Straight from the horse's mouth: agreement attraction effects with Turkish possessors

Sol Lago[1], Martina Gračanin-Yuksek[2], Duygu Fatma Şafak[3], Orhan Demir[2], Bilal Kırkıcı[2] & Claudia Felser[1]

[1]*Potsdam Research Institute for Multilingualism, Potsdam, Germany*

[2]*Department of Foreign Language Education, Middle East Technical University, Ankara, Turkey*

[3]*Technische Universität Braunschweig, Braunschweig, Germany*

Corresponding author: Sol Lago
marlago@uni-potsdam.de
University of Potsdam
Potsdam Research Institute for Multilingualism
Campus Golm, Haus 2
Karl-Liebknecht-Strasse 24-25
14476 Potsdam
Germany

Abstract

We investigated the comprehension of subject-verb agreement in Turkish-German bilinguals using two tasks. The first task elicited speeded judgments to verb number violations in sentences that contained plural genitive modifiers. We addressed whether these modifiers elicited attraction errors, which have supported the use of a memory retrieval mechanism in monolingual comprehension studies. The second task examined the comprehension of a language-specific constraint of Turkish against plural-marked verbs with overt plural subjects. Bilinguals showed a reduced application of this constraint, as compared to Turkish monolinguals. Critically, both groups showed similar rates of attraction, but the bilingual group accepted ungrammatical sentences more often. We propose that the similarity in attraction rates supports the use of the same retrieval mechanism, but that bilinguals have more problems than monolinguals in the mapping of morphological to abstract agreement features during speeded comprehension, which results in increased acceptability of ungrammatical sentences.

Keywords: agreement attraction; bilingualism; Turkish.

1. Introduction

Subject-verb agreement errors can provide insight into questions about how speakers license grammatical relations in comprehension. But whereas agreement errors have been widely studied in monolingual speakers [1] (Nicol, Forster, & Veres, 1997; Pearlmutter, Garnsey, & Bock, 1999; Acuña-Fariña, Meseguer, & Carreiras, 2014; Häussler, 2012; Tucker, Idrissi, & Almeida, 2015; Slioussar & Malko; Ristić, Molinaro, & Mancini, 2016), less is known about the comprehension of agreement in speakers whose native language differs from the societally dominant language, typically described as *heritage speakers*: bilingual speakers who are immigrants or children of immigrants and who were exposed to one language during childhood but then switched to the language of their host country (Scontras, Fuchs, & Polinsky, 2015). Here we use their agreement errors to examine whether number computations in a native language differ between monolingual and bilingual speakers, with the goal of better describing the mechanism that underlies the comprehension of agreement in the latter group.

The study of subject-verb agreement in heritage speakers is particularly relevant because they often make agreement errors in production (Benmamoun, Montrul, & Polinsky, 2013). For example, heritage speakers of morphologically rich languages such as Russian, Eastern Armenian, Lithuanian, and Polish often omit verbal agreement in their speech; for instance, English-dominant Russian speakers omit agreement inflections about 20% of the time (Polinsky, 1997, 2006). Subject-verb agreement errors are also found in heritage speakers of Egyptian and Palestinian Arabic (Albirini, Benmamoun, & Chakrani, 2013). However, problems with verbal morphology do not always occur: English-dominant children who are heritage Hungarian speakers show agreement error rates lower than 1% (Bolonyai, 2007; see also Fenveyesi, 2000; de Groot, 2005) and heritage Hindi speakers living in the United States produce subject-verb agreement in a native-like manner (Montrul, Bhatt, & Bhatia, 2012).

In contrast with production, less work has investigated whether agreement errors also occur in comprehension. Some evidence to this effect comes from a study on English-dominant speakers of Inuttitut (Sherkina-Lieber, 2011; Sherkina-Lieber, Perez-Leroux, & Johns, 2011). These speakers, who could not speak Inuttitut but rated their comprehension within the 80–90% range, were tested in an untimed grammaticality task with sentences containing subject-verb number mismatches. In contrast with monolingual speakers, who performed at ceiling detecting number mismatches, bilingual speakers were less accurate (98% vs. 75%), suggesting that agreement difficulties also affect comprehension, at least in cases where the dominant language, English, has poor agreement morphology.

Although this study did not speculate on the source of the differences between monolinguals and heritage speakers, some further evidence was provided by a self-paced reading study on English-dominant speakers of Spanish (Foote, 2011). Foote tested the processing of number and gender agreement in speakers who differed in the age at which they had acquired Spanish: heritage speakers (who had acquired Spanish from birth) and late bilinguals (who had acquired Spanish in the classroom at age 10 or later). Both groups were compared with Spanish monolinguals. In the number agreement conditions, participants read sentences where the subject head and verb were either adjacent (*I see that your father is/*are from Texas*) or separated by intervening material (*The watch of the man is/*are from Switzerland*). Monolinguals and both

---

[1] We use the term "monolinguals" to refer to speakers who grew up in households where only the societally-dominant language was spoken and who were schooled in the societally-dominant language (Kupisch & Rothman, 2016).

groups of bilinguals exhibited similar sensitivity to subject-verb agreement violations, with longer reading times after ungrammatical verbs. Furthermore, the size of the reading time disruptions was smaller in the non-adjacent conditions, suggesting reduced sensitivity to agreement violations when intervening material separated subjects and verbs. Foote suggested that if these effects were due to working memory limitations, these limitations affected monolingual and bilingual speakers alike, consistent with a similar underlying agreement mechanism.

The present study seeks to extend previous findings by examining a specific type of agreement errors, known as *agreement attraction*, which have been fruitfully used in studies on monolingual speakers to provide insight into the mechanisms responsible for agreement computations (Nicol, Forster, & Veres, 1997; Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009; Dillon, Mishler, Sloggett, & Phillips, 2013; Tanner, Nicol, & Brehm, 2014; Jäger, Engelmann, & Vasishth, 2017). In comprehension, attraction errors occur when speakers show reduced sensitivity to ungrammatical verbs, such as *are* in the sentence *The key to the cabinets <u>are</u> on the table*. The presence of the plural modifier, *cabinets*, misleads comprehenders, making them more likely to judge the sentence as acceptable as compared to ungrammatical sentences with a singular modifier (e.g. *cabinet*).

Whereas attraction errors in comprehension were initially thought to result from the same mechanisms as in production (e.g. feature percolation: Franck, Vigliocco, & Nicol, 2002; marking and morphing: Eberhard, Cutting, & Bock, 2005), recent studies have proposed that attraction errors in comprehension mainly occur due to interference during cue-based memory retrieval (Wagers et al., 2009; Dillon et al., 2013; Tanner et al., 2014; for proposals in production, see Badecker & Kuminiak, 2007; Lorimor, Jackson, & Foote, 2015). Further, several findings suggest that memory retrieval is engaged as a repair or error-driven process. Under this account, comprehenders predict the number of the verb based on the number of the subject phrase. When the form of the verb violates this prediction, comprehenders use memory retrieval to check whether a number-matching noun was somehow missed during first pass. When the plural modifier matches the verb in number, it is sometimes wrongly retrieved, allowing comprehenders to license the verb and fooling them into accepting an ungrammatical sentence.

The view of retrieval as a repair process has been supported by several sources of evidence, such as the observation that attraction errors are seen in ungrammatical but not in grammatical sentences, in which the verb matches the number prediction and thus no repair is needed (Wagers et al., 2009; Dillon et al., 2013; Tanner et al., 2014; Jäger et al., 2017). Other studies have supported the involvement of predictive processes in agreement computations (e.g. Tanner et al., 2014) and studies that have looked at the timecourse of attraction effects have found that reading disruptions associated with ungrammatical verbs (in the absence of plural attractors) occur earlier than attraction effects, suggesting that if attraction indexes memory retrieval, retrieval only occurs after number violations have already been detected (Lago, Shalom, Sigman, Lau, & Phillips, 2015; Tucker, Idrissi, & Almeida, 2015; Almeida & Tucker, 2017; but see Tanner, Grey, & van Hell, 2017).

Due to its support in previous studies, here we adopt the view that attraction errors in comprehension result from the use of memory retrieval. The present study sought to address whether bilingual speakers showed agreement errors in comprehension (given previous reports of these errors in production) and if so, whether these errors were due to the same retrieval mechanism proposed for monolingual speakers. With this goal, we used attraction errors to diagnose the use of memory

retrieval and we asked whether bilinguals and monolinguals showed similar attraction rates, consistent with a shared mechanism for subject-verb agreement computations. Further, we examined speakers whose dominant language had rich agreement morphology, in contrast with previous studies, which focused on English-dominant heritage speakers. We tested Turkish-German bilinguals who had learned Turkish from birth, but resided in Germany and spoke German daily. Bilinguals were compared with speakers living in Turkey who did not speak German and mainly used Turkish in their daily lives. Lastly, as previous work examined whether the processing of subject-verb agreement is influenced by the age of acquisition of the target language (e.g. Foote, 2011), we recruited bilingual speakers with different ages of acquisition of German and we addressed whether they modulated their agreement performance.

