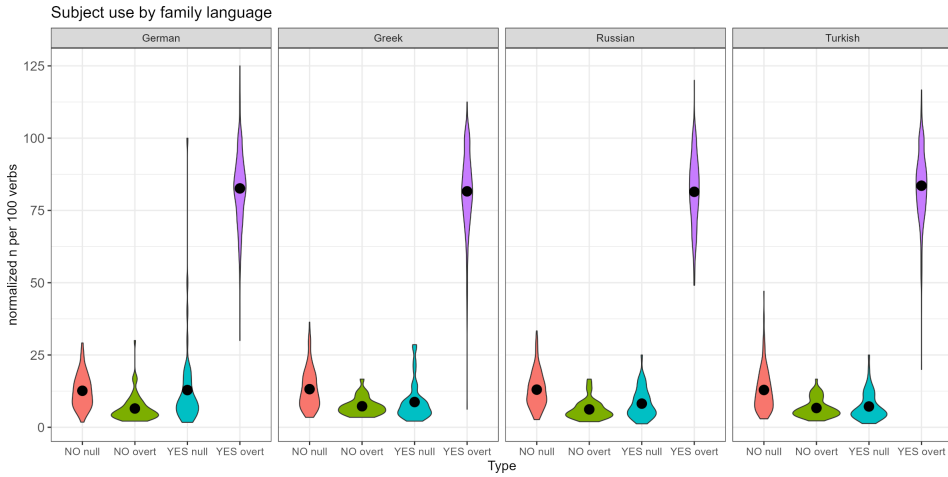
Figure 2: Normalized number of occurrences for four different combinations/categorizations in German data.

5.2 Intra- and extralinguistic factors

In order to verify which factors affect the realization of subjects we conducted a regression analysis and the results are shown below.

5.2.1 Country

One of the main aims of this study is the typological comparison between Greek, Russian, and Turkish as heritage languages (and partly German as a majority language). Since in our controlled design and study the participant groups live in similar sociodemographic and linguistic environments, this direct typological comparison can be drawn. For each language, there are three levels – Germany and the USA as the host countries of the heritage language, and the respective

homeland countries Greece, Russia and Turkey. We can report the effects of each of these levels per language because we sum-coded these variables which allows us to compare all groups equally.

For Greek, we observe that the effect of Country is statistically significant at the USA level (mean = 0.616, SD = 0.205, $p < 0.01$) and non-significant at the Germany (mean = 0.348, SD = 0.213, $p > 0.05$) and Greece (mean = 0.268, SD = 0.194, $p > 0.05$) levels. The directionality of the effect is similar in all groups. This indicates that heritage speakers of Greek in the USA show a higher use of overt subjects compared to the groups in the other countries.

For Russian, the effect of Country is significant at the Germany (mean = 0.472, SD = 0.099, $p < 0.01$) and Russia levels (mean = −0.599, SD = 0.101, $p < 0.01$), and non-significant at the USA level (mean = 0.127, SD = 0.101, $p > 0.05$). The directionality of the effect is positive for the heritage groups and negative for the Russia group. This indicates that heritage speakers of Russian are more likely to produce overt subjects compared to monolingual speakers who show a reverse effect meaning that they drop subjects more frequently.

For Turkish, the model estimates a significant effect at the Germany (mean = 0.328, SD = 0.115, $p < 0.001$) level and non-significant effects at the Turkey (mean = −0.154, SD = 0.114, $p > 0.05$) and USA (mean = −0.174, SD = 0.119, $p > 0.05$) levels. The directionality of the effect is positive for the Germany level and reversed for the Turkey and USA levels. This means that while heritage speakers of Turkish in Germany are likely to produce more overt subjects, this effect is not present for the other groups.

5.2.2 Formality and mode

To account for a broader repertoire of naturalistic language use, our study tests the effects of Mode and Formality which we utilize to capture register variation. For Greek, the effects are non-significant and non-meaningful (Mode: mean = −0.048, SD = 0.098, $p > 0.05$; Formality: mean = 0.118, SD = 0.097, $p > 0.05$). For Russian (Mode: mean = 0.376, SD = 0.048, $p < 0.001$; Formality: mean = 0.202, SD = 0.045, $p < 0.001$) and Turkish (Mode: mean = −0.114, SD = 0.054, $p < 0.001$; Formality: mean = 0.186, SD = 0.054, $p < 0.001$), the effects for Mode and Formality are all significant, and they partly overlap in their directionality. The effect for Formality indicates that Russian and Turkish speakers are more likely to overtly realize subjects in formal contexts compared to informal ones. The Mode effect goes in different directions for the two languages. Russian speakers are more likely to overtly realize subjects in spoken mode and drop more in

written mode. On the contrary, Turkish speakers drop subjects more in spoken mode and overtly realize the subject more when writing.

