

# Chapter 13

## Majority English of heritage speakers

Tatiana Pashkova<sup>a</sup>, Marlene Böttcher<sup>b</sup>, Kalliopi Katsika<sup>a</sup>,  
Sabine Zerbian<sup>c</sup> & Shanley E. M. Allen<sup>a</sup>

<sup>a</sup>University of Kaiserslautern-Landau <sup>b</sup>Kiel University <sup>c</sup>University of Stuttgart

Research on heritage speakers to date has focused largely on their heritage languages. In contrast, while their majority languages have been often examined in young children, less is known about adolescents and adults. Filling this gap, the current chapter presents our research on majority English of adolescent and adult heritage speakers of German, Greek, Russian, and Turkish compared to monolingually-raised English speakers. Using the RUEG corpus, we investigated four external interface areas prone to dynamicity in language contact – prosody, article use, reference, and clausal syntax in relation to information and discourse structure. In addition, we present findings on discourse organization. Based on our results, we discuss implications for majority language research.

### 1 Introduction

Most research on heritage speakers (HSs) to date focuses on their heritage language. In this chapter we focus on their other language – the majority language in the society – and ask whether the bilingualism resulting from heritage language (HL) maintenance leads to differences in the way heritage and monolingually-raised speakers use the majority language (ML).

HSs are bilinguals who typically speak one language within the family and other personal spheres (e.g., cultural or religious settings) – the HL – and another language that is the language of the larger society in which they live – the ML (Pascual y Cabo & Rothman 2012, Montrul 2016, Montrul & Polinsky 2021). They typically begin acquiring the HL at birth, and begin learning the ML sometime between birth and the onset of schooling. Through formal schooling they often



shift in dominance from the HL to the ML, such that they become dominant in the ML by adulthood. There is typically a wide variation across speakers in adult proficiency in the HL: they may retain full fluency, may gradually reduce dominance in HL so that they only comprehend but do not speak the HL, or anything in between.

The linguistic study of HSs has blossomed over the last two decades. However, as already noted, most research on HSs to date focuses on their HL (Montrul 2008, Kupisch 2013, Montrul 2016, Polinsky 2018, Montrul & Polinsky 2021). This is understandable from the cultural standpoint since there is considerable interest in supporting HL retention, and also from the linguistic research standpoint since it is both theoretically and practically relevant to explore what the HL looks like after dominance is lost.

However, research on the ML is also valuable for at least two reasons. First, it contributes to our understanding of cross-linguistic influence and effects of bilingualism. Many studies on HSs have focused on the influence of the ML on the HL – in other words, the more dominant language on the less dominant one. However, research on L2 learners has highlighted that the influence also occurs in the opposite direction, from the less dominant language to a more dominant one (see, for example, Schoonbaert et al. 2007, Chen et al. 2013 for late L2-to-L1 structural priming; Pavlenko & Jarvis 2002, Hohenstein et al. 2006, Gorba 2019 for late L2-to-L1 cross-linguistic influence).

Thus, we expect that research on the ML of HSs can provide insights into whether this type of cross-linguistic influence can also be attested in HL-to-ML situations. Moreover, we might observe general effects in the ML that are not traceable to any particular HL, which would add evidence that HL bilingualism in itself can lead to dynamic language patterns.

Second, ML research on adolescent and adult HSs is crucial to inform language education policies, which often argue against HL maintenance due to its harmful effects on ML ultimate attainment (e.g., McCardle 2015, Tracy & Gawlitzek 2023: 30–33). By comparing the ML of adult and adolescent HSs to their MS counterparts, we aim to provide practical evidence that HL retention does not negatively affect the ML in the long-term perspective.

Most research on the ML focuses on children given the goal that children become competent in the ML to facilitate school achievement, with primary focus on the lexicon and morphology (e.g., Bialystok et al. 2010, Marinis & Chondrogianni 2010, Paradis & Jia 2017). Research on the ML in adolescents and adults is also evident in the literature, but it is relatively less common. Most of it shows some differences between the ML as produced or comprehended by HSs compared to MSs. In the area of *semantics*, for example, both Scontras et al. (2017)

and Lee et al. (2011) showed differences in scope marking patterns in majority English between English MSs and Korean and Mandarin HSs, which the authors attribute to cross-linguistic influence from the scope patterns in Korean and Mandarin. With regard to *phonology*, HSs are typically found to not have a detectably different accent from MSs of the ML (e.g., Kupisch et al. 2014, Lloyd-Smith et al. 2020), although there is some evidence that they articulate more explicitly (Polinsky 2018: 142–144). Findings for perception are mixed: Chang (2016) finds that Korean HSs perceive unreleased stops more accurately in majority English than English MSs due to cross-linguistic influence from Korean, while Lee-Ellis (2012) shows that HSs with the same language pair perform like English MSs in perceiving an English-specific contrast in nonce words. For prosodic aspects of the ML of HSs, cross-linguistic influence is more consistently evidenced. This was shown for phonetic and phonological aspects of rises in ML German for Turkish HSs (Queen 2012), for broader prosodic features in the intonation of ML Dutch by Turkish HSs (van Rijswijk et al. 2017), and for intonational aspects of polar questions in ML English by Icelandic HSs (Dehé 2018).

In the domain of *morphosyntax*, most but not all studies have found effects of bilingualism in the ML of HSs. For example, three studies on referent tracking found differences between MSs and HSs in the ML. Azar et al. (2020) found that Turkish HSs produced more pronouns as compared to null forms for referent maintenance in majority Dutch compared to Dutch MSs, which can be viewed as higher explicitness of HSs than MSs in the ML. The authors attribute this explicitness to differences in likely input to MSs vs. HSs in the ML: HSs may have many more interlocutors who are L2 speakers of Dutch and who themselves are more explicit in their pronoun use. Contemori and colleagues (Contemori & Ivanova 2021, Contemori et al. 2023) found a somewhat opposite pattern – that Spanish HSs often produced pronouns for referent tracking in majority English in contexts where English MSs would typically produce full NPs. They attributed this to cross-linguistic influence from Spanish, since pronouns would have been appropriate in these contexts in Spanish. Other differences between HSs and MSs were found in the use of definite articles in ML German by Turkish HSs (Felser & Arslan 2019), acceptability judgements of definite and indefinite articles in ML English by Spanish HSs (Montrul & Ionin 2010), and use of subject-verb agreement in ML English by HSs with various HLs (Paradis 2019). However, Montrul & Ionin (2010) found that Spanish HSs patterned like English MSs in some aspects of interpretation of definiteness and possession in majority English, while Lee-Ellis (2012) found that Korean HSs patterned mostly like English MSs in their interpretation of locative alternation structures in majority English, and Kupisch et al. (2017) found that Turkish HSs performed like German MSs in their compre-

hension of subtle definiteness effects in majority German. Finally, two studies on a variety of morphosyntactic phenomena found that Spanish HSs patterned more closely to Swedish MSs in majority Swedish the higher their language aptitude (Bylund et al. 2012) and the earlier their age of onset of Swedish (Bylund et al. 2021).

In sum, most of the existing research on the ML use of HSs across linguistic domains shows some cross-linguistic influence from their HL or from language contact in general, although some research also shows no differences between HSs and MSs in ML use. Possible reasons for divergent results of various studies are the different language areas and different languages examined, different methods employed and different speakers tested. Our work reduces this methodological variation to different language areas only, while keeping the languages (majority English in contact with heritage German, Greek, Russian, and Turkish), the method (examination of elicited narratives)<sup>1</sup> and the speakers the same. This approach will show us whether there exists a more consistent pattern of difference/similarity between HSs and MSs in the ML.

Furthermore, our research addresses several open theoretical questions in the majority language literature. First, do we see consistent influence on majority English from typologically different HLs if speakers of these HLs are tested in the same methodological paradigm? Only a few studies investigate more than one HS group (Kupisch et al. 2014, Polinsky 2018, Paradis 2019, Labrenz 2023). This is important because it is only through comparing across HS groups that we can determine if a pattern really derives from cross-linguistic influence (different effect across groups depending on properties of HL) or from effects of bilingualism or language contact in general (similar effect across groups regardless of HL).

Second, is the effect of HS bilingualism equally evident in different registers of the ML and do HSs separate the registers in the ML the same way as MSs? The role of register has been examined in the HL, with some indications of register leveling (e.g., Alexiadou et al. 2022, Alexiadou & Rizou 2023, Tsehay 2023). However, few if any studies have investigated register differentiation in the ML (see Labrenz 2023 as an example of a recent RUEG-based study), so it is unclear if registers are dynamic not only in the HL but also in the ML.

Third, what differences between HSs and MSs can we discover outside of a strict experimental setting and in a more ecologically valid set up? Many of the previous ML studies that found a difference between HSs and MSs used carefully designed experiments to test specific language phenomena (e.g., Lee et al.

---

<sup>1</sup>Except for one study on article use reported in Section 5, which used a written fill-in-the-blank test.

2011 and Scontras et al. 2017 on scope; Contemori & Ivanova 2021 on referring expressions; Felser & Arslan 2019 on definiteness; Paradis 2019 on subject-verb agreement; Bylund et al. 2012, 2021 on various phenomena; Montrul & Ionin 2010 on article choice; van Rijswijk et al. 2017 on prosody). Some authors highlighted that the experiments were designed to be very demanding even for MSs (Bylund et al. 2012, Paradis 2019). In contrast, we are aware of only a few studies that discovered a difference between HSs and MSs in the ML in an elicited narrative set up (Azar et al. 2020, Contemori et al. 2023 on referring expressions). Other studies that used naturalistic data from tasks not specifically targeting a certain linguistic phenomenon (elicited narratives, interviews, picture descriptions) either did not find a difference between HSs and MSs in the ML (Kupisch et al. 2014 and Lloyd-Smith et al. 2020 on perceived foreign accent) or did not provide a conclusive comparison of HSs and MSs (Kupisch 2014 on adjective placement; Lein et al. 2016 on VOT production; Queen 2012 on intonation).

In the present chapter, we address these gaps by reviewing several studies examining the production patterns in majority English of German, Greek, Russian and Turkish HSs living in the USA. These studies were completed in the context of the six-year research project Research Unit *Emerging Grammars* (RUEG) funded by the German Research Foundation (FOR2357). Each of these studies examines a structure at the interface between a core area of language (e.g., phonology, morphology, syntax) and a non-core area of language (e.g., pragmatics, discourse). This follows the Interface Hypothesis (Sorace 2011, Tsimpli 2014), which states that phenomena at the interface between core and non-core areas of language are particularly open to the development of dynamic patterns in bilinguals. Interface structures crucially involve the integration of knowledge between different areas of language, and thus engender a higher cognitive load for processing than structures involving only one area of language. Each study looks for dynamic patterns in the ML productions of HSs and examines them in the light of potential sources of the pattern, including cross-linguistic influence or transfer from the HL to the ML as well as the presence of a language contact situation regardless of the particular languages involved (e.g., increased explicitness compared to MSs). Crucially, the data were collected via elicited narrative in four different communicative situations leading to production of different registers, thus allowing us to see productions in a more ecologically valid context than the experimental setups used in many of the studies to date. Finally, data from the same participants were analyzed in each of the studies on different linguistic phenomena, thus reducing inter-individual variation between studies.

The remainder of the chapter is structured as follows. Section 2 describes the RUEG corpus (Lüdeling et al. 2024) – the source of most of the ML data used in

the studies presented in this chapter. Section 3 looks at HSs' prosody, Section 4 examines reference, Section 5 covers article use, Section 6 explores syntax, and Section 7 turns to openings and closings in discourse. Section 8 critically discusses the reported findings and draws conclusions for heritage language and bilingualism research.

## 2 Corpus

The studies reported in this chapter are based on the data from the English subcorpus of the RUEG corpus (Wiese et al. 2021). The subcorpus contains the English data of 223 HSs with English as their majority language and 64 English MSs, all of whom were raised in the USA and resided there at the time of testing. The HS group includes 34 German, 65 Greek, 65 Russian, and 59 Turkish HSs. A total of 91% of HSs were first exposed to English at the age of 5 or earlier, with 46% having the first contact with English from birth. All speakers – HSs and MSs – comprise two age groups, adolescents (13–18 years old) and adults (20–37 years old); see Table 1.