2. Agreement attraction in Turkish

Our interest in agreement attraction in Turkish as a target language was also motivated by the properties of its rich case system. Previous production studies have shown that grammatical case modulates attraction in languages with case systems of different complexity (*Russian*: Lorimor et al., 2008; Nicol & Wilson, 1999; *Dutch*: Hartsuiker, Antón-Méndez, & van Zee, 2001; *Slovak*: Badecker & Kuminiak, 2007; *English*: Nicol & Antón-Méndez, 2009; Nicol, Barss, & Barker, 2016). One robust finding is that in languages where subjects always bear nominative case, a morphological non-nominative marker on the attractor reduces attraction, as compared to configurations where the subject head and the attractor share the same case form. Within a cue-based retrieval account, this fact is expected if speakers use case information as a retrieval cue for subject-verb agreement: configurations where the subject head and the attractor have the same case should render them more similar and increase their competition as retrieval targets, thus bolstering the likelihood of the attractor being misretrieved.

For instance, Nicol and colleagues examined possessive constructions in English, such as *the elf's house*, where the possessor noun *elf's* is marked with possessive/genitive case. They compared possessor attractors (e.g. *the elf's/elves' house with the tiny window)* with attractors within prepositional phrases (*the elf's house with the tiny window(s)*). They observed that plural prepositional modifiers reliably elicited attraction, consistent with previous studies. By contrast, attraction with possessors was either absent or significantly reduced (*Experiment 1*: 7% vs. 1%; *Experiment 2*: 24% vs. 7%). The authors concluded that the presence of the possessive marker acted as a morphological barrier and prevented agreement attraction in English.

The findings by Nicol and colleagues are relevant to our study because we also used possessor attractors. We were interested in studying Turkish possessive constructions, which parallel English ones in that a genitive marked possessor modifies a nominative marked subject head. In Turkish, the head agrees with the possessor in number and person:

(1) a. ben-im ev-im
    I-1SG.GEN house-1SG.
    my house

  b. biz-im   ev-imiz
    we-1PL.GEN house-1PL.
    our house

However, in contrast with English, where nouns with possessive case cannot act as clausal subjects, genitive-marked subjects in Turkish are very common. Genitive marking is required of subjects in non-finite nominalized embedded clauses, which, besides finite embedded clauses, whose subjects are nominative marked, constitute the canonical way of embedding in Turkish (Göksel & Kerslake, 2005; Coşkun, 2010; Kornfilt, 1997). Further, genitive subjects in nominalized embedded clauses control person and number agreement, with agreement markers showing up on embedded verbs.

(2)   Ben   [siz-in   Ankara'-ya   git-tiğ-iniz-i] bil-iyor-um.
        I-NOM you-2PL.GEN Ankara-DAT go-N-2PL-ACC know-PRES.PROG-1SG
        I know that you are going to Ankara/went to Ankara.

The presence of genitive subjects in Turkish suggests that whereas genitive case might provide a strong cue against subjecthood in English, in Turkish it might not, and thus attraction rates from genitive possessors might be higher in Turkish than in English. Therefore, we predicted that if Turkish speakers showed attraction in comprehension, they should be more likely to wrongly accept ungrammatical sentences when they contained a plural possessor. With regard to Turkish-German bilinguals, we hypothesized that if they used a cue-based retrieval mechanism to license subject-verb agreement in comprehension, then they should show attraction effects, thus sometimes accepting ungrammatical sentences when they contained a plural attractor. Further, given previous reports of reduced use of subject-verb agreement morphology in heritage production (Polinsky, 1997, 2006; Albirini et al., 2013) and comprehension (Sherkina-Lieber et al., 2011), we examined whether heritage speakers were overall less accurate at detecting subject-verb agreement violations. In this case, they should be more likely to accept ungrammatical sentences (with or without number attractors) as compared to monolingual speakers.

3. The current study

To examine the existence of attraction in Turkish, we adopted a design similar to previous comprehension studies (Pearlmutter et al., 1999; Wagers et al., 2009). We manipulated the grammaticality of the sentence (*grammatical/ungrammatical*: 3a,b vs. 3c,d) and the number of the attractor noun within the subject phrase (*singular/plural*: 3a,c vs. 3b,d). In the grammatical conditions, the subject head and the verb both agreed in number (note that singular verbs in Turkish appear without overt agreement morphology), while in the ungrammatical conditions, the subject head was singular and the verb was plural:

(3)   a. <u>Grammatical, SG attractor</u>

      Şarkıcının vokalisti sahnede sürekli zıpladı.
      The singer's backup vocalist non-stop on stage jumped.Ø.

    b. <u>Grammatical, PL attractor</u>

      Şarkıcıların vokalisti sahnede sürekli zıpladı
      The singers' backup vocalist non-stop on stage jumped.Ø.

    c. <u>Ungrammatical, SG attractor</u>

      *Şarkıcının vokalisti sahnede sürekli <u>zıpladılar</u>.
      The singer's backup vocalist non-stop on stage jumped.PL.

d. Ungrammatical, PL attractor

*<u>**Şarkıcıların**</u> vokalisti sahnede sürekli <u>zıpladılar</u>.
The singers' backup vocalist non-stop on stage jumped.PL.

We predicted that if Turkish comprehenders showed attraction, they should be more likely to wrongly accept ungrammatical sentences when they contained a plural attractor, as in (3d), compared with (3c). By contrast, the attractor noun should have a smaller or non-existent effect in grammatical sentences (3a vs. 3b), as found in previous comprehension studies. To increase the likelihood of agreement errors, our task introduced processing pressure by using a speeded visual presentation and requiring participants to provide their judgments under a response deadline.

Before examining agreement attraction, we need to consider one language-specific property of the Turkish agreement system. In the first and second persons, Turkish behaves like most languages: singular subjects require singular verbs and plural subjects require plural-marked verbs. However, in the third person, Turkish shows a different pattern: in sentences with overt plural subjects, verbs preferentially appear without number morphology, which makes them indistinguishable from singular forms (Sezer, 1978; Bamyacı, Häussler & Kabak, 2014):

(4)   a. Çilingir-**ler**   kapı-lar-ı   aç-ıyor.
           locksmith-PL door-PL-ACC open-PRES.PROG.
           Lit.: The *locksmiths is opening the doors*.

       b. ?Çilingir-**ler**   kapı-lar-ı   aç-ıyor-**lar**.
           locksmith-PL door-PL-ACC open-PRES.PROG-3PL
           The *locksmiths are opening the doors*.

Although the reasons that drive this unmarked number preference (henceforth *UNP*) are complex, it has been linked to the semantic and pragmatic properties of plural subjects (Bamyacı et al., 2014; Bamyacı, 2016; Schroeder, 1999). For example, Bamyacı and colleagues (2014) investigated the interaction between animacy and optional verb number marking in Turkish native speakers. They showed that unmarked verbs were considered the default in isolated short sentences, but that this preference varied gradually along the animacy hierarchy, with unmarked verbs being optional for animate nouns but obligatory for inanimate nouns. Further, the UNP gradually decreased with the increase in specificity and 'givenness' of the subject nouns (Bamyacı, 2016; Schroeder, 1999).

As regards Turkish-German bilinguals, Bamyacı (2016) found that they behaved similarly to monolinguals in that they preferred unmarked over plural-marked verbs in sentences with plural subjects. Like monolinguals, bilinguals were more accepting of plural-marked verbs with animate than inanimate subjects, and this acceptance increased with the subject's givenness. However, bilinguals and monolinguals differed in their sensitivity to animacy and givenness distinctions, with bilinguals making finer distinctions along these two dimensions. More relevantly to our study, Bamyacı also observed that heritage speakers displayed "a less pronounced plural dispreference" compared to monolinguals (2016: 110).

Since the UNP only concerns plural subjects, it should not affect agreement attraction, which involves agreement violations with singular subjects. To ensure that this was indeed the case, and that our participants did not reject plural verb forms across

the board, we ran a control experiment that directly compared agreement preferences with singular and plural subjects:

(5)   a. <u>SG subject, unmarked verb</u>

    Dedektif (kusursuz bir şekilde) çalıştı.
    The detective (flawlessly) worked.Ø.

    b. <u>SG subject, plural-marked verb</u>

    *Dedektif (kusursuz bir şekilde) çalıştılar.
    The detective (flawlessly) worked.PL.

    c. <u>PL subject, plural-marked verb</u>

    Dedektifler (kusursuz bir şekilde) çalıştılar.
    The detectives (flawlessly) worked.PL.

    d. <u>PL subject, unmarked verb</u>

    [?]Dedektifler (kusursuz bir şekilde) çalıştı.
    The detectives (flawlessly) worked.Ø.