5.2.3 Animacy

In addition to situational and discourse-dependent factors such as Mode and Formality, our study tests the grammatical factor of Animacy for the realization of subjects. Animacy is statistically significant for all the languages under investigation (Greek: mean = 0.398, SD = 0.201, $p < 0.01$; Russian: mean = 0.361, SD = 0.099, $p < 0.01$; Turkish: mean = 0.254, SD = 0.116, $p < 0.01$). The effect directionality indicates that inanimate subjects are more likely to be overtly realized and animate subjects are more likely to be null.

6 Discussion

Whether bilinguals drop subjects to similar degrees as other groups has sparked a lively debate in bilingualism research (Tsimpli et al. 2004, Argyri & Sorace 2007, Haznedar 2010, Gagarina 2012, Dubinina & Polinsky 2013, Ivanova-Sullivan 2014, 2015, Koban Koç 2016, Azar et al. 2020). A common observation has been that bilinguals tend to use a higher proportion of overt to null subjects than do monolingual speakers. Our study provides a comparative approach to three typologically different languages concerning the use of null subjects. We conducted an exploration in heritage and monolingual varieties of Greek, Russian, and Turkish to investigate intra- and inter-group convergence or divergence regarding pro-drop. The findings suggest that different patterns are observed depending on the type of pro-drop language in use. In Greek as a strict-pro-drop language, no differences are found between speakers in Germany and Greece, but speakers in the US produced more overt subjects. A more frequent realization of subjects is also observed in heritage Russian and heritage Turkish in Germany, which are partial pro-drop and topic pro-drop languages, respectively. No significant differences were observed in heritage Russian and heritage Turkish in the US. In addition, we controlled for other factors that might affect the production of overt or null subjects, and the most significant one across the languages turned out to be animacy. Furthermore, formality and mode appear to have a significant impact on the frequency of overt subjects in Russian and Turkish as well.

From early on, cross-linguistic influence has been discussed as a key facilitator of more overt subjects in heritage speakers of pro-drop languages with non-pro-drop majority languages (Gawlitzek-Maiwald & Tracy 1996, Döpke 2000). However, the literature for Greek as a pro-drop heritage language frequently showed

no differences between bilinguals and monolinguals (Di Domenico & Baroncini 2019, Daskalaki et al. 2019, Andreou 2020, Torregrossa et al. 2021, Giannakou & Sitaridou 2022). Our study adds to this literature with the Greek heritage group in Germany producing similar levels of null subjects as monolingually-raised Greek speakers. Heritage speakers of Greek in the US in our sample align with the general trend in the literature towards a higher use of overt subjects (though this effect was not confirmed in our Bayesian re-analysis of the data in Özsoy et al. (2023) which yielded no differences between the Greek groups). Similarly, Russian heritage speaker groups in the US and in Germany differ significantly from the monolingual variety regarding the frequency of overt subjects. This supports findings by Isurin (2011), Dubinina & Polinsky (2013) and contradict Nagy et al. (2011) who found no influence of majority English on subject drop in Russian. As for Turkish, the findings in the literature have been split with Haznedar (2010) and Koban Koç (2016) finding a higher use in overt subjects whereas Azar et al. (2020), Dikilitaş (2017) do not observe any change in heritage Turkishes. This is particularly interesting since Haznedar (2010), Koban Koç (2016) and Dikilitaş (2017) also investigated heritage Turkish speakers with English as a majority language and found mixed results. We find a higher use in Turkish in Germany but not in the US which confirms the mixed nature of the results.

To contextualize these results better, we draw on an explanation that goes beyond cross-linguistic influence and takes the different types of null subject languages in this study, namely, strict pro-drop Greek, partial pro-drop Russian and topic pro-drop Turkish into account (Philippaki-Warburton 1987, Öztürk 2008, Shushurin 2018). According to this perspective, we would find a continuum from more strict to less strict pro-drop languages. On the strict side of the continuum, we expect the pro-drop system to behave more stably under language contact influence in bilinguals. The more we move toward the less strict system, the more dynamicity in the pro-drop system we would expect. Dynamicity could be reflected as a higher use of overt subjects in bilingual speakers. Since the design of our study did not include pro-drop majority languages, we cannot infer whether the dynamics that we observed here are due to language internal change in bilingual grammars or due to cross-linguistic influence from the majority languages. We also were not able to attribute change to factors that we did not control for, such as sociolinguistic background factors.

There are fewer studies that investigate the dynamics in subject use from a perspective that goes from the heritage to the majority language. In this case, we could have expected bi-directional cross-linguistic influence given the interactions in the systems of bilingual grammars (Zhou et al. 2021). However, this does not seem to be the case as our German topic drop results are in line with previous

studies that found stability in heritage bilingual's majority German grammars (Schmitz et al. 2012, Flores & Rinke 2020, Brehmer et al. 2023). This confirms a general trend, in that more often the status of the language, majority or heritage, seems to play a role for linguistic innovation and change (Wiese et al. 2022).