Speakers were interviewed in large cities in the Eastern United States, or in nearby areas. These included New York City; Washington, D.C.; Chicago, Illinois; Boston, Massachusetts; Madison, Wisconsin; St. Paul and Minneapolis, Minnesota; Long Island, New York; Fort Lee and Bloomfield, New Jersey; and New Haven, Connecticut.

Table 1: Characteristics of speakers from the English subcorpus of the RUEG corpus

Group	Adolescents			Adults		
	N (Male)	Age Mean (sd)	Eng. AoO Mean (sd)	N (Male)	Age Mean (sd)	Eng. AoO Mean (sd)
English MSs	32 (13)	16.1 (1.4)	-	32 (13)	28.5 (3.9)	-
German HSs	27 (15)	5.5 (1.5)	0.3 (0.7)	7 (2)	25.3 (4.1)	0.9 (1.5)
Greek HSs	33 (16)	16.3 (1.4)	1.1 (1.8)	32 (13)	29.1 (3.4)	1.0 (1.8)
Russian HSs	32 (13)	15.8 (1.4)	2.5 (2.0)	33 (11)	27.5 (3.3)	3.7 (2.0)
Turkish HSs	32 (10)	16.0 (1.6)	2.6 (2.1)	27 (9)	26.2 (4.1)	2.2 (2.2)
Total	156 (67)	15.9 (1.5)	-	131 (48)	27.7 (3.8)	-

The data were collected using the Language Situations methodology (Wiese et al. 2025 [this volume], Wiese 2020), which allows eliciting comparable and naturalistic productions across communicative situations. Participants were shown a brief non-verbal video of a minor car accident and were asked to recount what they saw. The procedure consisted of two formality settings. In the formal setting, the participants were met by a formally dressed elicitor in an office-like room, whereas in the informal setting, the elicitor dressed casually and met the participant in a more relaxed environment, offering snacks and drinks. Prior to the informal session, the participant and the elicitor engaged in 10–15 minutes of conversation to create an easy-going atmosphere. The participant watched the video three times in total, twice in the first setting, once in the second setting, and recounted the event in spoken and written modes.

In the formal setting, participants left a voice message to a police hotline (spoken mode) and provided a written witness report to the police (written mode). In the informal setting, participants recorded a WhatsApp voice message to a friend (spoken mode) and typed a WhatsApp text message to a friend (written mode). The order of settings (formal and informal) and modes (spoken and written) was balanced among participants. English MSs accomplished all the tasks during one session. HSs completed the tasks in two sessions, one in English, their ML, and one in their HL, with a 3–5 day interval between sessions to reduce priming effects. The order of the language sessions was counterbalanced among HSs. After completing the narrative tasks, all participants filled out a language background questionnaire, which included a self-assessment of their language skills. The self-assessments showed that HSs and English MSs rated their English skills comparably high. In addition, HSs assessed their skills higher in their majority English than in their HL (Table 2).

The spoken data were transcribed in Praat (Boersma 2001) or EXMARaLDA (Schmidt & Wörner 2014). Subsequently, spoken and written data were annotated in EXMARaLDA for various phenomena of interest. The annotated data were accessed through the corpus tool ANNIS (Krause & Zeldes 2014; see Shadrova et al. 2025 [this volume] and Klotz et al. (2024) for further detail on the RUEG corpus).

To assess speakers' fluency and proficiency in English, we calculated their speech rate and lexical diversity (Table 3). Speech rate (in syllables per second) was calculated using a Praat script (de Jong et al. 2021) based on all spoken narratives in the English subcorpus. Lexical diversity was evaluated with two measures – the moving-average type-token ratio (MATTR) and measure of textual lexical diversity (MTLD), demonstrated to produce stable results for short texts

(Zenker & Kyle 2021). MATTR and MTLD were calculated based on all narratives in the English subcorpus using the lexical diversity package in Python (Kyle 2020)

Table 2: Language skill self-ratings of speakers from the English subcorpus of the RUEG corpus

Group	Spoken understanding Mean (sd)	Spoken production Mean (sd)	Written understanding Mean (sd)	Written production Mean (sd)
<b>In majority English</b>				
English MSs	4.92 (0.27)	4.81 (0.47)	4.80 (0.48)	4.77 (0.46)
German HSs	5.00 (0.00)	5.00 (0.00)	5.00 (0.00)	4.97 (0.17)
Greek HSs	4.97 (0.18)	4.95 (0.21)	4.88 (0.33)	4.89 (0.31)
Russian HSs	4.92 (0.32)	4.92 (0.27)	4.91 (0.29)	4.89 (0.36)
Turkish HSs	4.97 (0.18)	4.93 (0.25)	4.92 (0.38)	4.92 (0.34)
<b>In heritage languages</b>				
German HSs	4.47 (0.56)	3.65 (0.85)	3.71 (1.06)	2.88 (1.27)
Greek HSs	4.16 (0.78)	3.72 (1.05)	3.46 (1.20)	3.13 (1.29)
Russian HSs	4.38 (0.75)	3.87 (0.85)	3.25 (1.05)	2.83 (1.22)
Turkish HSs	4.28 (0.83)	3.90 (0.95)	3.37 (1.05)	3.30 (1.16)

Table 3: Speech rate (syll/sec), MATTR, and MTLD of speakers from the English subcorpus of the RUEG corpus (in majority English)

Group	Speech rate Mean (sd)	MATTR Mean (sd)	MTLD Mean (sd)
English MSs	3.26 (0.52)	0.678 (0.056)	38.80 (11.19)
German HSs	3.34 (0.46)	0.678 (0.049)	40.43 (10.45)
Greek HSs	3.23 (0.44)	0.670 (0.053)	37.66 (9.85)
Russian HSs	3.20 (0.47)	0.672 (0.049)	38.54 (10.06)
Turkish HSs	3.22 (0.45)	0.659 (0.050)	36.10 (10.63)

Most studies in this chapter examined the data of various subsets of speakers from the English subcorpus. The exact subset is specified in the description of each study. Furthermore, the studies used different corpus versions (e.g., 0.3.0 or 0.4.0), which is detailed in the cited papers. If a study did not employ the



English subcorpus, the characteristics of the data set are provided in the study's description.

### 3 Prosody

As noted earlier, previous studies have found that HSs tend to pattern with MSs in rating studies assessing their perceived foreign accent in the ML (e.g., Kupisch et al. 2014, Lloyd-Smith et al. 2020, but see Polinsky 2018: 138–153 for a review of segmental and suprasegmental differences). However, they often show effects of cross-linguistic influence in ML prosody (Queen 2012, van Rijswijk et al. 2017, Dehé 2018). In the RUEG project we contributed to research on prosodic interference in the ML of HSs (here majority English), investigating two phenomena – stressed pronouns and contrastive adjectives – at the interface of prosody and information structure.

One line of research looked at the prosodic realization of stressed pronouns in English MSs and HSs speaking majority English (for details, see Böttcher 2021, Böttcher & Zerbian 2020). In English, pronouns as discourse-given constituents are generally unaccented; thus, prominence on pronouns (stress) indicates focus or contrast (Krifka 2008, Selkirk 1996), as illustrated in (1) (small caps indicate prosodic prominence in the form of a pitch accent).

- (1) There was this couple walking down the sidewalk.
  - a. They<sub>GIVEN</sub> were about to cross the street.
  - b. [HE]<sub>CONTRAST</sub> had a ball while [SHE]<sub>CONTRAST</sub> was pushing a baby carriage.

In a pilot study, Böttcher & Zerbian (2020) investigated the realization of pronouns in majority English by six English MSs and 12 HSs (three Turkish, four Russian, and five Greek). The focus of this investigation was to analyze and explain the occurrence of stressed pronouns. Pronouns are not expected to be frequently stressed as given constituents (Krifka 2008, Selkirk 1996). However, the prosodic realization of given constituents is reportedly different in spontaneous compared to read speech (de Ruiter 2015). The spontaneous narrations from the RUEG corpus (two by each speaker, formal and informal) were, therefore, analyzed for the occurrence of stressed pronouns, prosodic phrasing, and information structure (following Krifka 2008, Baumann & Riester 2013, Himmelmann et al. 2018). The analysis revealed that stressed pronouns were produced frequently in spontaneous narrations by HSs and English MSs. Contrast was the most frequent factor for stressed pronoun realization ( $n = 52$ , 38%), while prosodic phrasing was also

important (separate or small phrase:  $n = 30$ , 22%; phrase finally:  $n = 16$ , 12%; and phrase initially:  $n = 33$ , 25%).

In a subsequent study, Böttcher (2021) explored whether HSs and English MSs differed in their prosodic realization of contrastively focused and not focused pronouns in majority English. This was done based on 40 narrations by 20 speakers from the RUEG corpus 0.3.0. – four English MSs and 16 HSs, with four speakers each of heritage Greek, Turkish, Russian, and German. Pronouns were analyzed for their information structure (Götze et al. 2007) and their prosodic realization (Beckman et al. 2006, Kügler & Baumann 2019). Overall, speakers show similar tendencies of realizing pronouns: unstressed if unfocused and stressed if contrastively focused (Figure 1). Overall, there is a tendency for fewer stressed pronouns in the narrations of HSs, independently of contrastive focus (see Böttcher 2021 for further detail). In the case of contrastive focus, only one pronoun in a monolingual production was unstressed, while heritage speakers leave these pronouns unstressed in about half of the cases. Since especially Turkish HSs produce contrastively focused pronouns frequently without stress, it is a topic for further research whether there is a relation between the absence of prosodic prominence and the pro-drop features of the languages in question (see Özsoy et al. 2025 [this volume]).

In sum, English MSs and HSs speaking majority English do not show qualitatively different patterns in the realization of stressed pronouns: both speaker groups stress pronouns in spontaneous speech, dependent on both information structure and prosodic phrasing (for majority German, see Zerbian & Böttcher 2019, Böttcher & Schubö 2021). However, there is a difference in distribution of unstressed pronouns: HSs do not always stress contrastively-focused pronouns in majority English and thus do not follow the pragmatic constraints that hold in English in all cases.

The second line of research investigated the relationship between prosody and contrastive focus of adjectives. The prosodic realization of modified NPs, such as *a blue car*, is dependent on focus and contrast, as shown in examples (2a-d). The placement of a pitch accent on the noun (indicated by small caps) allows for various focus interpretations while a pitch accent on the adjective (as in 2c) can only indicate narrow or contrastive focus on the adjective (Ladd 2008).

- (2) Pitch accent placement in modified noun phrases and its interpretation
  - a. wide/broad focus: Along came [a white CAR]<sub>FOCUS</sub>.
  - b. contrastive focus on noun: There was a white [VAN]<sub>FOCUS</sub> and a white [CAR]<sub>FOCUS</sub>.

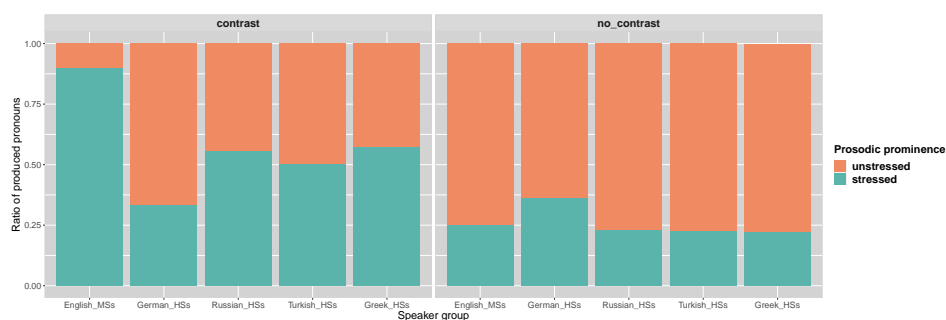


Figure 1: The occurrence of unstressed and stressed pronouns (ratio and in numbers) in the case of presence or absence of contrastive focus in English across different speaker groups in RUEG corpus 0.3.0

- c. contrastive focus on adjective: The [BLUE]<sub>FOCUS</sub> car crashed into the [WHITE]<sub>FOCUS</sub> car.
- d. contrastive focus on adjective and noun: The [BLUE]<sub>FOCUS</sub> [CAR]<sub>FOCUS</sub> rammed the [WHITE]<sub>FOCUS</sub> [VAN]<sub>FOCUS</sub>.