With plural subjects (5c,d), we expected Turkish speakers to favor unmarked over plural-marked verbs due to the UNP, as was previously shown by Bamyacı and colleagues in isolated short sentences (see also Sezer, 1978). But crucially, with singular subjects (5a,b), the UNP should not affect speakers' agreement profiles, resulting in a preference for unmarked over plural-marked verbs, similarly to other languages.

Based on Bamyacı's results, and together with previous reports that bilinguals sometimes show non-native morphosyntactic performance when it relies on semantic/pragmatic distinctions (Montrul et al., 2012; Sorace, 2004; Arslan, de Kok, & Bastiaanse, 2015), we expected Turkish-German bilinguals to show a reduced UNP in comprehension and to be less accepting of unmarked verbs with plural subjects, as compared to monolinguals. Thus, the data gathered in the UNP task allowed us to assess whether potential group differences in the attraction task were due to problems with subject-verb agreement in the bilingual group, i.e., to differences in the degree to which the two groups disliked plural-marked verbs.

4. Methods

4.1. Participants

Forty-four Turkish speakers (mean age = 20 years, age range = 18–25, 38 females, 43 right-handed) were recruited from the Middle East Technical University in Ankara. None of the participants reported any knowledge of German. In what follows, these speakers are described as "monolingual speakers", although some of them had been exposed to languages other than Turkish, including English (n = 43), French (n = 7), Italian (n = 4), Spanish (n = 2), Korean (n = 2), Russian (n = 1), Albanian (n = 1) and Arabic (n = 1). We did not collect participants' age of acquisition of these languages. However, they were always identified as foreign languages, and participants predominantly spoke Turkish in their daily lives.

Forty-five Turkish-German bilinguals (mean age = 27 years, age range = 18–51, 39 females, 44 right-handed) were recruited through word of mouth in Germany. All but one bilingual had completed their secondary education, with twenty-six having also obtained a university or tertiary degree and five a master's degree (the remaining

heritage speaker was attending the last year of their high-school studies). All bilinguals had acquired Turkish from birth and spoke Turkish and German daily (TABLE 1). The majority of the group (n = 31) had acquired Turkish and German before puberty, and thus were early bilinguals or heritage speakers. The remaining speakers (n = 14) were late bilinguals who had learnt German after puberty due to immigration to Germany. In order to avoid splitting bilinguals into two uneven subgroups, their results were analyzed jointly using their age of acquisition of German as a continuous predictor. Appendix I presents the results of supplementary analyses that examined a subset of early bilinguals who were age and education matched with the monolingual group. These analyses showed qualitatively similar results with the whole-group analyses.

Participants self-rated their language proficiency in both German and Turkish for each of the four skills (speaking, listening, reading and writing) using a scale from 1-10. These ratings were then averaged to get an overall measure of their language proficiency. Self-ratings were used to assess language proficiency because previous research has shown that they offer a good correspondence with formal language tests (Blanche & Merino, 1989; Ross, 1998; Marian, Blumenfeld, & Kaushanskaya, 2007; Lemhöfer & Broersma, 2012), they can be obtained quickly, and they avoid the problem of comparing tests that may vary in the design and level of difficulty of their materials across languages. As shown in TABLE 1, participants reported varying levels of Turkish and German proficiency. All participants provided informed consent and were compensated with course credit or payment.

Table 1. Demographic information of Turkish-German bilingual speakers.

	Mean [range]
Years in Germany	17 [1−49]
Age of acquisition Turkish (*years*)	0 [0−3]
Age of acquisition German (*years*)	9 [0−29]
Frequency Turkish	*'daily'* (93%), *'weekly'* (7%)
Frequency German	*'daily'* (91%), *'weekly'* (9%)
Self-rated Turkish proficiency (*%*)	91 [18−100]
Self-rated German proficiency (*%*)	84 [13−100]

4.2. Materials

Two sets of materials were constructed, as shown in (3) and (5). All materials were normed for plausibility. The first set consisted of forty items and was designed to address our main research question by assessing the existence of agreement attraction. We manipulated the grammaticality of the sentence (*grammatical/ungrammatical*) and the number of the attractor noun within the subject phrase (*singular/plural*). In the grammatical conditions the subject head was singular and the verb unmarked (i.e. lacking a plural marker), while in the ungrammatical conditions the subject head was singular and the verb was marked for plural. Of the experimental verbs, thirty-eight were intransitive (*unergatives*: n = 20; *unaccusatives*: n = 18) and two were transitive but appeared without an object (a grammatical option in Turkish).

The attractor was a genitive-marked possessor, which appeared in the singular form in half of the items and in the plural form in the other half (e.g. "*şarkıcının*" vs. "*şarkıcıların*", 'vocalist' vs. 'vocalists'). The head noun was always singular. Both the

attractor and subject head were third person nouns, always animate and human. The relationship between the possessor and the head noun was controlled for as it has been shown that it can influence the incidence of agreement errors (Nicol et al., 2016, Solomon & Pearlmutter, 2004). All subject noun phrases encoded a semantic relationship between the possessor and the head noun that could be appropriately paraphrased in English by using the possessive marker *'s* or the preposition *of*. The possessor was either professionally related to the head noun, with the head noun denoting an individual performing a service for the possessor (*the soldier's barber*) or an individual professionally superior to the possessor (*the policeman's commander*). Finally, the two nouns sometimes denoted individuals involved in the same kind of activity (*the painter's rival*).

The second set of materials consisted of 32 items that checked whether participants were sensitive to canonical agreement mismatches and to the UNP. We manipulated the number of the subject (*singular/plural*) and the verb (*singular or unmarked/plural*). All subjects were third person nouns, always animate and human. They appeared in sentence-initial position, as in the study by Bamyacı and colleagues (2014). In half of the materials subjects and verbs were adjacent, whereas in the other half, they were separated by a 2/3-word adverbial. As the linear distance between subjects and verbs might influence number agreement (Göksel, 1987; Bamyacı, et al., 2014), we compared the adjacent and non-adjacent conditions statistically to address whether they showed different judgment patterns. Neither Turkish nor Turkish-German speakers showed differences between these conditions, and thus they were collapsed in the final analyses (*monolinguals*: $\hat{\beta} = 0.02$; *SE* = 0.12; $z = 0.14$; $p = .889$; *bilinguals*: $\hat{\beta} = -0.04$; *SE* = 0.12; $z = -0.37$; $p = .713$). Further, the type of intransitive verb used across items (*unergatives*: n = 14; *unaccusatives*: n = 18) also did not modulate participants' judgments (*monolinguals*: $\hat{\beta} = 0.07$; *SE* = 0.14; $z = 0.55$; $p = .579$; *bilinguals*: $\hat{\beta} = 0.05$; *SE* = 0.16; $z = 0.31$; $p = .579$).

In addition, forty fillers were created for the attraction experiment. Half of the fillers were grammatical and the other half were ungrammatical. The grammatical fillers, like the experimental items, consisted of two-word subject phrases but did not contain any possessive constructions. The ungrammatical fillers consisted of violations other than subject-verb number agreement, including person agreement violations, case violations, word order violations and number violations within noun phrases. All experimental materials, as well as data for the experiments, can be found at the Center for Open Science Framework website (https://osf.io/).

4.3. Procedure and analysis

Participants were asked to judge whether the sentences sounded acceptable in Turkish. In the attraction task, sentences were presented one word at a time in the center of the screen with a rapid serial visual presentation rate of 300 ms per word (Wagers et al., 2009). In the UNP task, sentences were presented as a whole with unlimited time to answer. Both tasks were run on a web-based platform using Ibex Farm (Drummond, 2013). Web-based testing was used because it allowed us to expand our participant pool by recruiting Turkish-German bilinguals across Germany, and because this method has been found to yield reliable results in previous psycholinguistics studies (Enochson, & Culbertson, 2014; Sprouse, 2011; Gibson, Piantadosi, & Fedorenko, 2011; Dillon, Clifton, & Frazier, 2014; Chemla, Cummins & Singh, 2015; Wagers & Phillips, 2014).