Besides the effects of bilingualism, we explored different extra- and intra-linguistic factors which could influence the realization of pronominal subjects. Outside of the bilingualism literature, animacy has been extensively investigated as an intra-linguistic factor influencing subject omission, with the clear finding that animate subjects are omitted more often than inanimate subjects (Pu 1997, Meyerhoff 2009, Li et al. 2012, Vollmer 2019). However, to our knowledge, the present study is the first systematic investigation regarding the relation between animacy and its impact on the realization of subjects in (heritage) bilinguals. Our findings regarding animacy in Greek, Russian and Turkish indicate that this factor has a significant effect on subject realization. According to the animacy hierarchies given in Corbett (2000), animate subjects are more salient and thus can be easily omitted. Generally, the issue of animacy as a factor is related and overlapping with other concepts that have been found to be meaningful for the realization of subjects such as topicality or cognitive accessibility (Lee 2007, Allen et al. 2015, Laleko & Polinsky 2017). All of these concepts relate to how present the subject referent is in the speaker's mind or the general context.

As for the contexts of formality and mode in our methodology, we explored whether these extra-linguistic factors affect the realization of subjects. These factors are largely uncharted in the framework of bilingual null subject use. In a study with German (monolingual) corpora, Frick (2017) found that null subjects were more frequent in informal brief text messages such as SMS. In our study, the results suggest that the communicative situation does not affect the three languages in the same way. In Greek, neither formality nor mode seem to be significant factors. The picture is different for Russian and Turkish as formality and mode have an impact on subject realization. Specifically, in both languages participants produce more overt subjects in a formal communicative situation. However, the realization patterns diverge regarding mode. While Russian speakers produce more overt subjects in the spoken mode, the probability of overt subjects in the Turkish group is higher in the written mode. Generally, these findings support our suggested continuum from more strict and stable to less strict and dynamic pro-drop languages that align with the typological differences between Greek, Russian and Turkish. Specifically, the strict pro-drop language Greek does not seem to be affected by extra-linguistic factors and relies more on agreement-marking to license null subjects. Leaving aside the typological differences concerning null subjects in Russian and Turkish, being less explicit in the

informal communicative situation can be related to the relaxed tenor of discourse, the reduced attentiveness of the speaker and in general the nonchalant style of informal communicative situations (Stowell & Massam 2017). On the other hand, one would expect that explicitness conforms with written mode so any misunderstanding could be avoided as the paralinguistic features are absent from all communicative situations in written mode. Thus the question why the heritage groups diverge regarding the extra-linguistic factor mode remains an open issue and a desideratum for future research.

Based on our findings from the elicited production task at hand, we could build hypotheses for experimental tasks that focus on the processing and comprehension of pro-drop by speakers of languages that belong to different types of null subject languages. Specifically for the different groups of heritage Greek, Russian and Turkish speakers, we would expect that heritage speakers of Greek are more restricted in their acceptance of non-canonical overt subjects, but Russian and Turkish heritage speakers might allow a more free use of overt subjects. This would allow for testing whether the observation regarding typological differences in dynamicity in null subject use only holds for production or whether it extends to comprehension as well.

A core finding from our study on heritage Greek, Russian and Turkish is the observation that null subject use in the heritage language seems to be more stable or dynamic based on the underlying type of pro-drop language. While our study allows for more cross-linguistic comparisons than previous studies, it still largely relies on languages that can broadly be classified within a pro-drop theory which depends on agreement marking (Rizzi 1986). Therefore, expanding our observation to other languages within the typology of null subject languages would be interesting. For example, it would be interesting to see how radical-pro-drop languages like Mandarin Chinese, which lack agreement marking, behave in this cross-linguistic comparison (Neeleman & Szendrői 2007). Similarly, we would need to explore the role of partial pro-drop languages like Hebrew and Finnish, which allow null subjects in first and second person but not in third (Camacho 2013). This wider cross-linguistic comparison may allow us to test and generalize our hypothesis that languages which require more null subjects will behave more stably, whereas languages that allow overt subjects more freely will show more dynamicity under language contact. For example, the few studies that exist on the acquisition of null subjects by Chinese heritage language speakers indicate that their use of null subjects stays stable (Chou et al. 2020, Zhang 2021), which would be in line with our predictions. This would not only extend the typological diversity in the study of null subjects, but it could also open ways for

the formalization and theorizing regarding the change of the pro-drop parameter in bi- and multilingual grammmars of heritage speakers.

To our knowledge, this study is the first that compares subject realization in Greek, Russian and Turkish heritage speakers in Germany and the US and accounts for animacy, formality and mode variation. The study offers insights for the same phenomenon in three heritage varieties in contact with two majority languages, making the outcome comparable for the three populations of interest. So far, studies in the field of heritage languages have rarely used a unified methodology for different languages, which raises questions about the generalizability of the results (Winter & Grice 2021). In contrast, our study offers cross-linguistically comparable data from different heritage communities residing in the same countries and selected according to the same demographic criteria. Its wider implications for bilingualism are manifold. First, since in many languages morphological agreement and syntactic marking are crucial for allowing null subjects, heritage speakers' variability in this domain informs us about the grammatical mechanisms in the bilingual mind. Second, the three languages under scope of this study represent different types of pro-drop languages (strict, partial, topic pro-drop). Combining these different types with two non-pro-drop languages (German or English) in a bilingual's mind gives us insights into diverse outcomes. We can provide evidence of the divergence of heritage languages from the monolingual variety and confirm the unlikeliness of some scenarios in language, for example, that a strict pro-drop language completely abolishes this feature due to language contact. Rather, there are slight adjustments which we labeled dynamicity, but there is not a complete reorganization of the system.