Böttcher (2022) investigated the accent placement in noun phrases containing adjectives in English. Data analysis included 144 narrations from the RUEG corpus (0.4.0) by adolescent English MSs and Greek and Russian HSs speaking majority English. Each group comprised 24 speakers, each producing two narratives. Narratives were searched for modified noun phrases including the lexeme *car* and were analyzed for semantic contrast (Götze et al. 2007) and for the location of pitch accents and pitch accent type (ToBI guidelines, Beckman et al. 2006). As expected, the contrastive adjectives were accentuated in most cases (Figure 2). Surprisingly, adjectives were also accented when not explicitly expressing a contrast, which might still reflect the speaker's contrasting perspective (Kaland et al. 2014). Additionally, contrastively accented adjectives were also frequently followed by an accented noun. Such a double accent pattern was found more frequently in formal speech and in the speech of Russian HSs. The double accent also occurred frequently in the Russian heritage language of these HSs indicating a possible transfer (see Zerbian et al. 2025 [this volume], 2022). The analysis of pitch accent types as high (H\*) or rising tone accents (LH\*) did not show differences across speaker groups (see Böttcher 2022 for further detail).

The results of this study reveal that English MSs and HSs speaking majority English do not show qualitatively different patterns in the prosodic realization of contrastively focused adjectives: the adjectives generally carry a pitch accent, although at the same time double accent patterns are more frequently produced

by Russian HSs. In the light of the higher use of double accents in formal narra-  
tions by MSs and Greek HSs, this can possibly be interpreted as an effort towards  
being linguistically – in this case prosodically – more explicit in the context of  
formal situations (see also next section, for HSs see Polinsky 2018: 144).

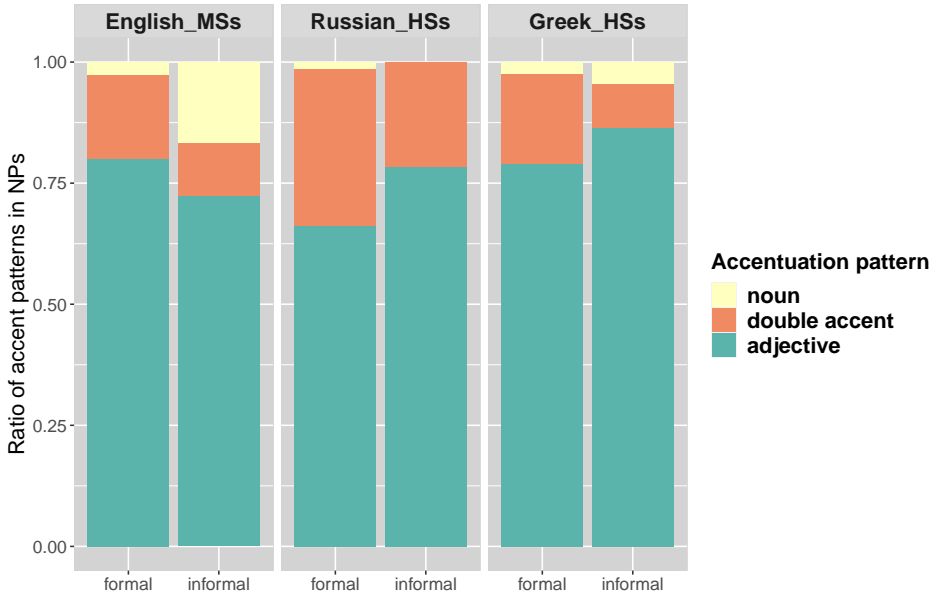


Figure 2: Distribution of accent patterns within modified noun phrases containing a contrastively focused adjective in English across speaker groups and communicative situations in the RUEG corpus 0.4.0

Overall, our results suggest that English MSs and HSs speaking majority English use similar prosodic strategies in realizing given pronouns and contrastive focus on adjectives, thus using available linguistic features along a continuum (cf. Wiese et al. 2022). Given that both structures are at the interface of prosody and semantics/pragmatics, the Interface Hypothesis (Sorace 2011) predicts these areas to be vulnerable and thus prone to change in language contact. However, our results do not lend strong support for this, at least in the ML.

## 4 Reference

Previous research has indicated that bilinguals might be more explicit than monolinguals in the domain of reference. In referent introduction, Barbosa et al. (2017)

have found that simultaneous bilingual children lexicalize more concepts in their narratives than monolinguals. In referent tracking, adult late L2 speakers tend to use full NPs instead of pronouns, compared to their L1 (Hendriks 2003, Gullberg 2006) and compared to L1 speakers of their L2 (Hendriks 2003, Yoshioka 2008). In addition, adult late L2 speakers and simultaneous bilingual children tend to over-produce and over-accept overt pronouns instead of null forms in pro-drop languages, compared to monolinguals (Serratrice et al. 2004, Sorace & Filiaci 2006, Sorace et al. 2009). Finally, adult Turkish HSs use more overt pronouns in their majority Dutch (a non-pro-drop language) compared to Dutch MSs (Azar et al. 2020). However, not all previous findings support this trend. Contemori et al. (2023) and Contemori & Ivanova (2021) found that Spanish HSs produced fewer nouns and more pronouns in their majority English compared to English MSs, and hence were less explicit.

Despite the disagreement in the literature, Pashkova and colleagues hypothesized that HSs might be more explicit either due to their frequent communication with L2 English speakers, who might benefit from extra detail (Polinsky 2018: 144) or due to more frequent input from L2 English speakers, who themselves might be more explicit (for results in majority Dutch see Azar et al. 2020). We conducted two studies to test this hypothesis: on referent introduction (Pashkova et al. 2020) and on referring expressions (Pashkova & Allen 2025).

In the referent introduction study, Pashkova et al. (2020) compared the number of introduced referents in a subset of the English subcorpus: 32 adolescent English MSs and 120 HSs (27 German, 31 Greek, 32 Russian, 30 Turkish). We identified 19 frequent referents, and coded each referent as introduced (lexicalized) or not in each narrative. Even though referent introduction is affected by various extra-linguistic factors (e.g., attention and memory play a role in which characters/objects will be mentioned in a narrative), it is an important first step of establishing reference (Vogels et al. 2019: 1). A higher proportion of introduced referents would be a sign of higher explicitness since the narrative would provide more detail about the events. Contrary to the hypothesis, a linear mixed model analysis showed no significant difference between HSs and English MSs in the number of introduced referents in majority English. In addition, we observed that HSs and MSs had the same effect of formality: both speaker groups introduced more referents in the formal setting than in the informal one.

The study of referring expressions (Pashkova & Allen 2025) consisted of two parts: modification of referring expressions and their form. In the modification part, we were interested whether HSs speaking majority English would modify referring expressions more often than English MSs, for instance, with a prepositional phrase (*the woman with the dog*) or a relative clause (*the woman who had a*

dog). A higher proportion of modified referring expressions would be an indication of higher explicitness because it would provide more information about the referent of the NP. We conducted a linear mixed effects model analysis based on the full English subcorpus. Contrary to our expectations, there was no significant difference between HSs and English MSs in the proportion of modified referring expressions. In addition, we observed that HSs and MSs had the same effect of formality: both speaker groups modified referring expressions more often in the formal setting than in the informal one.

In the part concerning the form of referring expressions, we hypothesized that HSs speaking majority English would be more explicit than English MSs in their use of expressions for referents that have been previously introduced; new referents are typically introduced by noun-headed NPs, so they do not exhibit much variation in referent form. We examined three types of referring expressions: noun-headed NPs (*the woman had a dog*), pronouns (*she also had some groceries*), and null anaphora (*she got scared and Ø dropped her groceries*). Each referent from the list of 19 frequent referents was annotated for its type of referring expression, irrespective of its syntactic position (i.e., subject, object, object of a preposition, etc.). Following Azar et al. (2020), we conducted two comparisons: (1) noun-headed NPs vs. pronouns and (2) pronouns vs. null anaphora. We expected that HSs would use more noun-headed NPs than MSs in the first comparison and more pronouns in the second comparison.

The first comparison was performed on the full English subcorpus, where each referent was annotated for its type of referring expression. The second comparison was done on a subset of the English subcorpus that included 40 English MSs, 42 Turkish and 40 Russian HSs (each speaker group equally split between adolescents and adults). The subsetting was done due to the necessity to annotate each referent for the type of clause it was used in – main or subordinate – since null anaphora can occur only in coordinate main clauses.

The confirmatory analyses with linear mixed effects models did not indicate a significant difference between MSs and HSs combined, neither in the noun-headed NP vs. pronoun comparison nor in the pronoun vs. null anaphora one. This shows that HSs in our sample are not more explicit than MSs across the board.

However, an exploratory analysis of noun-headed NP vs. pronoun referring expressions showed an interesting pattern that we had not previously hypothesized. This pattern has to do with formality and with a distinction that speakers make between two types of referent tracking – maintenance and reintroduction, which in our study are defined based on the syntactic roles of the referent in the current and previous finite clauses (following Hickmann & Hendriks 1999,

Serratrice 2007, Perniss & Özyürek 2015). The definitions of maintenance and reintroduction are presented in Table 4, along with examples.

Table 4: Definitions and examples of maintenance and reintroduction

Previous clause	Current clause	Referent tracking type
Subject <i>The dog ran.</i>	Any syntactic role <i>The driver saw it.</i>	Maintenance: Referent <i>dog</i> in current clause is maintained
Non-subject <i>The driver saw the dog.</i>	Non-subject <i>He didn't want to hit it.</i>	Maintenance
Non-subject <i>The driver saw the dog.</i>	Subject <i>The dog ran.</i>	Reintroduction: Referent <i>dog</i> in current clause is reintroduced
Absent <i>The driver turned right.</i>	Any syntactic role <i>He saw the dog.</i>	Reintroduction

The exploratory analysis individually compared German, Greek, Russian, and Turkish HSs to English MSs, as opposed to comparing all HSs together to MSs in the confirmatory analysis. Although there was considerable variability across speakers, the results (Figure 3) indicated two trends: first, in the formal setting, there was a larger difference between the proportions of full NPs in maintenance and reintroduction in narratives by German, Russian and Turkish HSs compared to English MSs (difference of 0.54 for German HSs, 0.48 for Russian HSs, 0.53 for Turkish HSs vs. 0.41 for English MSs). Second, in the informal setting, Russian and Turkish HSs produced significantly more NPs both in maintenance and reintroduction compared to English MSs. German and Greek HSs followed the same pattern but the difference did not reach significance.

These findings suggest that HSs taken together as one group are not more explicit than MSs across all contexts. The situation is more nuanced: when speaking or writing formally, some HS groups tend to separate maintenance and reintroduction more sharply than MSs. This is similar to a trend that we will present in Section 6 – there, heritage speakers show a tendency of a sharper distinction between formal and informal settings in their use of subordinate clauses and left dislocations, in comparison to MSs.

However, HSs also show a different tendency when speaking or writing informally: some HSs are more explicit in that they use more full NPs than MSs both

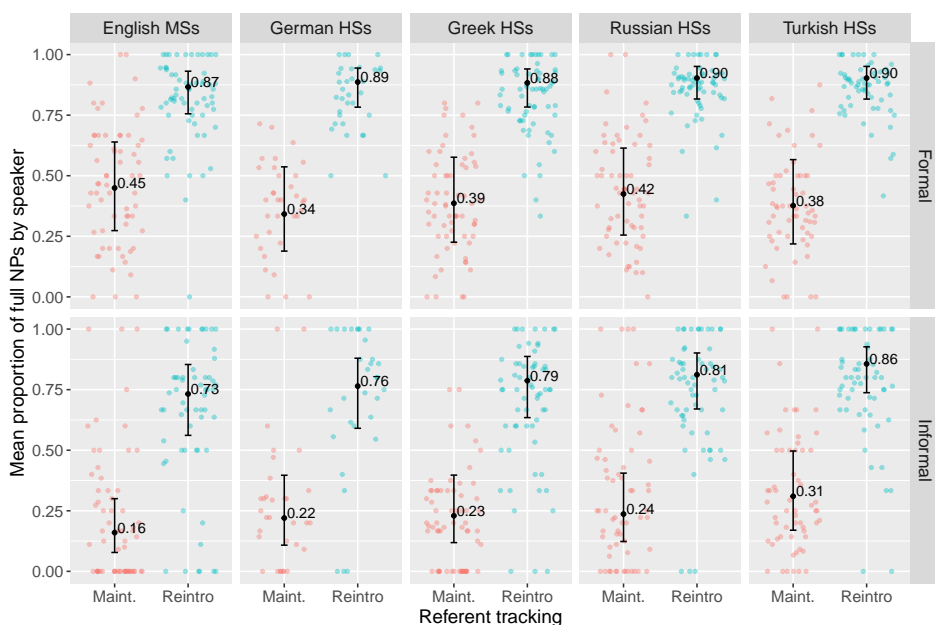


Figure 3: Predicted probabilities and individual proportions of noun-headed NPs by speaker group, formality and referent tracking. Predicted means and CIs are based on a linear model, individual proportions on raw data.

in maintenance and reintroduction. Overall, we see that HSs employ different strategies in the formal and informal settings: in a formal situation, they highlight the difference between maintenance and reintroduction, and in an informal situation they prefer to be more explicit regardless of the referent tracking status. We will come back to this trend in the general discussion.