The experimental session consisted of the attraction and UNP tasks presented in succession. Participants completed the attraction task first and the UNP task second. Before beginning the experimental session, they were given instructions with examples

of ungrammatical sentences and they completed six practice trials with feedback. At the end of each sentence, a response screen appeared for two seconds and participants gave an *"acceptable/unacceptable"* response. They were instructed to make their choice as quickly and accurately as possible. If they waited longer than two seconds before responding, they were given feedback that their response was too slow. An experimental session lasted approximately 40 minutes.

We analyzed the proportion of "acceptable" responses using mixed-effects logistic regression (Jaeger, 2008). Monolingual and bilingual speakers were analyzed separately, but when they showed different profiles, the differences were later quantified using a model with GROUP (*monolingual/bilingual*) as a factor. The attraction task was analyzed using a model with GRAMMATICALITY (*grammatical/ungrammatical*) and ATTRACTOR NUMBER (*singular/plural*) as fixed effects. When the two fixed effects interacted, we conducted follow-up pairwise comparisons. The UNP task was analyzed using a model with SUBJECT NUMBER (*singular/plural*) and (NUMBER) AGREEMENT (*agree/disagree*) as fixed effects. In the bilingual group, the variable AGE OF ACQUISITION OF GERMAN was centered and used as a continuous predictor.

The random effects structure of the models initially included random intercepts and slopes by participants and items. When the model did not converge, we followed the guidelines in Barr, Levy, Scheepers and Tily (2013) and gradually simplified the model until convergence was reached. The structure of the simplified models is provided in the results section. We report effect sizes using model coefficients in log odds ($\hat{\beta}$), standard errors (SE) and the *z*-statistic. Analyses were performed with R, an open-source programming language and environment for statistical computing (R Development Core Team, 2017). P-values were computed using the *lmerTest* package and the Satterthwaite's approximation for denominator degrees of freedom (Kuznetsova, Bruun Brockhoff & Haubo Bojesen Christensen, 2014).

5. Results

5.1. Attraction task

FIGURE 1 shows the proportion of "acceptable" responses across tasks. TABLE 2 shows the results of the statistical analysis. FIGURE 2 shows the effect of the age of acquisition of German of the bilingual group.

Monolingual speakers had 96% accuracy in the filler trials and all participants were above 70%. 2.22% of the experimental trials were removed because they exceeded the response deadline (*Gram, SG attractor*: 1.59%; *Gram, PL attractor*: 2.27%; *Ungram, SG attractor*: 2.50%; *Ungram, PL attractor*: 2.50%).

The final model structure for Turkish monolinguals included intercepts and slopes for both fixed effects by items and intercepts and slopes for ATTRACTOR NUMBER by participants. Acceptability judgments showed a main effect of GRAMMATICALITY: ungrammatical sentences were accepted less often than grammatical sentences. But crucially, there was an attraction effect, as supported by an interaction between GRAMMATICALITY and ATTRACTOR NUMBER: ungrammatical sentences with plural attractors were accepted more often than ungrammatical sentences with singular attractors. No effect of attraction was observed for grammatical sentences.

Bilingual speakers had 91% accuracy in the filler trials and all participants were above 70%. 5.11% of the experimental trials were removed because they exceeded the response deadline (*Gram SG attractor*: 2.22%; *Gram PL attractor*: 4.67%; *Ungram SG attractor*: 5.56%; *Ungram PL attractor*: 8.00%). Bilinguals showed main effects of

GRAMMATICALITY and ATTRACTOR NUMBER: ungrammatical sentences were accepted less often than grammatical sentences and sentences with plural attractors were accepted less often than sentences with singular attractors. Crucially, bilinguals also showed attraction: ungrammatical sentences were accepted more often with plural than singular attractors, and there was no difference in grammatical sentences.

In order to statistically assess differences between monolinguals and bilinguals, their judgments in ungrammatical sentences were directly compared using as predictors GROUP (*monolingual/bilingual*) and ATTRACTOR NUMBER (*singular/plural*) as well as their interaction. Overall, bilinguals accepted ungrammatical sentences significantly more often than monolinguals (30% vs. 20%), as shown by a main effect of GROUP ($\hat{\beta}$ = 0.89; *SE* = 0.35; *z* = 2.53; *p* = .011). However, their attraction rates were quantitatively similar: the acceptability rates of ungrammatical sentences with and without plural attractors did not differ statistically between monolinguals and bilinguals (*monolinguals*: 11%; *bilinguals*: 15%; ns. interaction between GROUP and ATTRACTOR NUMBER).

Lastly, the judgments of the bilinguals in ungrammatical conditions were not modulated by their age of acquisition of German. As can be seen in FIGURE 2, ungrammatical sentences showed a marginal tendency towards being rejected more often as participants' German age of acquisition increased ($\hat{\beta}$ = -0.04; *SE* = 0.02; *z* = -1.70; *p* = .090), but this pattern affected sentences with singular and plural attractors alike and thus it was not specific to attraction effects (ns. interaction between GROUP and ATTRACTOR NUMBER). These results suggest that attraction rates were unaffected by the age at which bilinguals acquired German.

Table 2. Model results for the judgments of monolingual and Turkish-German bilingual speakers. Negative coefficients index a reduction in the proportion of "*acceptable*" responses. Significant effects at the α = .05 level are bolded.

	Monolingual speakers				Bilingual speakers			
	$\hat{\beta}$	SE	*z*	*p*	$\hat{\beta}$	SE	*z*	*p*
Attraction task								
Grammaticality	**-5.51**	**0.33**	**-16.69**	**.000**	**-3.67**	**0.33**	**-11.29**	**.000**
Attractor number	0.14	0.25	0.57	.571	**0.56**	**0.18**	**3.07**	**.002**
Grammaticality × Attractor number	**1.69**	**0.53**	**3.19**	**.001**	**0.88**	**0.34**	**2.57**	**.010**
Attractor number: Ungram conditions	**0.94**	**0.26**	**3.68**	**.000**	**0.92**	**0.18**	**5.17**	**.000**
Attractor number: Gram conditions	-0.79	0.52	-1.51	.131	-0.07	0.25	-0.29	.769
UNP task								
Subject number	**2.72**	**0.90**	**3.02**	**.003**	**3.10**	**0.84**	**3.69**	**.000**
Agreement	**-5.95**	**1.07**	**-5.57**	**.000**	**-8.43**	**1.69**	**-5.00**	**.000**
Subject number × Agreement	**14.55**	**1.24**	**11.76**	**.000**	**14.93**	**2.98**	**5.01**	**.000**
Agreement: SG subject conditions	**-10.24**	**1.09**	**-9.42**	**.000**	**-11.01**	**1.55**	**-7.11**	**.000**
Agreement: PL subject conditions	**3.20**	**0.52**	**6.21**	**.000**	-0.03	0.34	-0.10	.924