Acknowledgements

References

Adamson, Luke James & Elena Anagnostopoulou. 2024. Interpretability and gender features in coordination: Evidence from Greek. In Robert Autry, Gabriela de la Cruz, Luis A. Irizarry Figueroa, Kristina Mihajlovic, Tianyi Ni, Ryan Smith & Heidi Harley (eds.), *Proceedings of the 39th West Coast Conference on Formal Linguistics*, 11–20. Somerville, MA: Cascadilla Proceedings Project.

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language & Linguistic Theory* 16(3). 491–539. DOI: 10.1023/A:1006090432389.

Allen, Shanley E. M., Mary Hughes & Barbora Skarabela. 2015. The role of cognitive accessibility in children's referential choice. In Ludovica Serratrice & Shanley E. M. Allen (eds.), *The acquisition of reference*, 123–153. Amsterdam: John Benjamins.

Andreou, Maria. 2020. The effects of updating and proficiency on overspecification in American-Greek children. *European Journal of Psychological Research* 7(1). 26–32.

Argyri, Efrosyni & Antonella Sorace. 2007. Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition* 10(1). 79–99. DOI: 10.1017/S1366728906002835.

Ariel, Mira. 2001. Accessibility theory: An overview. In Ted J.M. Sanders, Joost Schilperoord & Wilbert Spooren (eds.), *Text representation: Linguistic and psycholinguistic aspects* (Human Cognitive Processing 8), 29–87. Amsterdam: John Benjamins. DOI: 10.1075/hcp.8.04ari.

Arnold, Jennifer E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4(4). 187–203. DOI: 10.1111/j.1749-818X.2010.00193.x.

Austin, Jennifer, Liliana Sánchez & Silvia Perez-Cortes. 2017. Null subjects in the early acquisition of English by child heritage speakers of Spanish. *Romance Languages and Linguistic Theory* 11. 209–27.

Azar, Zeynep, Aslı Özyürek & Ad Backus. 2020. Turkish-Dutch bilinguals maintain language-specific reference tracking strategies in elicited narratives. *International Journal of Bilingualism* 24(2). 376–409. DOI: 10.1177/1367006919838.

Baayen, Rolf Harald, Douglas Davidson & Douglas Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4). Special Issue: Emerging data analysis, 390–412. DOI: 10.1016/j.jml.2007.12.005.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.

Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39(3-4). 129–181. DOI: 10.1515/tl-2013-0009.

Bizzarri, Camilla. 2015. Russian as a partial pro-drop language: Data and analysis from a new study. *Annali di Ca' Foscari: Serie occidentale* (49). 335–362.

Boon, Erin Diane. 2014. *Heritage Welsh: A study of heritage language as the outcome of minority language acquisition and bilingualism.* Harvard University. (Doctoral dissertation).

Brehmer, Bernhard, Aldona Sopata & Massimiliano Canzi. 2023. Same or different? Subject realization in the majority and the heritage language of Polish-German bilingual children. *Linguistics Vanguard* 9(2). 179–189. DOI: 10.1515/lingvan-2022-0061.

Brezina, Vaclav & Miriam Meyerhoff. 2014. Significant or random? A critical review of sociolinguistic generalisations based on large corpora. *International Journal of Corpus Linguistics* 19(1). 1–28. DOI: 10.1075/ijcl.19.1.01bre.

Britain, David, Carmen Llamas & Dominic Watt. 2010. Supralocal regional dialect levelling. *Language and identities.* 193–204.

Budennaya, Evgeniya. 2019. Russian contrastive pronouns in Polish translations: A corpus-based study [Russkie kontrastivnye mestoimenija v pol'skix perevodax: Opyt korpusnogo issledovanija]. *Tipologija morfosintaksičeskix parametrov* 2(2). 11–26.

Budennaya, Evgeniya. 2020. Sub"ekt nastojaščego vremeni v ustnyx i pis'mennyx narrativax [The subject of the present tense in oral and written narratives]. In A. A. Kibrik, Ks. P. Semënova, D. V. Sičinava, S. G. Tatevosov & A. Ju. Urmančieva (eds.), *VAProsy jazykoznanija: Megasbornik nanostatej. Sb. st. k jubileju V. A. Plungjana [Problems of linguistics: A mega-collection of Nanostatas. Collected articles for the anniversary of V. A. Plungian]*, 648–653.