## 5 Article use

We investigated article use by HSs speaking majority English and English MSs in narrative data from the English subcorpus and in an additional experimental study focusing specifically on article choice. The study on article use in narratives compared the proportions of unexpected articles in new and given referents based on the full English subcorpus (Wiese et al. 2022: 10–14). We extracted all strictly new referents, that is, referents that were mentioned for the first time and did not have any related anchor in the previous discourse (such as *the car* would be for *the driver*). We also extracted all given referents that were already



mentioned in the previous discourse. Next, we identified the referents that featured unexpected article choice, that is, “*the* + new” and “*a* + given” referents. All other referents, including the most expected (“*a* + new” and “*the* + given”), were marked as “other.” Comparing HSs to MSs, we observed no effect of bilingualism: the two groups produced similar proportions of unexpected “*the* + new” referents out of all new referents and of unexpected “*a* + given” referents out of all given referents. In addition, we found that HSs and MSs were influenced by spoken/written mode in a similar fashion: irrespective of the speaker group, there were significantly more unexpected “*the* + new” and “*a* + given” referents in the spoken mode than in the written mode. Overall, this study indicated no evidence of differences between HSs and English MSs in majority English, with both groups producing a small proportion of unexpected articles with new and given referents, especially in the spoken mode.

The second study on article use (Pashkova & Allen 2022) featured more constrained contexts and is the only study mentioned in this chapter that did not use the RUEG corpus data. Speakers were given a fill-in-the-blank task eliciting determiners, where each item was a short dialogue between two people (adopted from Ionin et al. 2008). This study tested the Fluctuation Hypothesis (originally developed for late L2 speakers; Ionin 2006) on adult and adolescent HSs speaking English as their ML. The Fluctuation Hypothesis posits that L2 English speakers whose L1 does not have definiteness-based articles (e.g., Korean) fluctuate between definiteness and specificity in their article choice in English: sometimes they attend to definiteness of the NP, and sometimes to specificity, resulting in productions that are different from those of L1 English speakers. On the other hand, L2 English speakers whose L1 has definiteness-based articles (e.g., Spanish) are expected to perform similarly to L1 English speakers and not exhibit any fluctuation.

HSs are similar to late L2 speakers because both speaker groups learn English while having experience of using an additional language (HL for HSs, and the only L1 for L2 speakers), so similar transfer effects could be observed in HSs as well. We hypothesized that Russian HSs in our sample ( $n = 63$ ) would exhibit the most fluctuation since their HL does not have articles, followed by Turkish HSs ( $n = 58$ ), whose HL has a complex definiteness- and specificity-based differential object marking system. Finally, we predicted that Greek HSs ( $n = 78$ ) would be the most similar to English MSs ( $n = 46$ ) since Greek has definiteness-based articles.

The results did not support the Fluctuation Hypothesis: none of the HS groups exhibited the fluctuation pattern attested in late L2 speakers. As a sole result, this would be quite unsurprising since exposure to English by the HS begins much

earlier than that of typical late L2 speakers. However, we also observed two unpredicted patterns: first, all speaker groups, including MSs, produced more unexpected articles (*the*) in the indefinite contexts (slightly above or under 1% in indefinite contexts, close to zero in definite ones). Second, Turkish HSs were the only bilingual group significantly different from English MSs: they produced more unexpected articles in both definite and indefinite contexts (*the* in definite and *a* in indefinite).

The first unpredicted pattern is similar to the finding from the narrative study described above: there the percentage of unexpected articles was also numerically higher for the new referents (= indefinite contexts) than for given ones (= definite contexts), even though this comparison was not statistically tested. However, it is hard to attribute these two results to the same cause since the narrative task and the fill-in-the-blank task were quite different. In the narrative task, it is plausible that speakers fell back to their own perspective of being able to uniquely identify the referents (since they were familiar with the characters and objects in the story after watching the video) and did not consider the addressee's perspective. This is especially likely given the fact that unexpected articles were produced more often in the spoken than in the written mode because the spoken mode does not offer a possibility to plan ahead and revise. In the fill-in-the-blank task, this explanation does not seem to apply: first, the task was written and untimed, removing any performance pressure, and second, the speakers did not have a perspective of being familiar with the referents: they were reading the items for the first time. In this case, a possible explanation would be that speakers found the task too easy and engaged in creating background contexts that would justify the use of definite articles in indefinite contexts (e.g., prior shared knowledge or experience between the interlocutors in the dialogues). This would explain why indefinite articles were rare in definite contexts: once the uniqueness of the referent has been established, it seems almost impossible to create a context where it would disappear.

The second unpredicted pattern, the difference between Turkish HSs and English MSs, is not straightforward to explain. One possibility is that Turkish HSs engaged in creating the background contexts that justify definite articles more often than any other group. Another potential account is that acquiring a definiteness-based system in English along with a definiteness- and specificity-based system in Turkish led to the formation of a unique system that is different from both. This finding requires further investigation to see if the pattern persists in other speaker samples, and if so, to understand its underlying cause.

Summing up, in the area of article use we observed no difference between the HSs and English MSs in majority English when examining the narrative data,

and more unexpected articles produced by Turkish HSs than English MSs in an experimental task (with no difference between other HS groups and MSs). At the same time, all speakers, including English MSs, produced a small proportion of unexpected articles in indefinite contexts, which might point to task effects.

## 6 Syntax

In the domain of syntax, we compared HSs speaking majority English and English MSs for three phenomena – the use of various clause types across registers, the use of left dislocations across registers, and word order variation caused by information status of referents.

We conducted two studies on clause types across registers. Pashkova et al. (2022) compared the use of independent main clauses (*I was walking down the street. I saw a couple.*) and coordinate main clauses (*I was walking down the street, and I saw a couple.*) in a subset of the English subcorpus: 20 adolescent English MSs and 20 German HSs. Since previous research has focused on clause-type variation mainly in the HL of HSs (Keskin 2025 [this volume], Özsoy et al. 2022, Schleppegrell & Colombi 1997, Sanchez Abchi & de Mier 2017), we aimed to create a more comprehensive picture of HSs' repertoires and examine their ML as well.

We did not observe a significant difference between German HSs and English MSs: both groups used independent main clauses and coordinate main clauses with similar frequencies. In addition, HSs and MSs alike used more independent main clauses in written narratives than in spoken ones, and more coordinate main clauses in spoken narratives than in written ones. Finally, both speaker groups showed the same effect of formality: they used more coordinate main clauses in informal settings than in formal ones.

The second study on the use of clause types (Tsehaye et al. 2021) looked into the use of subordinate clauses in a larger subset of the English subcorpus: 32 adolescent English MSs and 27 adolescent German HSs. We found that bilinguals and monolinguals used subordinate clauses differently depending on the formality and mode (Figure 4). German HSs made a distinction between the formal and informal setting in both spoken and written modes, while English MSs made a distinction between the settings only in the written mode.

Our second syntactic structure of interest is left dislocation (LD), in which a noun phrase precedes its core clause and the argument position within the core clause is filled with a pronoun that is co-referential with the dislocated noun phrase (e.g., *The woman<sub>NP</sub>, she<sub>PRONOUN</sub> spilled her groceries*). LDs are indicative of an informal register (Keenan 1977, Geluykens 1992; see Bunk et al. 2025 [this volume] for a comparative review of LD in English, German, and Russian).

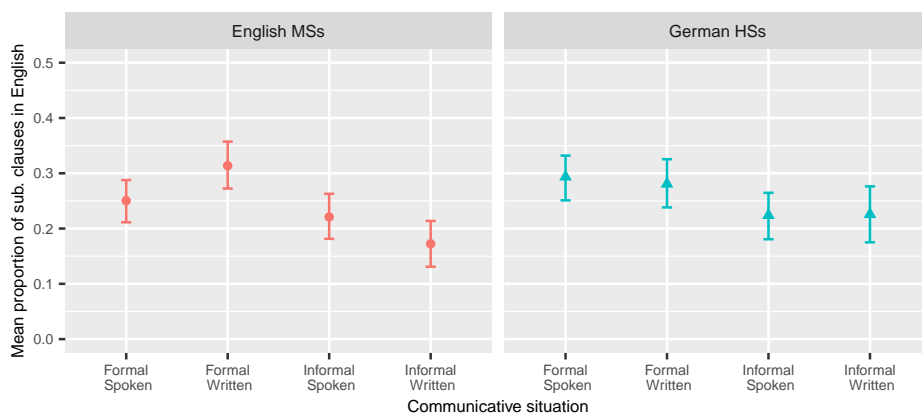


Figure 4: Mean proportions of subordinate clauses in English by speaker group and communicative situation. Means and bootstrapped CIs are based on raw data.

Pashkova et al. (submitted) hypothesized that the use of LD in the majority English of HSs might be different than that of English MSs because previous studies have found differences in the frequency of LD use by French-English bilinguals and English MSs (Hervé et al. 2016, Tagliamonte & Jankowski 2019). We also hypothesized that HSs may approach the formality constraint on LD usage differently from English MSs, given that divergent approaches to formality have been observed in other register-dependent phenomena, such as subordinate clauses (Tsehaye et al. 2021, see above).

We investigated the use of LD in the full English subcorpus. The findings reveal that in the formal setting, Greek and Turkish HSs use fewer LDs than English MSs, while German and Russian HSs do not show any significant difference from the MSs. In the informal setting, Turkish HSs use more LDs than English MSs, while Greek, German, and Russian HSs do not demonstrate any significant difference from the MSs.

As to the difference between the formal and informal settings, Greek and Turkish HSs have a dissimilar approach to the formality distinction in comparison to English MSs. These HSs use slightly more LDs in informal than formal settings, whereas English monolinguals use slightly fewer LDs in informal than formal settings. That is, English monolinguals have a reverse formality effect compared to Greek and Turkish HSs. In fact, Greek and Turkish HSs adhere more strictly to the expected pattern of using LDs predominantly in the informal context. German and Russian HSs do not differ from English MSs statistically, even though

numerically they pattern with Greek and Turkish HSs because they have slightly more LDs in the informal setting than in the formal one (Figure 5). Also noticeable in Figure 5 is that Turkish HSs differentiate the formalities quite strictly, while English MSs and other HS groups do not show such a strict differentiation.

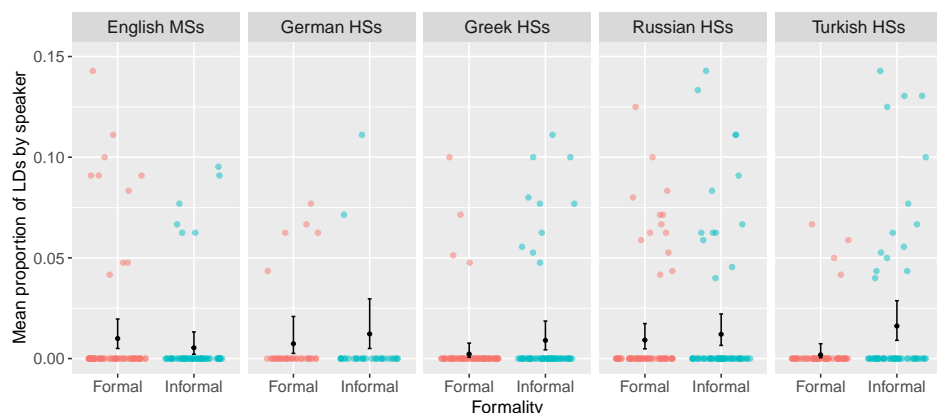


Figure 5: Predicted probabilities and individual mean proportions of LDs by speaker group and formality. Predicted means and CIs are based on a linear model, individual proportions on raw data. The y-axis is zoomed from the original size of 0–30% to the size of 0–15% to make the model predictions more visible. When zooming, 10 data points that were above the 15% mark were removed: two in English MSs, three in German HSs, two in Russian HSs, and three in Turkish HSs.