Figure 1. *(A) Attraction task*: Monolinguals and bilinguals were more likely to accept ungrammatical sentences when they contained a plural attractor. Both groups showed similar rates of attraction. *(B) UNP task*: With third person plural subjects, monolingual speakers showed a UNP and preferred unmarked to plural-marked verbs. By contrast, Turkish-German bilinguals accepted unmarked and plural-marked verbs to a similar degree.
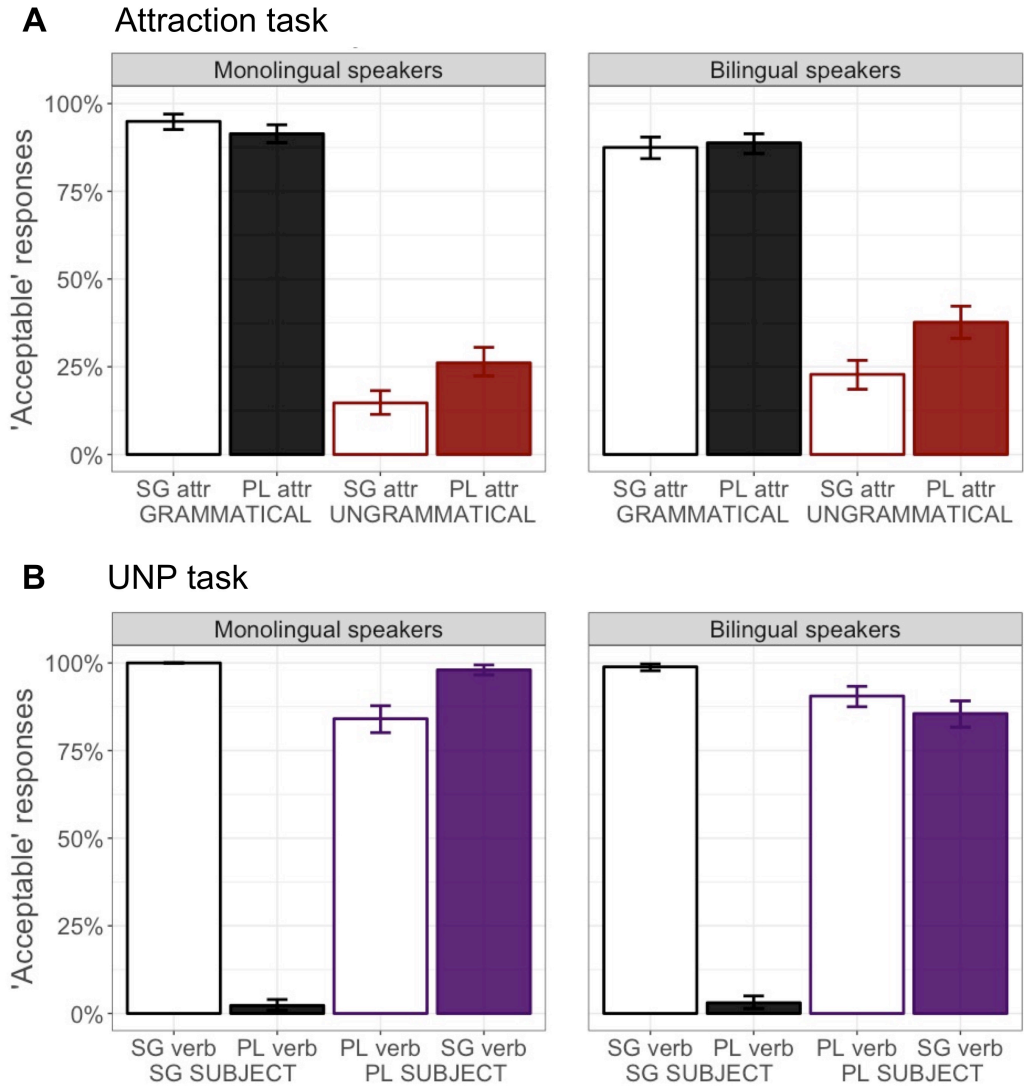
Figure 2. *(A) Attraction task*: the attraction rates of bilinguals were not modulated by their age of acquisition of German. Empty triangles show by-subject averages for the singular attractor conditions and filled triangles for the plural attractor conditions. *(B) UNP task*: with third person plural subjects, bilinguals who had acquired German later showed a stronger UNP, being more likely to prefer unmarked over plural-marked verbs, as compared with participants with an earlier German age of acquisition. Empty triangles show by-subject averages for the plural verb conditions and filled triangles for the singular verb conditions. Note that although the y-axis shows back-transformed percentages for easier readability, distances are expressed in a log scale, to adequately reflect the estimates of the logistic regression model.
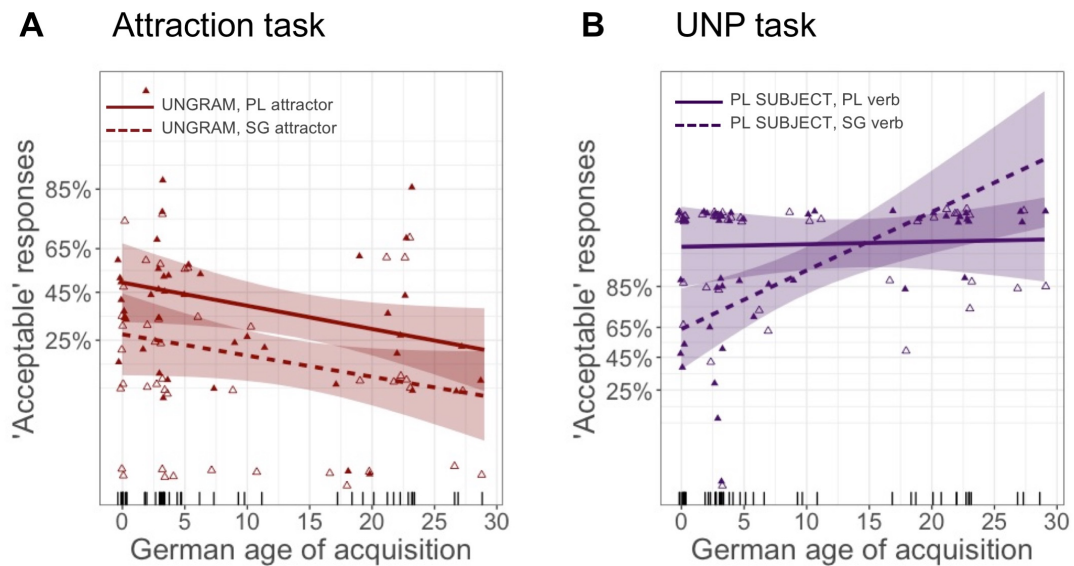
5.2. UNP task

The final model structure for Turkish monolinguals included intercepts and slopes for both fixed effects by participants, and intercepts and slopes for SUBJECT NUMBER by items.[2] The responses of Turkish monolinguals showed a main effect of AGREEMENT: sentences where subjects and verbs disagreed in number were accepted less often than sentences where subjects and verbs agreed in number. Further, there was a main effect of SUBJECT NUMBER: sentences with plural subjects were accepted more often than sentences with singular subjects. But crucially, there was a significant interaction between SUBJECT NUMBER and AGREEMENT. In sentences with singular subjects, unmarked verbs were accepted more often than plural-marked verbs (i.e. a canonical agreement effect). But in sentences with plural subjects, plural-marked verbs were accepted less often than unmarked verbs (i.e. a reversed agreement effect). Thus, with plural subjects, monolingual speakers preferred unmarked to plural-marked verbs, consistent with a UNP.

Bilingual speakers showed a somewhat different pattern. With singular subjects, they showed a canonical agreement effect by preferring singular to plural-marked verbs, similarly to monolinguals. But with plural subjects, plural-marked and unmarked verbs

---

[2]The model failed to converge due to floor effects in the singular subject conditions, where singular verbs were almost always judged as acceptable and plural verbs were almost always judged as unacceptable (98% and 2% respectively). To address this problem of complete separation, we adopted the penalized or bias-corrected solution proposed by Firth (1993) and added a weak prior to the fixed-effect parameters in the statistical model (Gelman, Jakulin, Pittau & Su, 2008; Abrahantes & Aerts, 2012).

were accepted to similar extents (91% and 86% respectively). Thus, although bilinguals largely accepted unmarked verbs with plural subjects, they did not prefer them to plural-marked verbs.

In order to statistically assess differences between monolinguals and bilinguals, their judgments in the conditions relevant to the UNP, which contained plural subjects, were compared using as predictors GROUP (*monolingual/bilingual*) and AGREEMENT (*agree/disagree*) as well as their interaction. The difference between the two groups was supported by a marginal GROUP × AGREEMENT interaction ($\hat{\beta}$ = -2.63; *SE* = 1.33; *z* = -1.98; *p* = .048).

Lastly, the judgments of the bilinguals in the UNP conditions showed a significant correlation with their age of acquisition of German ($\hat{\beta}$ = 0.15; *SE* = 0.05; *z* = 3.39; *p* = .000). Participants who had acquired German later showed a stronger UNP, being more likely to prefer unmarked over plural-marked verbs with plural subjects, as compared with participants with an earlier German age of acquisition. Thus, bilinguals who acquired German later in life behaved more similarly to monolinguals than bilinguals who acquired German earlier. Note, however, that participants' age of acquisition of German was significantly correlated with their German (*r* = -0.52; *t* = -3.95; *p* = .000) and Turkish proficiencies (*r* = 0.40; *t* = 2.77; *p* = .008): Bilinguals who acquired German earlier ranked their German proficiency more highly (and their Turkish proficiency more lowly) than participants who acquired German later. Therefore, differences in the UNP between bilinguals might have been due to their German age of acquisition, Turkish proficiency, German proficiency, or a combination of these.

6. Discussion

This study examined the comprehension of subject-verb agreement by comparing attraction rates between Turkish monolinguals and Turkish-German bilinguals. We were interested in whether bilinguals would show attraction effects, which have been used to support the use of a memory retrieval mechanism in monolingual speakers. In addition, we examined a language-specific constraint, the unmarked number preference (UNP). The findings of each task are discussed below.

To elicit agreement attraction, we used genitive modifiers in order to take advantage of the properties of the case system of Turkish. Whereas possessive case marking in languages like English and German precludes nouns from being grammatical subjects, Turkish requires genitive-marked subjects in non-finite embedded clauses. We hypothesized that if case information played a role in the selection of an appropriate controller for agreement, then attraction effects with possessors should occur in Turkish, in contrast with languages like English, where attraction from possessor modifiers has been reported to be non-existent or minimal (Nicol et al., 2016).