Camacho, José. 2013. *Null subjects.* Cambridge: Cambridge University Press.

Chafe, Wallace L. 1987. Cognitive constraints on information flow. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse: Outcome of a symposium, Eugene, Oregon, June 1984* (Typological Studies in Language 11), 21–51. Amsterdam: John Benjamins. DOI: 10.1075/tsl.11.03cha.

Chociej, Joanna. 2011. Polish null subjects: English influence on heritage Polish in Toronto. *Ms., University of Toronto.*

Chou, Chao-Ting Tim, Tsung-Ying Chen & Acrisio Pires. 2020. Acquisition of null objects in Mandarin Chinese by heritage speakers: Syntax-pragmatics interface knowledge without inflectional morphology. *International Journal of Chinese Linguistics* 7(2). 223–269. DOI: 10.1075/ijchl.19016.cho.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology.* The University of Chicago Press.

Corbett, Greville G. 1979. The agreement hierarchy. *Journal of Linguistics* 15(2). 203–224.

Corbett, Greville G. 2000. *Number.* Cambridge: Cambridge University Press.

Daskalaki, Evangelia, Vasiliki Chondrogianni, Elma Blom, Froso Argyri & Johanne Paradis. 2019. Input effects across domains: The case of Greek subjects in child heritage language. *Second Language Research* 35(3). 421–445. DOI: 10.1177/0267658318787.

Di Domenico, Elisa & Ioli Baroncini. 2019. Age of onset and dominance in the choice of subject anaphoric devices: Comparing natives and near-natives of two null-subject languages. *Frontiers in Psychology* 9. 2729. DOI: 10.3389/fpsyg.2018.02729.

Dikilitaş, Kenan. 2017. Acquisition of pro drop by a Turkish-English bilingual child. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* 32(1). 33–45.

Döpke, Susanne. 2000. Generation of and retraction from cross-linguistically motivated structures in bilingual first language acquisition. *Bilingualism: Language and Cognition* 3(3). 209–226. DOI: 10.1017/S1366728900000341.

Dryer, Matthew S. 2013. Expression of pronominal subjects. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/101.

Dubinina, Irina & Maria Polinsky. 2013. Russian in the USA. In Michael Moser & Maria Polinsky (eds.), *Slavic languages in migration* (Slavische Sprachgeschichte 6). Wien: Lit Verlag.

Enç, Mürvet. 1986. Topic switching and pronominal subjects in Turkish. *Studies in Turkish linguistics* 8. 195.

Fehrmann, Dorothee & Uwe Junghanns. 2008. Subjects and scales. *Linguistische Arbeitsberichte* 86. 189–220.

Flores, Cristina & Esther Rinke. 2020. Factors constraining subject expression in European Portuguese spoken in Hamburg: A bi-generational corpus investigation. *Discours* (26). 3–31. DOI: 10.4000/discours.10648.

Fougeron, I. I. & Žan Breillard. 2004. Mestoimenie *ja* i postroenie diskursivnyx svjazej v sovremennom russkom jazyke. [The pronoun *ja* and the construction of discourse relations in modern Russian]. In Tat'jana Mixajlovna Nikolaeva, I.I. Fužeron, V.S. Efimova, R.F. Kasatkina, L.L. Kasatkin, A.V. Vencov & Žan Brejar (eds.), *Verbal'naja i neverbal'naja opory prostranstva mežfrazovyx svjazej [Verbal and nonverbal supports of interphrase space]*. Jazyki slavjanskoj kul'tury.

Frascarelli, Mara. 2018. The interpretation of pro in consistent and partial null-subject languages: A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in generative grammar: A synchronic and diachronic perspective*, 211–239. Oxford: Oxford University Press.

Frick, Karina. 2017. *Elliptische Strukturen in SMS: Eine korpusbasierte Untersuchung des Schweizerdeutschen*. Berlin: De Gruyter. DOI: 10.1515/9783110517859.

Fries, Norbert. 1988. Über das Null-Topik im Deutschen. *Sprache und Pragmatik: Arbeitsberichte* 3. 19–49.

Gagarina, Natalia. 2012. Discourse cohesion in the elicited narratives of early Russian-German sequential bilinguals. In Kurt Braunmüller & Christoph Gabriel (eds.), *Multilingual individuals and multilingual societies* (Hamburg Studies on Multilingualism 13), 101–120. Amsterdam: John Benjamins. DOI: 10.1075/hsm.13.08gag.

Gawlitzek-Maiwald, Ira & Rosemarie Tracy. 1996. Bilingual bootstrapping. *Linguistics* 34(5). 901–926. DOI: 10.1515/ling.1996.34.5.901.

Giannakou, Aretousa & Ioanna Sitaridou. 2022. (In)felicitous use of subjects in Greek and Spanish in monolingual and contact settings. *Glossa: A journal of general linguistics* 7(1). DOI: 10.16995/glossa.5812.