In summary, the LD study shows that some HS groups use LDs differently in their majority English compared to English MSs, while other groups do not. The difference does not lie in the overall frequency of LDs, but in the way speakers distribute LDs in formal and informal settings: some HSs are actually more in line with the expected patterns reported in the literature than English MSs. It is a matter of future research to determine the source of this difference: while we cannot exclude cross-linguistic influence from the Greek and Turkish Hs, a general HS pattern seems more likely for two reasons. First, it is quite surprising that cross-linguistic influence from two Hs would make HSs' productions more aligned with the expected LD pattern in majority English. Second, all four HS groups trend in the same direction numerically, which is also somewhat unlikely if the HL influence was the source of difference.

The third phenomenon in the domain of syntax that we examined is the variability in word order caused by the information status of referents (see Bunk

et al. (2025 [this volume]) for more on word order dynamics in language contact and Tsehay et al. (2025 [this volume]) for word order phenomena in heritage German). English follows a fixed SVO word order, which can be changed to accommodate pragmatic factors, leading to various information packaging constructions (Huddleston 2002: 67–68). One such factor is the given-before-new principle, which suggests that a clause should begin with given/background information and end with new information (Biber et al. 2021: 888–889). However, this principle creates a conflict when it comes to new subjects in English. While the SVO word order demands that a new subject be placed at the beginning of a clause, it would be more helpful to position it closer to the end because the referent is new.

To address this conflict, English speakers may use different constructions that deviate from the SVO order, such as existential *there* (e.g., *There was a woman unloading her groceries on the other side of the street*), presentational *there* (e.g., *There stood a woman unloading her groceries on the other side of the street*), locative inversion (e.g., *On the other side of the street was a woman unloading her groceries*), or passivization (e.g., *Some groceries were being unloaded by a woman*).

Research suggests that bilingual English speakers may use information-packaging constructions differently than English MSs, potentially due to cross-linguistic influence. For instance, speakers of Singapore or Jamaican English use fewer existential *there* constructions than speakers of monolingual varieties (Winkle 2015).

Building on these previous findings, Pashkova & Allen (in preparation) investigated the syntactic structures employed for the introduction of new subjects in majority English. We asked whether HSs of Russian and Turkish, languages with more flexible word orders, utilize more non-SVO structures than English MSs, potentially due to cross-linguistic influence from their heritage languages. To test this hypothesis, we compared the syntactic patterns employed for the introduction of new subjects, based on a subset of the English subcorpus. It included 82 HSs (40 Russian and 42 Turkish) and 40 English MSs. Each subject was annotated for its information status (new or given) as well as the syntactic structure it appeared in. The syntactic structures were classified into two categories: SVO-type structures (SVO and non-inverted copular clauses), and non-SVO-type structures, encompassing existential and presentational *there*, locative and non-locative inversions, right and left dislocations, passives, questions, and *it*- and pseudo-clefts.

The results revealed no significant difference in the frequency of non-SVO structures employed by Russian and Turkish HSs and English MSs. However, a distinction emerged in the frequency of non-SVO structures in new and given

subjects: new subjects appeared in non-SVO structures more often compared to given subjects. There was no interaction between the subject information status and the speaker group (HSs vs. MSs). These findings indicate that new subjects are more commonly associated with information packaging constructions (i.e., non-SVO structures) than given subjects, and both HS and MSs in our sample follow this trend in majority English in a similar way.

Summing up, in the domain of syntax we have not observed qualitatively different patterns in HSs' and MSs' productions. If there are dissimilarities between bilinguals and monolinguals, they are connected to different use of syntactic structures in formal and informal contexts (subordinate clauses in German HSs, LDs in Greek and Turkish HSs).

## 7 Discourse

Discourse organization refers to the way in which speakers organize their ideas in a coherent way to convey a meaningful message. This is typically achieved through the use of various linguistic devices such as discourse markers, as well as through macro-level textual components such as openings and closings. Discourse organization can vary across countries and cultures, social groups, age groups, genre, speech communities, and registers, with sometimes very subtle differences (e.g., Fox Tree 2010). Previous studies have shown that HSs who are highly proficient in both their HL and ML exhibit similar patterns of discourse organization as MSs (Silva-Corvalán 2014), while those who are less proficient in one or both languages often display different patterns (Montrul 2010).

To examine this further, we investigated whether the Greek, Turkish, German and Russian HSs use similar organization strategies as English MSs in their majority English, and whether any differences are a result of cross-linguistic influence. We focus on both openings and closings and discourse markers in the English subcorpus of the RUEG corpus (see Katsika et al. (2025 [this volume]) for an analysis of openings and closings in various languages and Labrenz et al. (2025 [this volume]) for an analysis of discourse markers).

Discourse openings and closings are defined as any material preceding or following the core phase of the discourse (for English, see Schegloff & Sacks 1973, Schegloff 1968, 2007; for other languages, see Pavlidou 2014, Luke & Pavlidou 2002). By core phase, we mean the part of the text that constitutes the narrative, which in our case is the description of the events presented in the video. They are used to define textual boundaries that set the main text apart from the framing, introduce or close off a new topic, orient the addressee to what is coming next,

make the speaker's stance clear, and interpret and evaluate linguistic material (Labov 1972, Tolchinsky et al. 2002, Berman & Katzenberger 2004).

We identified the functions of openings and closings within three main categories: intersubjective (oriented towards the imaginary interlocutor), subjective (oriented towards the speaker), and textual (focusing on the organization of the text). Intersubjective, subjective and textual functions and their respective examples are presented in Table 5.

Table 5: List of intersubjective, subjective and textual functions with examples

Functions	Examples
<i>Intersubjective</i>	
Greeting	<i>hello</i>
Summoning	<i>bro</i>
Initial inquiry	<i>How are you?</i>
Identification	<i>My name is...</i>
Justification	<i>I am calling to report an accident</i>
Attention getter	<i>You wouldn't believe what happened!</i>
Reaction seeking	<i>Can you call me back?</i>
Inquiry response	<i>I'm fine too!</i>
Give advice	<i>Be careful when you're driving</i>
Common ground	<i>Remember that day...</i>
Valediction	<i>bye</i>
<i>Subjective</i>	
Reaction	<i>whoa</i>
Evaluation	<i>Totally crazy</i>
Personal statement	<i>I was walking as I always do in the mornings</i>
<i>Textual</i>	
Initialize narrative	<i>Here's what I saw:</i>
Contextual	<i>I just saw an accident</i>
Episodic	<i>It happened at 12pm today</i>
Resolution	<i>Nobody got hurt</i>
Encapsulation	<i>It was all an accident</i>
Coda	<i>That's the end</i>

A descriptive frequency analysis revealed that all groups of speakers in our



sample used more openings (~80%) than closings (~50%) across all communicative situations in the majority English texts. The overwhelming use of an opening in speakers' narrations indicates that openings were used as a means to establish the communicative situation (e.g., formal vs. informal). In contrast, the relatively low percentage of closings shows that closings have a more optional character in speakers' narratives.

German and Russian HSs showed the highest frequency of openings and closings, while Turkish HSs showed the lowest frequency including many texts in which they only described the video events with no introductory or closing material (Figure 6). This indicates that although English is the dominant language for all speaker groups, there are differences in macro discourse strategies across groups which may originate from contact with the HL (see Katsika et al. (2025 [this volume]) for analyses of openings and closings in Turkish MSs and HSs) or from other sources that may be uncovered in further research. In terms of functions, we find that intersubjective functions are mostly used in informal settings whereas textual functions are primarily used in formal settings.

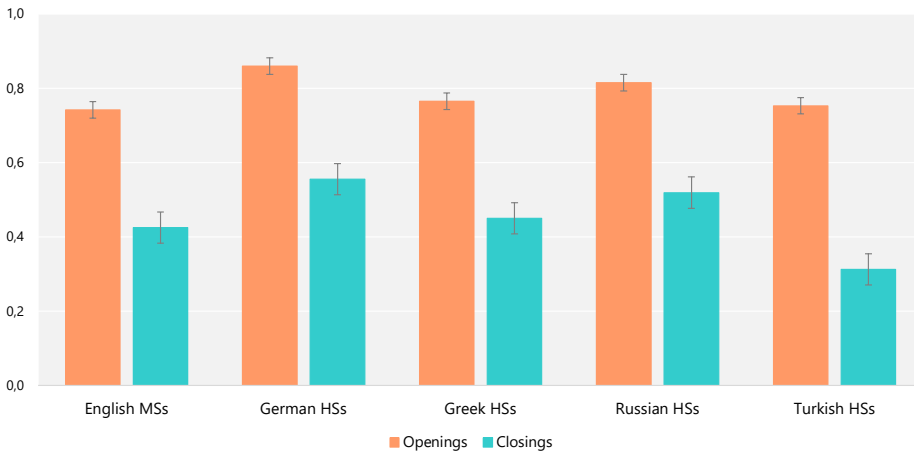


Figure 6: Frequency distribution of openings and closings in the English texts. Error bars represent standard errors. Means and error bars are based on raw data.

Discourse markers also play a crucial role in the organization of discourse. We analyzed the frequency and discourse functions of two discourse markers in the full English corpus: *so* and *yeah*. The literature reports that *so* as a discourse marker is mostly used with an inferential function to indicate “result” or “inference” (Schiffrin 1987). The MSs in our corpus used *so* with a higher frequency than all groups of HSs in majority English, although the functions of *so*

are distributed similarly across all speaker groups (see Labrenz et al. (2025 [this volume]) for a more detailed analysis of the German *so* and other discourse markers). In terms of function, *so* was primarily used across groups as a connector (3a), then to initialize the narrative (3b), and third most frequently to switch the theme (3c).

- (3) a. it was cause this guy like lost control of his soccer ball and I know you love soccer *so* thought you'd be interested (USmo11FE\_isE)
- b. *so* i just witnessed an accident it seemed like a minor accident there was a woman on the right side of the street opening her trunk (USmo05FE\_isE)
- c. and then a dog starts barking and he runs in the street *so* this guy is driving down the street and he sees a dog (USmo10ME\_isE)

The analysis of the discourse marker *yeah* showed that its frequency is much more restricted than that of *so*. Interestingly, *yeah* is used most frequently by German HSs, followed by MSs, and then Greek and Turkish HSs. The elevated use of *yeah* by German HSs may stem from the high frequency of *ja* in the German language, as well as the functional similarity between *yeah* and the German *ja* ('yeah') (Trotzke et al. 2020).<sup>2</sup> The most frequent discourse functions of both *ja* and *yeah* across groups are initializing and ending the narrative, evaluation, and hesitation. (4) shows an example of *yeah* ending a narrative.

- (4) the car stop and hit the other car so it's really all that happened um *yeah* (USmo69ME\_isE)

In sum, we see that the patterns of use of discourse openings and closings in majority English are similar across HS and MS groups, although the frequency of use differs somewhat. This extends to the discourse markers *so* and *yeah* as well; both are used with similar functions across groups but with somewhat different frequencies. In the case of *yeah*, this appears to be driven by cross-linguistic influence.

## 8 Discussion

We began this chapter by asking whether the majority English of adolescent and adult HSs showed any dynamic patterns due to bilingualism or whether it looked

---

<sup>2</sup>Although there is phonological similarity between the English *yeah* and the Turkish discourse marker *ya*, we did not find inflated use of *yeah* in Turkish HSs. This possibly stems from the fact that the English *yeah* and the Turkish *ya* do not have full functional overlap (as is the case between the German *ja* and the English *yeah*) (Şimşek 2012).

similar to patterns found in English MSs. In short, we found that the four groups of HSs patterned remarkably similarly to each other and to the English MSs in most respects, with some quantitative differences. We elucidate these similarities and differences in the following paragraphs.

## 8.1 Similarities

We begin with the similarities between HSs and MSs in majority English. In the area of phonology, HSs performed like MSs in their use of stressed pronouns based on information structure and phrasing, and in their use of pitch accents marking contrastive adjectives prosodically.