The results of monolinguals speakers confirmed this prediction by showing robust attraction effects. To our knowledge, these results provide the first demonstration of agreement attraction in Turkish. Further, the size of attraction effects was indeed stronger than in previous English experiments: Whereas in English attraction rates ranged between 1%–7%, attraction rates in our study almost doubled, ranging between 11%–15% across monolingual and bilingual speakers. Therefore, although this cross-linguistic comparison should be taken with caution due to several between-studies differences (e.g. production vs. comprehension tasks, different participants and materials), these patterns suggest that case information can indeed affect agreement computations: In English, genitive marking on a noun should provide a strong cue against subjecthood. By contrast, genitive case in Turkish is compatible with

subjecthood, and thus genitive-marked attractors might be more likely to mislead Turkish than English comprehenders.

The results of the Turkish-German group showed an interesting pattern of similarities and differences with monolingual speakers. First, bilinguals were more likely to accept ungrammatical sentences overall (i.e., collapsing across the singular and plural attractor conditions), consistent with the reduced use of subject-verb agreement in heritage populations (Polinsky, 1997, 2006; Albirini et al., 2013). However, the examination of errors in ungrammatical sentences showed similar rates of attraction in monolingual and bilingual speakers. These results support the view that both groups use the same mechanism to license subject-verb agreement in comprehension. Specifically, within a framework where attraction arises when participants detect an ungrammatical verb and initiate memory retrieval to license it, the similar attraction rates across groups provide evidence that monolinguals and bilinguals use retrieval in the same way and to the same extent in comprehension.

To reconcile the two patterns found for bilingual speakers (similar attraction rates but more errors overall compared to monolinguals) we suggest that two processes are differentially affected in the computation of agreement: the mapping of morphological verb forms to abstract agreement features, such as number and person, and the subsequent use of these features to license agreement during memory retrieval. As mentioned in the Introduction, a well-established account of attraction in comprehension proposes that memory retrieval is an error-driven process: when comprehenders detect that the number of the verb does not match the number predicted by the subject phrase, they attempt to license it by retrieving other previously encountered constituents, a process that may result in misretrieval and thus give rise to attraction.

But crucially, in order to detect the number mismatch between the subject and the verb, participants need to have successfully mapped the morphological information on the verb to abstract agreement features (e.g. '3rd person singular'), in order to compare them with their number prediction. We propose that this first stage of agreement computation may be vulnerable in bilingual speakers, because the mapping of morphological forms to agreement features might be error-prone in situations that require the rapid decomposition of fast-evolving input, such as everyday speech or the speeded presentation format of our acceptability judgment task. A reduced ability to perform this process would result in verbal representations with unspecified abstract features, which, compared with the features of the subject head, would not yield a mismatch. This would, in turn, result in bilinguals' occasional inability to recognize ungrammatical verb forms, thus making them more likely to accept ungrammatical sentences than monolinguals.

By contrast, the later stage of agreement licensing, involving memory retrieval, is likely to be similarly implemented by monolingual and bilingual speakers, as diagnosed by their similar attraction rates. Bilinguals' behavior may have therefore resulted from a mixture of two different types of trials: First, a minority of trials, in which participants might have failed to recognize ungrammatical verb forms altogether, thus failing to notice ungrammatical sentences (with both singular and plural attractors) and to initiate memory retrieval. In these trials, bilinguals behaved differently from monolinguals. Secondly, a majority of trials, in which bilinguals were able to appropriately parse the verb and extract its agreement features, using them as cues for memory retrieval and resulting in agreement attraction. In these trials, bilinguals behaved like monolinguals. In short, what the similar attraction rates across groups indicate is that despite the fact that bilinguals detected fewer ungrammatical sentences overall, the influence of the

plural attractor noun with respect to the total number of ungrammatical rejections was quantitatively indistinguishable from that of the monolingual group. Heritage speakers' errors with subject-verb agreement in comprehension might thus be due to reduced sensitivity to morphological inflection rather than to different underlying mechanisms for computing agreement.

A current limitation of our findings is that we did not measure the processing of agreement at the verb: our tasks only probed for the downstream consequences of agreement computations after participants had already finished reading the sentences. Further work using incremental measures such as eye-tracking or event-related potentials will be able to address this limitation. If our suggestion is on the right track, heritage speakers should show reduced processing disruptions when encountering ungrammatical verbs (as compared to monolinguals), but facilitation in ungrammatical sentences due to plural modifiers should be quantitatively similar to that of monolingual speakers. The use of online measurements would also represent a useful addition to the heritage literature, which has often focused on *whether* monolingual and bilingual speakers differ, rather than on *how* they differ and under what circumstances (Rothman & Treffers-Daller, 2014; Kupisch & Rothman, 2016).

In addition to agreement attraction, we found support for a previously reported language-specific constraint, the UNP, which describes Turkish speakers' preference for unmarked over plural-marked verbs in sentences with overt plural subjects. To verify that the UNP did not cause Turkish speakers to disprefer plural verbs across the board, thus potentially affecting the results of the attraction task, the UNP task examined participants' judgments to number mismatches with singular and plural subjects. Monolinguals showed evidence of the UNP and accepted plural subject sentences more often with unmarked than plural-marked verbs (i.e. a reversed agreement preference). By contrast, bilingual speakers judged unmarked and plural verbs as similarly acceptable, showing a reduced UNP (Bamyacı, 2016). As the UNP is influenced by semantic and pragmatic factors, these results support previous claims that bilinguals' morphosyntactic performance may be non-native-like when it relies on semantic/pragmatic distinctions (Montrul et al., 2012; Sorace, 2004; Arslan et al., 2015). However, this contrast should be treated as quantitative rather than qualitative, because both monolinguals and bilinguals accepted plural subjects with unmarked verbs to a very large extent. The key difference was that whereas monolinguals preferred unmarked to plural verbs, bilinguals accepted both to similar extents. Finally, with singular subjects, both groups preferred singular to plural verbs (i.e. a canonical agreement preference), thus validating our agreement attraction design.

The reduced UNP in bilingual speakers could be due to several factors. One is cross-linguistic influence from German, a language in which plural subjects require plural-marked verbs. Due to their exposure to German, bilinguals may have overgeneralized this requirement to Turkish. However, this explanation is unlikely, because bilinguals only showed a weakening of the UNP, rather that its reversal, which would have been expected if they had simply transferred their German agreement constraints. A more likely explanation for bilinguals' behavior might draw on the properties of their Turkish input. A previous study (Bamyacı et al., 2014) found that the UNP was more prevalent in younger than older generations of Turkish monolingual speakers (mean age 28 vs. 43 years). Since most of the bilinguals in our study (mean age 27 years) grew up in Germany, their Turkish input is likely to have come primarily from their parents, whose grammar is probably similar to that of the older speakers in Bamyacı's study. Under this account, the UNP was already reduced in the input received by the bilinguals, possibly leading to their failure in developing a robust

preference for unmarked verbs with plural subjects. However, this explanation should also remain speculative because it relies exclusively on the results reported in Bamyaci and colleagues (2014) and because we did not collect information about the bilinguals' caregivers, or about the type of Turkish input that they had received while growing up.

Lastly, the weaker UNP in the bilingual group may have resulted from both the influence of German and the properties of their Turkish input. A joint role for these factors would be consistent with results reported for the comprehension and production of differential object marking (DOM) in heritage speakers of Spanish living in the United States (Montrul, 2014). Montrul compared first generation Spanish immigrants, who had grown up as monolingual Spanish speakers but were currently living in heritage conditions, with simultaneous and sequential bilinguals. Because all three groups showed higher omission rates for DOM than a control group of monolingual Spanish speakers, Montrul argued that these were caused by both cross-linguistic influence from English, which lacks DOM (i.e., in the first-generation group) and reduced-input conditions (i.e., in the simultaneous and sequential bilinguals, who had switched to English after childhood, and whose parents exhibited signs of attrition). Further research is needed to examine whether our findings may have similarly resulted from the combination of both the availability of German and the quantity and quality of Turkish input.

The influence of participants' German age of acquisition on the UNP cannot discriminate between the possibilities above because it was significantly correlated with both their German and Turkish proficiencies. A more interesting aspect of the age of acquisition effect is that it was selective: it affected participants' behavior with the UNP but not with agreement attraction. This dissociation supports the generalization that heritage speakers performed differently from monolinguals when tested on a language-specific, pragmatic or discourse-related constraint, the UNP, but were mostly similar to monolinguals when tested on agreement attraction, which probes for a cross-linguistically robust constraint on subject-verb agreement. This dissociation resembles previous developmental findings in monolingual children, which indicate a difference between universal, cross-linguistically robust constraints (acquired early) and language-specific ones, acquired later (Phillips & Ehrenhofer, 2015). Together with these findings, our study suggests that taking into account the cross-linguistic status of grammatical constraints provides a useful way to predict areas of variation between monolingual and heritage populations.