Guijarro-Fuentes, Pedro & Katrin Schmitz. 2015. The nature and nurture of heritage language acquisition. *Lingua* 164. 239–250. DOI: 10.1016/j.lingua.2015.05.008.

Gundel, Jeanette K, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69. 274–307.

Haegeman, Liliane. 2010. Non-overt subjects in diary contexts. In Joan Mascaró & Marina Nespor (eds.), *Grammar in progress*, 167–174. Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110867848.167.

Haegeman, Liliane & Tabea Ihsane. 2001. Adult null subjects in the non-pro-drop languages: Two diary dialects. *Language Acquisition* 9(4). 329–346. DOI: 10.1207/S15327817LA0904_03.

Harvie, Dawn. 1998. Null subject in English: Wonder if it exists? *Cahiers Linguistique Inguistiques d'Ottawa*. 15–26.

Haznedar, Belma. 2010. Transfer at the syntax-pragmatics interface: Pronominal subjects in bilingual Turkish. *Second Language Research* 26(3). 355–378. DOI: 10.1177/02676583103657.

Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36(4). 533–564. DOI: 10.1162/002438905774464322.

Hulk, Aafke & Natascha Müller. 2000. Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition* 3(3). 227–244. DOI: 10.1017/S1366728900000353.

Isurin, Ludmila. 2011. *Russian diaspora: Culture, identity, and language change* (Contributions to the Sociology of Language 99). Walter de Gruyter. DOI: 10.1515/9781934078457.

Ivanova-Sullivan, Tania. 2014. Interpretation of anaphoric subject pronouns in heritage Russian. In Ryan T. Miller, Katherine I. Martin, Chelsea M. Eddington, Ashlie Henery, Nausica Marcos Miguel, Alison M. Tseng, Alba Tuninetti & Daniel Walter (eds.), *Selected proceedings of the 2012 second language research forum: Building bridges between disciplines*, 145–154. Somerville, MA.

Ivanova-Sullivan, Tania. 2015. *Theoretical and experimental aspects of syntax-discourse interface in heritage grammars*. Leiden: Brill.

Koban Koç, Didem. 2016. Social variables and Turkish subject pronoun use in New York City: The effect of language contact. *Poznan Studies in Contemporary Linguistics* 52(3). 431–453.

Krause, Thomas & Amir Zeldes. 2014. ANNIS3: A new architecture for generic corpus query and visualization. *Digital Scholarship in the Humanities* 31(1). 118–139. DOI: 10.1093/llc/fqu057.

Laleko, Oksana & Maria Polinsky. 2017. Silence is difficult: On missing elements in bilingual grammars. *Zeitschrift für Sprachwissenschaft* 36(1). 135–163. DOI: 10.1515/zfs-2017-0007.

Lee, Chungmin. 2007. Contrastive (predicate) topic, intonation, and scalar meanings. In Chungmin Lee, Matthew Gordon & Daniel Büring (eds.), *Topic and focus*, 151–175. Springer. DOI: 10.1007/978-1-4020-4796-1_9.

Li, Xiaoshi, Xiaoqing Chen & Wen-Hsin Chen. 2012. Variation of subject pronominal expression in Mandarin Chinese. *Sociolinguistic Studies* 6(1). 91. DOI: 10.1558/sols.v6i1.91.

Lüdeling, Anke, Artemis Alexiadou, Shanley E. M. Allen, Oliver Bunk, Natalia Gagarina, Sofia Grigoriadou, Rahel Gajaneh Hartz, Kateryna Iefremenko, Esther Jahns, Kalliopi Katsika, Mareike Keller, Martin Klotz, Thomas Krause, Annika Labrenz, Maria Martynova, Onur Özsoy, Tatiana Pashkova, Maria Pohle, Judith Purkarthofer, Vicky Rizou, Christoph Schroeder, Anna Shadrova, Luka Szucsich, Rosemarie Tracy, Wintai Tsehaye, Heike Wiese, Sabine Zerbian, Yu-

lia Zuban & Nadine Zürn. 2024. *RUEG Corpus*. Version 1.0. Zenodo. DOI: 10. 5281/zenodo.11234583.

Madariaga, Nerea. 2018. Diachronic change and the nature of pronominal null subjects. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in generative grammar: A synchronic and diachronic perspective*, 171–198. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198815853.003.0007.

Madariaga, Nerea. 2022. Referential null subjects in Russian: A synchronic and diachronic overview. In Gréte Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.), *Null subjects in Slavic and Finno-Ugric: Licensing, structure and typology*, 75–103. Boston/Berlin: de Gruyter Mouton. DOI: 10.1515/9781501513848-003.

Margaza, Panagiota & Anna Gavarró. 2020. Null/overt subject alternations in L2 Spanish and L2 Greek. *Glossa: A journal of general linguistics* 5(1). DOI: 10.5334/gjgl.969.

Martynova, Maria, Onur Özsoy, Vasiliki Rizou, Artemis Alexiadou, Natalia Gagarina & Luka Szucsich. 2024. Demonstratives in heritage Greek, Russian, and Turkish in Germany and the US. *International Journal of Bilingualism*. DOI: 10.1177/1367006924126105.