In terms of reference, they introduced and modified new referents with the same frequency as MSs, and, similarly to MSs, introduced and modified more referents in formal than informal settings. They also patterned like MSs in the proportion of NPs vs. pronouns and pronouns vs. null forms that they used to realize already-introduced referents, as well as in using more NPs for referent re-introduction than for referent maintenance.

In the domain of article use, HSs produced unexpected structures at the same rate as MSs (regarding use of both *a* for definite and *the* for indefinite), and produced more unexpected articles in indefinite than definite contexts and in spoken than written modes (both with similar frequency to MSs). For syntax, they produced independent vs. coordinate main clauses at similar frequencies to MSs, and patterned like MSs in preferring independent main clauses for the spoken mode vs. coordinated main clauses for the written mode, and in producing more coordinated main clauses in informal than formal settings. Additionally, HSs used a similar frequency of non-SVO structures to introduce new subjects compared to MSs. Finally, HSs structured their discourse similarly to MSs: all of the speaker groups produced more openings than closings, and preferred inter-subjective functions for openings and closings in the informal setting vs. textual functions in the formal setting. The functions of the discourse markers *so* and *yeah* were also distributed similarly across the HS and MS groups. In sum, we found no differences between our four HS groups and English MSs in many patterns of majority English.

It may seem surprising that our results show so many similarities in majority English between HSs and MSs, given a large number of differences found in the literature reviewed in the Introduction. We offer several possible reasons for this. One potential reason is that our studies focus on a particular type of heritage speaker for whom many demographic factors favor a positive language learning situation. They were all fluent and literate in the HL (which is not the case in

several studies), typically from mid to high SES families (50–89% of mothers in each speaker group had a postsecondary degree), and without indication of traumatic immigration history in their recent background (e.g., refugee from a war zone). In addition, 91% of them started acquiring English from the age of 5 or earlier, and nearly half from birth, so they had very early exposure to the ML in a society where the ML is highly dominant.

Another main factor relates to the methods of our study. Most of our data derive from naturalistic narrative productions, for which the cognitive demands are relatively low. The speaker is also in control of the language they produce, so they can easily keep to language that is comfortable for them. It may be that experiments or other tasks with higher cognitive demands, or focusing more narrowly on a particular structure, would provide more possibility to see dynamic patterns resulting from bilingualism. For example, Bylund et al. (2012) used materials designed to be particularly challenging for the HS participants, and Lee et al. (2011) and Scontras et al. (2017) focused narrowly on differences in scope phenomena in Korean/Mandarin vs. English that appear relatively rarely in narrative contexts such as those we elicited. Further, we examined only a subset of possible structures; examining other structures may lead to different results.

In general, however, we have added evidence to the literature that several structures in bilinguals' more dominant language (here, ML) are not influenced by cross-linguistic influence or general bilingualism effects. In addition, we have shown that for at least our sample of HSs, bilingualism is not an impediment to performance in the majority language that is consistent with that of MSs (Flege 1999, Ventureyra et al. 2004, MacWhinney 2005). This suggests that there is no reason to prohibit heritage speakers from maintaining their HL, and supports the view that heritage speakers are indeed native speakers of the ML rather than "deficient versions" of speakers of the ML (as already expressed in Rothman & Treffers-Daller 2014, Montrul 2016, Kupisch & Rothman 2018, Tsehay et al. 2021, Wiese et al. 2022).

## **8.2 Differences**

However, the productions of HSs in majority English also differed significantly in several respects from those of English MSs. Many of the differences are consistent with the idea put forth in the literature that HSs might be more explicit than MSs. In the domain of phonology, Russian HSs used double-accented structures in contrastive accent situations – accenting both the adjective and the noun – where MSs and Greek HSs typically accented only the adjective. The occasional use of double accent by MSs and HSs of Greek in formal situations gives an

indication that this phenomenon could be related to an effort to be prosodically more explicit by making the structure more salient (e.g., producing an accent on more lexical constituents than expected). Both patterns suggest that they were trying to make the structures more salient than MSs did. In referent use, Russian and Turkish HSs used more NPs in the informal setting than MSs – again a sign of greater explicitness. In one area, however, we found that HSs were less explicit than MSs: they stressed fewer contrastively-focused pronouns than MSs.

In other cases, the HS bilinguals seemed to make a stricter distinction between different linguistic and extra-linguistic contexts than MSs did. In the area of reference, for example, German, Russian, and Turkish HSs had a larger difference in the proportion of NPs between maintenance and reintroduction of referents in the formal setting than MSs did, signaling a stricter distinction between these functions. We also found an exaggerated distinction in HSs vs. MSs in their use of subordinate clauses depending on the formality and mode: German HSs distinguished between the formal and informal setting in both spoken and written modes, while MSs only made a distinction between the settings in the written mode. Finally, when producing left dislocation structures, Turkish HSs differentiated between the formal and informal settings more than MSs did.

Related to the stricter distinction between extra-linguistic contexts, in some situations HSs appeared to follow the differentiation between the contexts that is expected from the literature more strictly than MSs. In the case of left dislocations, all HS groups used more LDs in the informal setting than in the formal one (in line with the literature), while MSs had the reverse pattern – they produced slightly more LDs in the formal setting than the informal one. While the reversal of the pattern is visible in all HSs, it reached significance only in the Greek and Turkish HSs (see similar results for left dislocations in majority German in Bunk et al. (2025 [this volume])).

The finding in our data that HSs are more explicit in some contexts and tend to differentiate contexts more strictly is striking, and also consistent with other reports in the literature noted earlier (e.g., Serratrice et al. 2004, Barbosa et al. 2017, Polinsky 2018, Azar et al. 2020). It suggests that the dominant language of bilinguals can show signs of general bilingualism effects, in addition to potential cross-linguistic influence from the less dominant language.

Why might this be the case for our data? It may well be that HSs want to show that they can speak the ML “properly”, especially in a test-like situation. It could also be an effect of their typical environment – that HSs are more exposed to input containing explicit forms from L2 speakers than MSs are, as suggested by Azar et al. (2020), or that they are more explicit to make their speech clearer to their L2 interlocutors, as suggested by Polinsky (2018). Finally, it may be that

HSs choose to be explicit in a more nuanced way, such as signaling different contextual factors to the fullest degree, rather than in the straightforward way of simply being more explicit at all costs. This may explain why we see effects of explicitness or greater differentiation in some contexts but not in others where we might have expected them.

Cross-linguistic influence does not seem to be a main driver of dynamic patterns in our data, but may account for two of our findings. First, Turkish HSs produced more unexpected articles in our fill-in-the-blank study than did MSs, which may be related to the fact that Turkish has a very different article system than English. Second, German HSs used the discourse marker *yeah* more frequently than MSs, which we hypothesize results from its functional similarity with the German *ja* "yeah". However, the reduced use of the discourse marker *so* by HSs compared to MSs does not seem to stem from cross-linguistic influence.

The fact that we did not see more cross-linguistic influence in our data overall may be for at least two reasons. First, while we selected structures that differ broadly across the languages in our study (e.g., Russian and Turkish have more free word order than English so we expected more use of non-SVO utterances to introduce new subjects), we did not specifically target a narrow point on which there were clear differences. Second, the type of data we collected – elicited narratives – allows speakers to avoid structures they are uncomfortable with. While elicited narratives provide ecologically valid data in various registers, they are likely not the ideal type of data to see effects of cross-linguistic influence that are necessarily subtle due to the high proficiency of the speakers.

Another perplexing point is that we sometimes found dynamic patterns in one or two of the four groups of HSs but not all of them, and the groups that differed changed from study to study. For example, the German and Russian HSs patterned like the MSs in production of left dislocations while the Greek and Turkish HSs did not. Similarly, the Greek HSs patterned like the MSs in their use of noun phrases vs. pronouns to refer to given referents, but the other three groups of HSs patterned differently. On the surface, it does not seem like cross-linguistic influence is a likely explanation because there are no apparent salient differences between the languages. However, future research should carefully examine the relevant structures in data from monolingually-raised speakers of those languages. Another possible explanation for differences across groups is non-linguistic factors such as subtle SES differences or psychological factors. Nonetheless, we highlight the importance of comparison between different groups of HSs using the same task as the performance across groups allows us to see more clearly if a particular phenomenon is related to bilingualism in general or to some feature of the language(s) in question.

## 9 Conclusion

In sum, we examined the production of majority English by heritage speakers of German, Greek, Russian and Turkish in a variety of studies focusing on different interface phenomena covering several areas of language. The studies revealed many phenomena where heritage speakers and monolingually-raised speakers showed identical patterns, underlining the growing perception in the field that heritage speakers can reasonably be considered native speakers of their majority language (as well as their heritage language). We also found several differences between the productions of (certain groups of) heritage speakers and monolingually-raised speakers, most of which could be attributed to either cross-linguistic influence or some form of explicitness or greater differentiation between contexts than is present in monolingually-raised speakers. Overall, the studies reported here reveal the richness and complexity of patterns in the majority language of heritage speakers.

## Acknowledgements

We are very grateful to all the research assistants who contributed to data collection, annotation, and analysis of the majority English data reported in this chapter, as well as the preparation of this chapter: Iliuza Akhmeztianova, Amy Amoakuh, Aban Alehojat, Chris Allison, Ricarda Bothe, Mert Can, Ryan Carroll, Franziska Cavar, Leah Doroski, Mary Elliott, Harshaa Gopalakrishnan, Gajaneh Hartz, Luc Henriquez, Abigail Hodge, Yuliia Ivashchyk, Janie Fink, Hanna Kim, Hannah Lee, Mark Murphy, Mariia Naumovets, Gizem Öskürçi, Sadaf Rezai, Simge Sargın Kısacık, Jasmine Segarra, Golshan Shakeebae, Selena Song, Shreya Srivastava, Rhiannon Stewart, Mateo Vargas-Nunez, Fiona Wong and Barbara Zeyer. We further thank Lea Coy for help with pre-publication formatting.

We also thank the members on the RUEG project and Mercator Fellows for their many contributions to the RUEG method and for very helpful discussions of our studies over the years of the project. We would also like to thank two anonymous reviewers and three reviewers from the RUEG project (Annika Labrenz, Onur Özsoy, and Wintai Tsehaye) for their constructive feedback and insightful comments.

The studies in this chapter were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation); the grant was awarded to the Research Unit *Emerging Grammars in Language Contact Situations* (FOR 2537), grant numbers 394837597 (P2), 394995401 (P5), 394844953 (P7), 313607803 (P8 and P9).

## References

- Alexiadou, Artemis & Vasiliki Rizou. 2023. The use of periphrasis for the expression of aspect by Greek heritage speakers: A case study of register variation narrowing. *Register Studies* 5(1). 82–110. DOI: 10.1075/rs.20022.ale.
- Alexiadou, Artemis, Vasiliki Rizou & Foteini Karkaletsou. 2022. A plural indefinite article in heritage Greek: The role of register. *Languages* 7(2). 115. DOI: 10.3390/languages7020115.
- 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.
- Barbosa, Poliana, Elena Nicoladis & Margaux Keith. 2017. Bilingual children's lexical strategies in a narrative task. *Journal of Child Language* 44(4). 829–849. DOI: 10.1017/S030500091600026X.
- Baumann, Stefan & Arndt Riester. 2013. Referential and lexical givenness: Semantic, prosodic and cognitive aspects. In Gorka Elordieta & Pilar Prieto (eds.), *Prosody and meaning*, 119–162. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110261790.119.
- Beckman, Mary E., Julia Hirschberg & Stefanie Shattuck-Hufnagel. 2006. The original ToBi system and the evolution of the ToBi framework. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199249633.003.0002.
- Berman, Ruth A. & Irit Katzenberger. 2004. Form and function in introducing narrative and expository texts: A developmental perspective. *Discourse Processes* 38(1). 57–94. DOI: 10.1207/s15326950dp3801\_3.
- Bialystok, Ellen, Gigi Luk, Kathleen F. Peets & Sujin Yang. 2010. Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition* 13(4). 525–531. DOI: 10.1017/S1366728909990423.
- Biber, Douglas, Stig Johansson, Geoffrey N. Leech, Susan Conrad & Edward Finegan. 2021. *Grammar of spoken and written English*. Amsterdam & Philadelphia: John Benjamins. DOI: 10.1075/z.232.
- Boersma, Paul. 2001. Praat: A system for doing phonetics by computer. *Glott International* 5(9/10). 341–345.
- Böttcher, Marlene. 2021. Semantic vs. prosodic prominence: Pronouns realisation in spontaneous mono- and bilingual English. In *Proc. 1st International Conference on Tone and Intonation (TAI)*, 127–131. DOI: 10.21437/TAI.2021-26.
- Böttcher, Marlene. 2022. A comparison of pitch accent patterns in contrastive adjective+noun structures in bilingual Englishes. *Proceedings of Speech Prosody, Lisbon*. 807–811. DOI: 10.21437/SpeechProsody.2022-164.