References

Acuña-Fariña, J.C, Meseguer, E., & Carreiras, M. (2014). Gender and number agreement in comprehension in Spanish. *Lingua*, 143, 108–128.

Albirini, A., Benmamoun, E., & Chakrani, B. (2013). Gender and number agreement in the oral production of Arabic heritage speakers. *Bilingualism: Language and Cognition*, 16, 1–18.

Almeida, D., & Tucker, M. (2017). The Complex Structure of Agreement Errors: Evidence from Distributional Analyses of Agreement Attraction in Arabic. *Proceedings of the 47th Meeting of the North-Eastern Linguistic Society*.

Arslan, S., de Kok, D. & Bastiaanse, R. (2015). Processing grammatical evidentiality and time reference in heritage and monolingual Turkish speakers. *Bilingualism: Language and Cognition*, 20, 3, 457–472.

Au, T., Knightly, L. Jun, S-A. & Oh, J. (2002). Overhearing a language during childhood. *Psychological Science* 13, 238–243.

Badecker, W., and Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: evidence from gender and case in Slovak. *Journal of Memory and Language*, 56, 65–85.

Baker, M.C. (2008). The Syntax of Agreement and Concord, Cambridge University Press, Cambridge.

Bamyacı, E. (2016) Competing Structures in the Bilingual Mind: A Psycholinguistic Investigation of Optional Verb Number Agreement. Springer.

Bamyacı, E., Häussler, J., Kabak, B. (2014) The interaction of animacy and number agreement: An experimental investigation. *Lingua*, 148, 254–277.

Barr D.J., Levy R., Scheepers C., & Tily, H. (2013) Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 3, 255–278.

Benmamoun, E., Montrul, S., & Polinsky, M. (2013). Heritage Languages and Their Speakers: Opportunities and Challenges for Linguistics. *Theoretical Linguistics*, 39, 129–181.

Blanche, P. & Merino, B.J. (1989). Self-Assessment of Foreign-Language Skills: Implications for Teachers and Researchers. Language Learning, 39, 3, 313–338.

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45–93.

Bolonyai, A. (2007). (In)vulnerable agreement in incomplete bilingual L1 learners. *The International Journal of Bilingualism*, 11, 3–2.

Chemla, E., Cummins, C., & Singh, R. (2016). Training and timing local scalar enrichments under global pragmatic pressures. *Journal of Semantics*, 34, 1, 107–126.

Coşkun, H. (2010). Question elements in Turkish complement clauses. *Turkic Languages*, 14: 43–68.

de Groot, C. (2005). The grammars of Hungarian outside Hungary from a linguistic-typological perspective. In A. Fenyvesi (Ed.). *Hungarian language contact outside Hungary*, 351–370. Amsterdam: John Benjamins.

Dillon, B., Clifton, Jr., C., & Frazier, L. (2014) Pushed aside: Parentheticals and Processing. *Language, Cognition and Neuroscience*, 29, 4, 483–498.

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting interference profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69, 2, 85–103.

Drummond, A. (2013) Ibex Farm. Available at http://spellout.net/ibexfarm/.

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number

agreement in sentence production. *Psychological Review*, 112, 531−559.

Enochson K, & Culbertson, J. (2015) Collecting Psycholinguistic Response Time Data Using Amazon Mechanical Turk. *PLOS ONE* 10, 3, doi: 10.1371/journal.pone.0116946.

Fenyvesi, A. (2000). The affectedness of the verbal complex in American Hungarian. In A. Fenyvesi & K. Sándor (Eds.). Language contact and the verbal complex of Dutch and Hungarian. *Working papers from the 1st Bilingual Language Use Theme Meeting of the Study Centre on Language Contact*, November 11–13, 1999, Szeged, Hungary, 94–107. Szeged: JGyTF Press.

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics*, 32, 187–220.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Attraction in sentence production: The role of syntactic structure. *Language and Cognitive Processes*, 17, 371–404.

Gibson, E., Piantadosi, S., & Fedorenko, K. (2011) Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistic Compass*, 5, 509–524.

Göksel, A. (1987). Distance restrictions on syntactic processes. In H. E. Boeschoten & L. T. Verhoeven (Eds.), Studies on Modern Turkish: Proceedings of the Third Conference on Turkish Linguistics (pp. 69–81). Tilburg: Tilburg University Press.

Göksel, A., & Kerslake, C. (2005). *Turkish: A Comprehensive Grammar.* New York: Routledge.

Hartsuiker, R. J., Antón-Méndez, I., & van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, 45, 546–572.

Haznedar, B., & Schwartz, B. (1997) Are there optional infinitives in child L2 acquisition? In E. Hughes, M. Hughes, & A. Greenhill, (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development (BUCLD)*, pp. 293–306. Somerville, MA: Cascadilla Press.

Häussler, J. (2012). The emergence of attraction errors during sentence comprehension. Doctoral Dissertation, Universität Konstanz.

Jacob, G., & Kırkıcı, B. (2016). The processing of morphologically complex words in a specific speaker group: A masked-priming study with Turkish heritage speakers. *The Mental Lexicon*, 11, 2, 308–328.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Jäger, L., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language,* 94, 316–339.

Kornfilt, J. (1997). *Turkish*. London, New York: Routledge.

Kupisch, T. & Rothman, J. (2016) Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 1–9.

Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0–11. http://CRAN.R-project.org/package=lmerTest.

Lago, S., Shalom, D., Sigman, M., Lau, E., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language*, 82, 133–149.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44,* 325–343.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 375–419.

Lorimor, H., Bock, J. K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Number Agreement and Attraction in Russian. Language and Cognitive Processes, 23, 769–799.

Lorimor, H., Jackson, C. N., & Foote, R. (2015). How gender affects number: Cue-based retrieval in agreement production. *Language, Cognition and Neuroscience*, 30, 947–954.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 4, 940–967.

Montrul, S. (2012). Is the heritage language like a second language? EUROSLA Yearbook, 12, 1–29.

Montrul, S. (2014). Structural changes in Spanish in the United States: Differential object marking in Spanish heritage speakers across generations. *Lingua*, 151, 177–196.

Montrul, S., Bhatt, R. & Bhatia, A. (2012). Erosion of case and agreement in Hindi heritage speakers. *Linguistic Approaches to Bilingualism*, 2, 141–176.

Montrul, S., Foote, R., & Perpiñán, S. (2008) Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58, 503–553.

Nicol, J., & Antón-Méndez, I. (2009). "The effect of case marking on subject– verb agreement errors in English," in *Time and Again: Theoretical Perspectives on Formal Linguistics in Honor of D. Terence Langendoen,* eds W. D. Lewis, S. Karimi, H. Harley, and S. O. Farrar (Amsterdam: John Benjamins Publishing), 135–150.

Nicol, J. L., Barss, A., & Barker, J. E. (2016). Minimal Interference from Possessor Phrases in the Production of Subject-Verb Agreement. *Frontiers in Psychology*. http://doi.org/10.3389/fpsyg.2016.00548

Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, *36*, 569–587.

Nicol, J., & Wilson, R. (1999). "Agreement and case-marking in Russian: a psycholinguistic investigation of agreement errors in production," in *The Eight Annual Workshop on Formal Approaches to Slavic Languages. The Philadelphia Meeting* (Ann Arbor, MI: Michigan Slavic Publications), 314–327.

Odlin, T. (2010). Cross-linguistic influence. In C. J. Doughty and M. H. Long (eds), The *Handbook of Second Language Acquisition*. Oxford: Blackwell Publishing, 436–287.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–456.

Phillips, C., & Ehrenhofer, L. (2015). The role of processing in language acquisition. *Linguistic Approaches to Bilingualism*, 5, 4, 409–453.

Pires, A., & Rothman, J. (2009). Disentangling contributing variables to incomplete acquisition competence outcomes: what differences across Brazilian and European Portuguese heritage speakers tell us. *International Journal of Bilingualism*, 13, 211–238.

Polinsky, M. (1997). Cross-linguistic parallels in language loss. *Southwest Journal of Linguistics*, 14/1–2, 87–123.

Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics,* 14, 191–262.

Polinsky, M., & Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass*, 1, 5, 368–395.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 103–33.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rinke, E. & Flores, C. (2014). Heritage Portuguese bilinguals' morphosyntactic knowledge of clitics. *Bilingualism: Language and Cognition*, 17, 4, 681–699.