McCloskey, Jim. 1997. Subjecthood and subject positions. In Liliane Haegeman (ed.), *Elements of grammar: Handbook in generative syntax*, 197–235. Springer.

Meyerhoff, Miriam. 2009. Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change* 21(3). 297–317. DOI: 10.1017/S0954394509990196.

Montrul, Silvina. 2015. *The acquisition of heritage languages*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139030502.

Müller, Natascha & Aafke Hulk. 2001. Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition* 4(1). 1–21. DOI: 10.1017/S1366728901000116.

Nagy, Naomi G., Nina Aghdasi, Derek Denis & Alexandra Motut. 2011. Null subjects in heritage languages: Contact effects in a cross-linguistic context. *University of Pennsylvania Working Papers in Linguistics* 17(2). 16.

Neeleman, Ad & Kriszta Szendrői. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38(4). 671–714. DOI: 10.1162/ling.2007.38.4.671.

Oppermann, Sophia. 2021. Subject gap coordination. *Journal of Historical Syntax* 5(16-25). 1–38. DOI: 10.18148/hs/2021.v5i16-25.58.

Özsoy, Onur & Mariya Burbelko. In preparation. *Topic drop in mono- and bilingual native German speakers*.

Özsoy, Onur, Maria Martynova, Vasiliki Rizou, Natalia Gagarina, Luka Szucsich & Artemis Alexiadou. 2023. *Pro-drop realization in heritage Greek, Russian and Turkish is modulated by typological differences*. Paper presented at Deutsche

Gesellschaft für Sprachwissenschaft (DGfS), Universität zu Köln, Cologne, Germany.

Öztürk, Balkız. 2001. Turkish as a non-pro-drop language. In Eser Erguvanlı Taylan (ed.), *The verb in Turkish* (Linguistik Aktuell/Linguistics Today 44), 239–259. Amsterdam: John Benjamins. DOI: 10.1075/la.44.10ozt.

Öztürk, Balkız. 2008. Non-configurationality: Free word order and argument drop in Turkish. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 411–440. Amsterdam: John Benjamins.

Paradis, Johanne & Samuel Navarro. 2003. Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of the input? *Journal of Child Language* 30(2). 371–393. DOI: 10.1017/S0305000903005609.

Pashkova, Tatiana, Mark Murphy & Shanley E. M. Allen. 2020. *Explicitness in referent introduction in heritage speakers' majority English*. Poster presented at Boston University Conference on Language Development (BUCLD 45). Boston, USA.

Pekelis, Olga E. 2018. Èllipsis podležaščego glavnoj klauzy v russkom jazyke i ego svjaz' s tipologiej korreljatov [Subject ellipsis in the main clause in Russian and the typology of correlatives]. *Voprosy jazykoznanija* (6). 31–59.

Perlmutter, David M. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart & Winston.

Philippaki-Warburton, Irene. 1987. The theory of empty categories and the pro-drop parameter in modern Greek. *Journal of Linguistics* 23(2). 289–318. DOI: 10.1017/S0022226700011282.

Polinsky, Maria. 2015. Heritage languages and their speakers: State of the field, challenges, perspectives for future work, and methodologies. *Zeitschrift für Fremdsprachwissenschaft* 26. 7–27.

Polinsky, Maria. 2018. *Heritage languages and their speakers* (Cambridge Studies in Linguistics 159). Cambridge: Cambridge University Press. DOI: 10.1017/9781107252349.

Pu, Ming-Ming. 1997. Zero anaphora and grammatical relations in Mandarin. *Typological Studies in Language* 35. 281–322. DOI: 10.1075/tsl.35.08pu.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17(3). 501–557.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in Minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in Minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13(2). 155–163. DOI: 10.1177/1367006909339814.

Rothman, Jason, Fatih Bayram, Vincent DeLuca, Grazia Di Pisa, Jon Andoni Duñabeitia, Khadij Gharibi, Jiuzhou Hao, Nadine Kolb, Maki Kubota, Tanja Kupisch, Tim Laméris, Alicia Luque, Brechje van Osch, Sergio Miguel Pereira Soares, Yanina Prystauka, Deniz Tat, Aleksandra Tomić, Toms Voits & Stefanie Wulff. 2023. Monolingual comparative normativity in bilingualism research is out of "control": Arguments and alternatives. *Applied Psycholinguistics* 44(3). 316–329. DOI: 10.1017/S0142716422000315.

Schäfer, Lisa. 2021. Topic drop in German: Empirical support for an information-theoretic account to a long-known omission phenomenon. *Zeitschrift für Sprachwissenschaft* 40(2). 161–197. DOI: 10.1515/zfs-2021-2024.

Schmid, Helmut. 1997. Probabilistic part-of-speech tagging using decision trees. In Daniel B. Jones & Harold Somers (eds.), *New methods in language processing*, 154–164. London: Routledge.