- Böttcher, Marlene & Fabian Schubö. 2021. Pronoun vs. preposition: Where is stress assigned in German prepositional phrases? In *Proc. 1st International Conference on Tone and Intonation (TAI)*, 88–91. DOI: 10.21437/TAI.2021-18.
- Böttcher, Marlene & Sabine Zerbian. 2020. Stressed pronouns in spontaneous English. In *Proc. Speech Prosody 2020*, 131–135. DOI: 10.21437/SpeechProsody.2020-27.
- Bunk, Oliver, Shanley E. M. Allen, Sabine Zerbian, Tatiana Pashkova, Yulia Zuban & Erica Conti. 2025. Information packaging and word order dynamics in language contact. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 379–412. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775181.
- Bylund, Emanuel, Niclas Abrahamsson & Kenneth Hyltenstam. 2012. Does first language maintenance hamper nativelikeness in a second language? *Studies in Second Language Acquisition* 34(2). 215–241. DOI: 10.1017/S0272263112000034.
- Bylund, Emanuel, Kenneth Hyltenstam & Niclas Abrahamsson. 2021. Age of acquisition – not bilingualism – is the primary determinant of less than native-like L2 ultimate attainment. *Bilingualism: Language and Cognition* 24(1). 18–30. DOI: 10.1017/S1366728920000188.
- Chang, Charles B. 2016. Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition* 19(4). 791–809. DOI: 10.1017/s1366728914000261.
- Chen, Baoguo, Yuefang Jia, Zhu Wang, Susan Dunlap & Jeong-Ah Shin. 2013. Is word-order similarity necessary for cross-linguistic structural priming? *Second Language Research* 29(4). 375–389. DOI: 10.1177/0267658313491962.
- Contemori, Carla & Iva Ivanova. 2021. Bilingual referential choice in cognitively demanding situations. *Bilingualism: Language and Cognition* 24(1). 83–95. DOI: 10.1017/S1366728920000176.
- Contemori, Carla, Naoko Tsuboi & Alma L. Armendariz Galaviz. 2023. Referential choice in two languages: The role of language dominance. *International Journal of Bilingualism*. DOI: 10.1177/13670069231170545.
- de Jong, Nivja H., Jos Pacilly & Willemijn Heeren. 2021. PRAAT scripts to measure speed fluency and breakdown fluency in speech automatically. *Assessment in Education: Principles, Policy & Practice* 28(4). 456–476. DOI: 10.1080/0969594X.2021.1951162.
- de Ruiter, Laura E. 2015. Information status marking in spontaneous vs. read speech in story-telling tasks: Evidence from intonation analysis using GTToBI. *Journal of Phonetics* 48: *The impact of stylistic diversity on phonetic and phono-*

- logical evidence and modeling. Petra Wagner, Jürgen Trouvain & Frank Zimmerer (eds.). 29–44. DOI: 10.1016/j.wocn.2014.10.008.
- Dehé, Nicole. 2018. The intonation of polar questions in North American (“heritage”) Icelandic. *Journal of Germanic Linguistics* 30(3). 213–259. DOI: 10.1017/S1470542717000125.
- Felser, Claudia & Seçkin Arslan. 2019. Inappropriate choice of definites in Turkish heritage speakers of German. *Heritage Language Journal* 16(1). 22–43. DOI: 10.46538/hlj.16.1.2.
- Flege, James Emil. 1999. Age of learning and second language speech. In David Birdsong (ed.), *Second language acquisition and the critical period hypothesis*, 101–131. Mahwah: Lawrence Erlbaum Associates Publishers.
- Fox Tree, Jean E. 2010. Discourse markers across speakers and settings. *Language and Linguistics Compass* 4(5). 269–281. DOI: 10.1111/j.1749-818X.2010.00195.x.
- Geluykens, Ronald. 1992. *From discourse process to grammatical construction: On left-dislocation in English* (Studies in Discourse and Grammar 1). Amsterdam & Philadelphia: John Benjamins. DOI: 10.1075/sidag.1.
- Gorba, Celia. 2019. Bidirectional influence on L1 Spanish and L2 English stop perception: The role of L2 experience. *The Journal of the Acoustical Society of America* 145(6). 587–592. DOI: 10.1121/1.5113808.
- Götze, Michael, Thomas Weskott, Cornelia Endriss, Ines Fiedler, Stefan Hinterwimmer, Svetlana Petrova, Anne Schwarz, Stavros Skopeteas & Ruben Stoel. 2007. Information structure. *Interdisciplinary Studies on Information Structure: ISIS* (7). 147–187.
- Gullberg, Marianne. 2006. Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning* 56(1). 155–196. DOI: 10.1111/j.0023-8333.2006.00344.x.
- Hendriks, Henriette. 2003. Using nouns for reference maintenance: A seeming contradiction in L2 discourse. In Anna Giacalone Ramat (ed.), *Typology and second language acquisition*, 291–326. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110891249.291.
- Hervé, Coralie, Ludovica Serratrice & Martin Corley. 2016. Dislocations in French–English bilingual children: An elicitation study. *Bilingualism: Language and Cognition* 19(5). 987–1000. DOI: 10.1017/S1366728915000401.
- Hickmann, Maya. & Henriette Hendriks. 1999. Cohesion and anaphora in children’s narratives: A comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language* 26(2). 419–452. DOI: 10.1017/S0305000999003785.

- Himmelman, Nikolaus P., Meytal Sandler, Jan Strunk & Volker Unterladstetter. 2018. On the universality of intonational phrases: A cross-linguistic interrater study. *Phonology* 35(2). 207–245. DOI: 10.1017/s0952675718000039.
- Hohenstein, Jill, Ann Einsenbergh & Letitia Naigles. 2006. Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish–English bilingual adults. *Bilingualism: Language and Cognition* 9(3). 249–261. DOI: 10.1017/S1366728906002616.
- Huddleston, Rodney. 2002. Syntactic overview. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 43–70. Cambridge: Cambridge University Press. DOI: 10.1017/9781316423530.003.
- Ionin, Tania. 2006. This is definitely specific: Specificity and definiteness in article systems. *Natural Language Semantics* 14(2). 175–234. DOI: 10.1007/s11050-005-5255-9.
- Ionin, Tania, Maria Luisa Zubizarreta & Salvador Bautista Maldonado. 2008. Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua* 118(4). 554–576. DOI: 10.1016/j.lingua.2006.11.012.
- Kaland, Constantijn, Emiel Krahmer & Marc Swerts. 2014. White bear effects in language production: Evidence from the prosodic realization of adjectives. *Language and Speech* 57(4). 470–486. DOI: 10.1177/0023830913513710.
- Katsika, Kalliopi, Annika Labrenz, Kateryna Iefremenko & Shanley E.M. Allen. 2025. Discourse openings and closings across languages in contact. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 527–576. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775189.
- Keenan, Elinor O. 1977. Why look at planned and unplanned discourse? In Elinor Ochs & Tina Bennet (eds.), *Discourse across time and space*, 1–42. University of Southern California, Los Angeles.
- Keskin, Cem. 2025. Balkan Turkic as a model for understanding contact-induced change in Turkish. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 287–323. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775175.
- Klotz, Martin, Rahel Gajaneh Hartz, Annika Labrenz, Anke Lüdeling & Anna Shadrova. 2024. Die RUEG-Korpora: Ein Blick auf Design, Aufbau, Infrastruktur und Nachnutzung multilingualer Forschungsdaten. *Zeitschrift für germanistische Linguistik* 52(3). 578–592. DOI: 10.1515/zgl-2024-2026.
- 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.

- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4). 243–276.
- Kügler, Frank & Stefan Baumann. 2019. *Annotationsrichtlinien DIMA*. 4th. <http://dima.uni-koeln.de/>.
- Kupisch, Tanja. 2013. A new term for a better distinction? A view from the higher end of the proficiency scale. *Theoretical Linguistics* 39(3-4). 203–214. DOI: 10.1515/tl-2013-0012.
- Kupisch, Tanja. 2014. Adjective placement in simultaneous bilinguals (German–Italian) and the concept of cross-linguistic overcorrection. *Bilingualism: Language and Cognition* 17(1). 222–233. DOI: 10.1017/S1366728913000382.
- Kupisch, Tanja, Dagmar Barton, Katja Hailer, Ewgenia Klaschik, Ilse Stangen, Tatjana Lein & Joost van de Weijer. 2014. Foreign accent in adult simultaneous bilinguals. *Heritage Language Journal* 11(2). 123–150. DOI: 10.46538/hlj.11.2.2.
- Kupisch, Tanja, Alyona Belikova, Öner Özçelik, Ilse Stangen & Lydia White. 2017. Restrictions on definiteness in the grammars of German-Turkish heritage speakers. *Linguistic Approaches to Bilingualism* 7(1). 1–32. DOI: 10.1075/lab.13031.kup.
- Kupisch, Tanja & Jason Rothman. 2018. Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism* 22(5). 564–582. DOI: 10.1177/1367006916654355. (27 April, 2023).
- Kyle, Kristopher. 2020. *Lexical-diversity*. Version 0.1.1. <https://pypi.org/project/lexical-diversity/>.
- Labov, William. 1972. The transformation of experience in narrative syntax. In William Labov (ed.), *Language in the inner city: Studies in the black English vernacular* (Conduct and communication 3), 355–396. Philadelphia: Univ. of Pennsylvania Press.
- Labrenz, Annika. 2023. Functional variation of German *also* across registers and speaker groups. *Contrastive Pragmatics* 4(2). 289–320. DOI: 10.1163/26660393-bja10077.
- Labrenz, Annika, Kateryna Iefremenko, Kalliopi Katsika, Shanley E.M. Allen, Christoph Schroeder & Heike Wiese. 2025. Dynamics of discourse markers in language contact. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 493–526. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775187.
- Ladd, D. Robert. 2008. *Intonational phonology*. 2nd edn. (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

- Lee, Miseon, Hye-Young Kwak, Sunyoung Lee & William O'Grady. 2011. Processing, pragmatics, and scope in Korean and English. In *Proceedings of 19th Japanese-Korean linguistic conference*. Stanford: Center for the Study of Language & Information.
- Lee-Ellis, Sunyoung. 2012. *Looking into bilingualism through the heritage speaker's mind*. University of Maryland (College Park). (Doctoral dissertation).
- Lein, Tatjana, Tanja Kupisch & Joost van de Weijer. 2016. Voice Onset Time and global foreign accent in German–French simultaneous bilinguals during adulthood. *International Journal of Bilingualism* 20(6). 732–749. DOI: 10.1177/1367006915589424.
- Lloyd-Smith, Anika, Marieke Einfeldt & Tanja Kupisch. 2020. Italian-German bilinguals: The effects of heritage language use on accent in early-acquired languages. *International Journal of Bilingualism* 24(2). 289–304. DOI: 10.1177/1367006919826867.
- 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, Anika Labrenz, Maria Martynova, Onur Özsoy, Tatiana Pashkova, Maria Pohle, Judith Purkarthofer, Vicky Rizou, Christoph Schroeder, Anna Shadrova, Luka Szucsich, Rosemarie Tracy, Wintai Tsehay, Heike Wiese, Sabine Zerbian, Yulia Zuban & Nadine Zürn. 2024. *RUEG Corpus*. Version 1.0. Zenodo. DOI: 10.5281/zenodo.11234583.
- Luke, Kang-kwong & Theodossia-Soula Pavlidou (eds.). 2002. *Telephone calls: Unity and diversity in conversational structure across languages and cultures* (Pragmatics & Beyond New Series 101). Amsterdam: John Benjamins. DOI: 10.1075/pbns.101.
- MacWhinney, Brian. 2005. A unified model of language acquisition. In *Handbook of bilingualism*, 50–70. Oxford: Oxford University Press. DOI: 10.1093/oso/9780195151770.003.0004.
- Marinis, Theodoros & Vasiliki Chondrogianni. 2010. Production of tense marking in successive bilingual children: When do they converge with their monolingual peers? *International Journal of Speech-Language Pathology* 12(1). 19–28. DOI: 10.3109/17549500903434125.
- McCardle, Peggy. 2015. Bilingualism: Research and policy. *New Directions for Child and Adolescent Development* (147). 41–48. DOI: 10.1002/cad.20088.
- Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism* (Studies in Bilingualism). Amsterdam: John Benjamins. DOI: 10.1075/sibil.39.