Schmidt, Thomas & Kai Wörner. 2014. EXMARaLDA. In Jacques Durand, Ulrike Gut & Gjert Kristoffersen (eds.), *Handbook on corpus phonology*, 402–419. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199571932.013.030.

Schmitz, Katrin, Marisa Patuto & Natascha Müller. 2012. The null-subject parameter at the interface between syntax and pragmatics: Evidence from bilingual German–Italian, German–French and Italian–French children. *First Language* 32(1-2). 205–238. DOI: 10.1177/0142723711403880.

Shushurin, Philip. 2018. Null pronouns in Russian embedded clauses. In Pritty Patel-Grosz, Patrick Georg Grosz & Sarah Zobel (eds.), *Pronouns in embedded contexts at the syntax-semantics interface*, 145–169. Cham: Springer. DOI: 10.1007/978-3-319-56706-8_5.

Silverstein, Michael & Robert MW Dixon. 1986. Hierarchy of features and ergativity. *Features and Projections* 25. 163–232.

Sorace, Antonella. 2011. Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism* 1(1). 1–33. DOI: 10.1075/lab.1.1.01sor.

Stowell, Tim & Diane Massam. 2017. Introducing register variation and syntactic theory. *Linguistic Variation* 17(2). 149–156. DOI: 10.1075/lv.00001.mas.

Torregrossa, Jacopo, Maria Andreou, Christiane Bongartz & Ianthi-Maria Tsimpli. 2021. Bilingual acquisition of reference: The role of language experience, executive functions and cross-linguistic effects. *Bilingualism: Language and Cognition* 24(4). 694–706. DOI: 10.1017/S1366728920000826.

Tracy, Rosemarie. 2000. Language mixing as a challenge for linguistics. In *Cross-linguistic structures in simultaneous bilingualism* (Studies in Bilingualism 21), 11–36. Amsterdam: John Benjamins. DOI: 10.1075/sibil.21.04tra.

Tracy, Rosemarie. 2014. Mehrsprachigkeit: Vom Störfall zum Glücksfall. In Manfred Krifka, Joanna Błaszczak, Annette Leßmöllmann, André Meinunger, Barbara Stiebels, Rosemarie Tracy & Hubert Truckenbrodt (eds.), *Das mehrsprachige Klassenzimmer: Über die Muttersprachen unserer Schüler* (Springer Spektrum), 13–33. Berlin Heidelberg: Springer VS. DOI: 10.1007/978-3-34315-5.

Trutkowski, Ewa. 2016. *Topic drop and null subjects in German* (Linguistics & Philosophy 6). Berlin: de Gruyter. DOI: 10.1515/9783110446173.

Tsimpli, Ianthi-Maria, Antonella Sorace, Caroline Heycock & Francesca Filiaci. 2004. First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism* 8(3). 257–277. DOI: 10.1177/13670069040080030601.

Valdés, Guadelupe. 2005. Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal* 89(3). 410–426. DOI: 10.1111/j.1540-4781.2005.00314.x.

Vollmer, Maria C. 2019. *How radical is pro-drop in Mandarin? A quantitative corpus study on referential choice in Mandarin Chinese.* University of Bamberg. (Doctoral dissertation).

Wiese, Heike. 2020. Language situations: A method for capturing variation within speakers' repertoires. In Yoshiyuki Asahi (ed.), *Methods in dialectology XVI* (Bamberg Studies in English Linguistics 59), 105–117. Frankfurt am Main: Peter Lang.

Wiese, Heike, Artemis Alexiadou, Shanley E. M. Allen, Oliver Bunk, Natalia Gagarina, Katerina Iefremenko, Maria Martynova, Tatiana Pashkova, Vicky Rizou, Christoph Schroeder, Anna Shadrova, Luka Szucsich, Rosemarie Tracy, Wintai Tsehaye, Sabine Zerbian & Yulia Zuban. 2022. Heritage speakers as part of the native language continuum. *Frontiers in Psychology* 12. 5982. DOI: 10.3389/fpsyg.2021.717973.

Winter, Bodo. 2011. Pseudoreplication in phonetic research. In *17th International Congress of Phonetic Sciences (ICPhS)*, 2137–2140. Hong Kong, China.

Winter, Bodo. 2019. *Statistics for linguists: An introduction using R.* Routledge.

Winter, Bodo & Martine Grice. 2021. Independence and generalizability in linguistics. *Linguistics* 59(5). 1251–1277. DOI: 10.1515/ling-2019-0049.

Zhang, Xinye. 2021. Language variation in mandarin as a heritage language: Subject personal pronouns. *Heritage Language Journal* 18(1). 1–29. DOI: 10.1163/15507076-12340020.

Zhou, Jiangling, Ziyin Mai & Virginia Yip. 2021. Bidirectional cross-linguistic influence in object realization in Cantonese–English bilingual children. *Bilingualism: Language and Cognition* 24(1). 96–110. DOI: 10 . 1017 / S1366728920000231.