- Montrul, Silvina. 2010. How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics* 31(1). 167–207. DOI: 10.1017/S014271640999021X.
- Montrul, Silvina. 2016. The language of heritage speakers. In *The acquisition of heritage languages*, 41–89. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139030502.003.
- Montrul, Silvina & Tania Ionin. 2010. Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition* 13(4). 449–473. DOI: 10.1017/S1366728910000040.
- Montrul, Silvina & Maria Polinsky (eds.). 2021. *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press. DOI: 10.1017/9781108766340.
- Özsoy, Onur, Kateryna Iefremenko & Christoph Schroeder. 2022. Shifting and expanding clause combining strategies in heritage Turkish varieties. *Languages* 7(3). 242. DOI: 10.3390/languages7030242.
- Özsoy, Onur, Vasiliki Rizou, Maria Martynova, Natalia Gagarina, Luka Szucsich & Artemis Alexiadou. 2025. Null subjects in heritage Greek, Russian and Turkish. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 179–217. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775167.
- Paradis, Johanne. 2019. English second language acquisition from early childhood to adulthood: The role of age, first language, cognitive, and input factors. In Megan M. Brown & Brady Dailey (eds.), *Proceedings of the 43rd Boston University Conference on Language Development*, 11–26. Somerville, MA: Cascadilla Press.
- Paradis, Johanne & Ruiting Jia. 2017. Bilingual children’s long-term outcomes in English as a second language: Language environment factors shape individual differences in catching up with monolinguals. *Developmental science* 20(1). DOI: 10.1111/DESC.12433.
- Pascual y Cabo, Diego & Jason Rothman. 2012. The (il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics* 33(4). 450–455. DOI: 10.1093/applin/ams037.
- Pashkova, Tatiana & Shanley E. M. Allen. 2022. Specificity and definiteness in article choice in heritage speakers’ majority English. *Talk presented at Heritage Languages and Variation, September 23–25 in Limassol, Cyprus*.
- Pashkova, Tatiana & Shanley E. M. Allen. 2025. Explicitness of referring expressions in heritage speakers’ majority English. *Lingua* 314. DOI: 10.1016/j.lingua.2024.103854.



- Pashkova, Tatiana & Shanley E. M. Allen. In preparation. *Syntactic structures in the introduction of new subjects in h-speakers' majority English in the US*.
- Pashkova, Tatiana, Hannah Lee, Mark Murphy & Shanley E. M. Allen. Submitted. *Left dislocations in heritage speakers' majority English in the US*.
- 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.
- Pashkova, Tatiana, Wintai Tsehay, Shanley E. M. Allen & Rosemarie Tracy. 2022. Syntactic optionality in heritage language use: Clause type preferences of German heritage speakers in a majority English context. *Heritage Language Journal* 19(1). 1–41. DOI: 10.1163/15507076-12340022.
- Pavlenko, Aneta & Scott Jarvis. 2002. Bidirectional transfer. *Applied Linguistics* 23(2). 190–214. DOI: 10.1093/applin/23.2.190.
- Pavlidou, Theodossia-Soula. 2014. Phases in discourse. In Klaus P. Schneider & Anne Barron (eds.), *Pragmatics of discourse*, 353–384. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110214406-014.
- Perniss, Pamela & Aslı Özyürek. 2015. Visible cohesion: A comparison of reference tracking in sign, speech, and co-speech gesture. *Topics in Cognitive Science* 7(1). 36–60. DOI: 10.1111/tops.12122.
- Polinsky, Maria. 2018. *Heritage languages and their speakers* (Cambridge Studies in Linguistics 159). Cambridge: Cambridge University Press. DOI: 10.1017/9781107252349.
- Queen, Robin. 2012. Turkish-German bilinguals and their intonation: Triangulating evidence about contact-induced language change. *Language* 88(4). 791–816.
- Rothman, Jason & Jeanine Treffers-Daller. 2014. A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics* 35(1). 93–98. DOI: 10.1093/applin/amt049.
- Sanchez Abchi, Veronica Soledad & Mariela Vanesa de Mier. 2017. Syntactic complexity in narratives written by Spanish heritage speakers. *Vial: Vigo International Journal of Applied Linguistics* 14. 125–148.
- Schegloff, Emanuel A. 1968. Sequencing in conversational openings. *American Anthropologist* 70(6). 1075–1095. DOI: 10.1525/aa.1968.70.6.02a00030.
- Schegloff, Emanuel A. 2007. *Sequence organization in interaction: A primer in conversation analysis*. 1st edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511791208.
- Schegloff, Emanuel A. & Harvey Sacks. 1973. Opening up closings. *Semiotica* 8(4). DOI: 10.1515/semi.1973.8.4.289.

- Schiffrin, Deborah. 1987. *Discourse markers* (Studies in Interactional Sociolinguistics 5). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511611841.
- Schlepppegrell, Mary J. & M. Cecelia Colombi. 1997. Text organization by bilingual writers: Clause structure as a reflection of discourse structure. *Written Communication* 14(4). 481–503. DOI: 10.1177/0741088397014004003.
- 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.
- Schoonbaert, Sofie, Robert J. Hartsuiker & Martin J. Pickering. 2007. The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language* 56(2). 153–171. DOI: 10.1016/j.jml.2006.10.002.
- Scontras, Gregory, Maria Polinsky, Cheng-Yu Edwin Tsai & Kenneth Mai. 2017. Cross-linguistic scope ambiguity: When two systems meet. *Glossa: A Journal of General Linguistics* 2(1). 36. DOI: 10.5334/gjgl.198.
- Selkirk, Elisabeth. 1996. The prosodic structure of function words. In James L. Morgan & Katherine Demuth (eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 187–215. Mahwah: Lawrence Erlbaum Associates, Inc.
- Serratrice, Ludovica. 2007. Referential cohesion in the narratives of bilingual English-Italian children and monolingual peers. *Journal of Pragmatics* 39(6). 1058–1087. DOI: 10.1016/J.PRAGMA.2006.10.001.
- Serratrice, Ludovica, Antonella Sorace & Sandra Paoli. 2004. Crosslinguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition* 7(3). 183–205. DOI: 10.1017/S1366728904001610.
- Shadrova, Anna, Martin Klotz, Rahel Gajaneh Hartz & Anke Lüdeling. 2025. Mapping the mappings and then containing them all: Quality assurance, interface modeling, and epistemology in complex corpus projects. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 69–110. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775161.
- Silva-Corvalán, Carmen. 2014. *Bilingual language acquisition: Spanish and English in the first six years* (Cambridge Approaches to Language Contact). Cambridge University Press.
- Şimşek, Yazgül. 2012. *Sequenzielle und prosodische Aspekte der Sprecher-Hörer-Interaktion im Türkendeutschen*. Münster: Waxmann. 328.



- 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.
- Sorace, Antonella & Francesca Filiaci. 2006. Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22(3). 339–368. DOI: 10.1191/0267658306sr271oa.
- Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci & Michela Baldo. 2009. Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua* 119(3). 460–477. DOI: 10.1016/j.lingua.2008.09.008.
- Tagliamonte, Sali A. & Bridget L. Jankowski. 2019. Grammatical convergence or microvariation? Subject doubling in English in a French dominant town. *Proceedings of the Linguistic Society of America* 4(1). 1–15. DOI: 10.3765/plsa.v4i1.4514.
- Tolchinsky, Liliana, Victoria Johansson & Anita Zamora. 2002. Text openings and closings in writing and speech: Autonomy and differentiation. *Written Language & Literacy* 5(2). 219–253. DOI: 10.1075/wll.5.2.05tol.
- Tracy, Rosemarie & Ira Gawlitzek. 2023. *Mehrsprachigkeit und Spracherwerb*. Tübingen: Narr.
- Trotzke, Andreas, Ermenegildo Bidese & Manuela Caterina Moroni. 2020. German discourse particles in the second language classroom: Teasing apart learning problems at the syntax-pragmatics interface. *Pedagogical Linguistics* 1(2). 184–210. DOI: 10.1075/pl.20008.tro.
- Tsehay, Wintai. 2023. Light-weights placed right: Post-field constituents in heritage German. *Frontiers in Psychology* 14. DOI: 10.3389/fpsyg.2023.1122129.
- Tsehay, Wintai, Tatiana Pashkova, Rosemarie Tracy & Shanley E. M. Allen. 2021. Deconstructing the native speaker: Further evidence from heritage speakers for why this horse should be dead! *Frontiers in Political Science* 12. 1–14. DOI: 10.3389/fpsyg.2021.717352.
- Tsehay, Wintai, Rosemarie Tracy & Johanna Tausch. 2025. Inter- and intra-individual variation: How it materializes in Heritage German and why it matters. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 141–177. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775165.
- Tsimpli, Ianthi-Maria. 2014. Early, late or very late? Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism* 4(3). 283–313. DOI: 10.1075/lab.4.3.01tsi.
- van Rijswijk, Remy, Antje Muntendam & Ton Dijkstra. 2017. Focus marking in Dutch by heritage speakers of Turkish and Dutch L1 speakers. *Journal of Phonetics* 61. 48–70. DOI: 10.1016/j.wocn.2017.01.003.

- Ventureyra, Valérie A.G, Christophe Pallier & Hi-Yon Yoo. 2004. The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics* 17(1). Attrition, 79–91. DOI: 10.1016/S0911-6044(03)00053-8.
- Vogels, Jorrig, Emiel Krahmer & Alfons Maes. 2019. Accessibility and reference production. In Jeanette Gundel & Barbara Abbott (eds.), *The Oxford handbook of reference*, 336–364. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199687305.013.16.
- 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 Tsehay, 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.
- Wiese, Heike, Artemis Alexiadou, Shanley E. M. Allen, Oliver Bunk, Natalia Gagarina, Katerina Iefremenko, Esther Jahns, Martin Klotz, Thomas Krause, Annika Labrenz, Anke Lüdeling, Maria Martynova, Katrin Neuhaus, Tatiana Pashkova, Vicky Rizou, Tracy Rosemarie, Christoph Schroeder, Luka Szucsich, Wintai Tsehay, Sabine Zerbian & Yulia Zuban. 2021. *RUEG Corpus*. Version 0.4.0. Zenodo. DOI: 10.5281/zenodo.3236068.
- Wiese, Heike, Annika Labrenz & Albrun Roy. 2025. Tapping into speakers’ repertoires: Elicitation of register-differentiated productions across groups. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 33–67. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775159.
- Winkle, Claudia. 2015. *Non-canonical structures, they use them differently: Information packaging in spoken varieties of English*. Freiburg: Albert-Ludwigs-Universität Freiburg. (Dissertation). <https://freidok.uni-freiburg.de/data/10600>.
- Yoshioka, Keiko. 2008. Gesture and information structure in first and second language. *Gesture* 8(2). 236–255. DOI: 10.1075/gest.8.2.07yos.
- Zenker, Fred & Kristopher Kyle. 2021. Investigating minimum text lengths for lexical diversity indices. *Assessing Writing* 47. DOI: 10.1016/j.asw.2020.100505.
- Zerbian, Sabine & Marlene Böttcher. 2019. Stressed pronouns in mono-and bilingual German. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren

- (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, 2640–2644. Australasian Speech Science & Technology Association Inc.
- Zerbian, Sabine, Marlene Böttcher & Yulia Zuban. 2022. Prosody of contrastive adjectives in mono-and bilingual speakers of English and Russian: A corpus study. In *Proceedings 11th international conference on speech prosody. International speech communication association (ISCA) online archive*, 812–816. DOI: 10.21437/SpeechProsody.2022-165.
- Zerbian, Sabine, Yulia Zuban, Marlene Böttcher & Oliver Bunk. 2025. Intonation in heritage languages. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 413–448. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775183.