Ristic, B., Molinaro, N., & Mancini, S., (2016) Agreement attraction in Serbian: Decomposing markedness. *The Mental Lexicon*, 11, 2, 242–276.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, 5, 1–20.

Rothman, J. (2007). Heritage speaker competence differences, language change, and input type: Inflected infinitives in heritage Brazilian Portuguese. *The International Journal of Bilingualism*, 11, 359–389.

Rothman, J. & Treffers-Daller, J. (2014) A prolegomenon to the construct of the native speaker: heritage speaker bilinguals are natives too! Applied Linguistics, 35, 93–98.

Schroeder, C., (1999). *The Turkish Nominal Phrase in Spoken Discourse*. Harrassowitz Verlag, Wiesbaden.

Scontras, G., Fuchs, Z., & Polinsky, M. (2015) Heritage language and linguistic theory. *Frontiers in Psychology*. http://doi.org/10.3389/fpsyg.2015.01545.

Sezer, E., (1978). Eylemlerin çoğul öznelere uyumu. Genel Dilbilim Dergisi, Ankara Dilbilim Çevresi Derneği, Ankara, 25–32.

Sherkina-Lieber, M. (2011). Comprehension of Labrador Inuttitut functional morphology by receptive bilinguals. Doctoral Dissertation, University of Toronto.

Sherkina-Lieber, M., Perez-Leroux, A. T. & Johns, A. (2011). Grammar without speech production: The case of Labrador Inuttitut heritage receptive bilinguals. *Bilingualism: Language and Cognition*, 14, 301–317.

Slioussar, N. & Malko, A. (2016). Gender Agreement Attraction in Russian: Production and Comprehension Evidence. *Frontiers in Psychology*, 7, 1651–1671.

Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, 49, 1– 46.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretation and methods. *Bilingualism: Language and Cognition*, 7, 143–146.

Sprouse, J. (2011) A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavioral Research Methods*, 43, 155–167.

Tanner, D., Grey, S., & van Hell, J. (2017). Dissociating retrieval interference and reanalysis in the P600 during sentence comprehension. *Psychophysiology*, 248–259.

Tanner, D., Nicol, J., & Brehm, L. (2014). The time course of feature interference in

agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language*, 76, 195–215.

Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing Number in the Real-Time Processing of Agreement: Self-Paced Reading Evidence from Arabic. *Frontiers in Psychology*. http://doi.org/ 0.3389/fpsyg.2015.00347.

Wagers, M., Lau, E., & Phillips, C. (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language, 61,* 206–237.

Wagers, M.W., Phillips, C. (2014). Going the distance: memory and control processes in active dependency construction. *Quarterly Journal of Experimental Psychology*, 67, 1274–1304.

Appendix I: Supplementary Analyses

This appendix reports supplementary analyses conducted on a subset of bilingual speakers who were similar in age and level of education with the monolingual group. We note that the criteria for defining "heritage speaker" differ substantially with regard to whether or not they make reference to the age of acquisition of the dominant language. For example, Polinsky and Kagan (2007) define heritage speakers as "people raised in a home where one language is spoken who subsequently switch to another dominant language". This definition includes speakers who moved to the host country in early adulthood. By contrast, other accounts restrict the term to speakers who are exposed to both the heritage and societally dominant languages from early childhood (Montrul, 2012; Scontras, Fuchs, & Polinsky, 2015).

For the supplementary analyses, we adopted the stricter definition of heritage speakers and considered only those who had acquired both Turkish and German from early childhood. This group of twenty heritage speakers (mean age = 23 years, age range = 18–29, 18 females, all right-handed) had acquired Turkish from birth and German from pre-school, although half of participants reported already having some German exposure in the household (mean German age of acquisition = 2 years, range = 0–7 years). Heritage participants resided in Germany at the time of testing, where they had lived an average of 22 years (range = 14–28 years). They rated themselves as more proficient in German (mean = 97%, range = 85–100%) than Turkish (mean = 86%, range = 55–98%) and reported speaking both languages daily. All but one heritage speakers had completed their secondary education, with eight having also obtained a Bachelor's degree and one a Master's degree (the remaining heritage speaker was attending the last year of high-school).

In order to compare heritage speakers with a similarly sized group, the first twenty participants of the monolingual group were selected (mean age = 20 years, age range = 18–21, 15 females, all right-handed). All monolinguals were college students at the Middle East Technical University in Ankara. The same within- and between-group analyses as reported in the paper were performed. The results replicated the critical patterns reported for the bilingual whole-group analyses in both the agreement attraction and unmarked number preference (UNP) tasks. They are reported below. FIGURE S1 shows the proportion of "acceptable" responses across tasks and TABLE S2 shows the statistical results.

Figure S1. Results of supplementary analyses on a subset of twenty heritage speakers. *(A) Attraction task*: Monolingual and heritage speakers were more likely to accept ungrammatical sentences when they contained a plural attractor, showing similar rates of attraction. *(B) UNP task*: With third person plural subjects, monolingual speakers showed an unmarked number preference (UNP). In contrast, there was no statistical difference in the heritage group, although they showed a numeric dispreference for unmarked verbs.

A   Attraction task
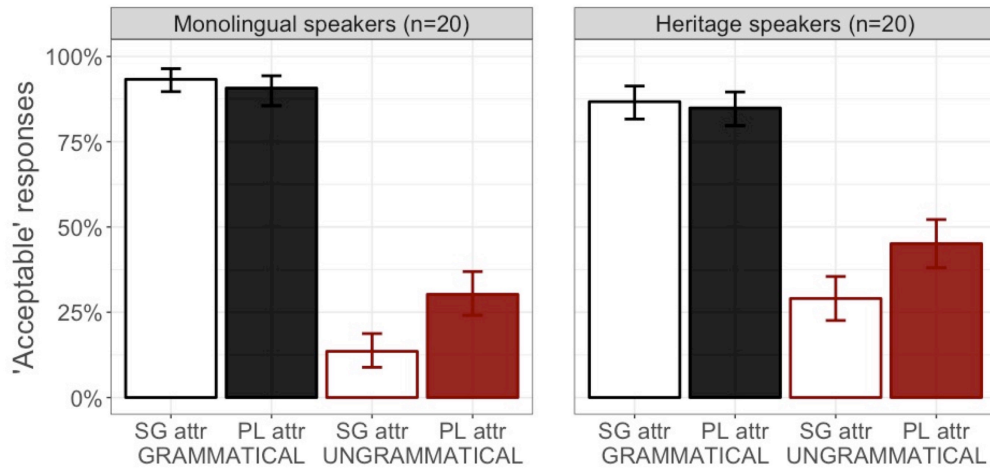
B   UNP task
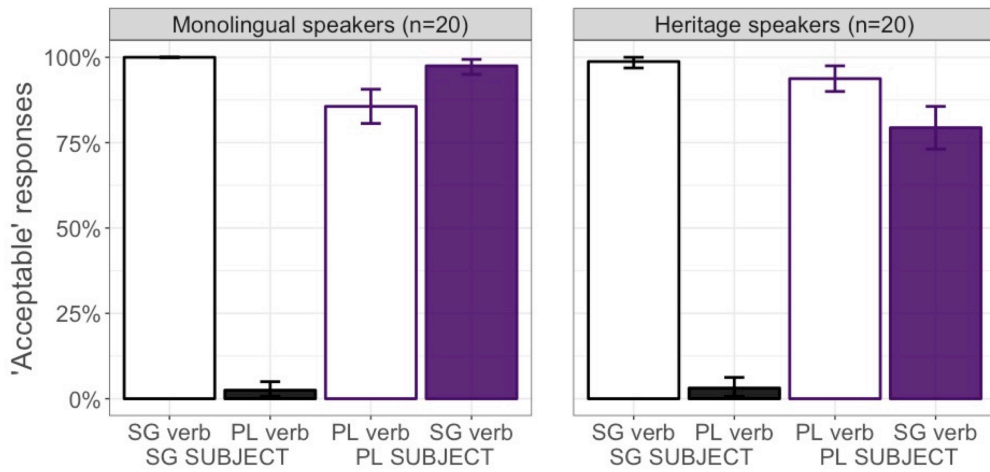
Attraction task

As in the bilingual whole-group analyses, heritage speakers showed an interaction between GRAMMATICALITY and ATTRACTOR NUMBER, consistent with attraction: ungrammatical sentences were accepted more often with plural than singular attractors, and there was no difference in grammatical sentences.

In order to statistically assess differences between heritage and monolingual speakers, judgments in ungrammatical sentences were directly compared using as predictors GROUP (*monolingual/heritage*) and ATTRACTOR NUMBER (*singular/plural*) as well as their interaction. Overall, heritage speakers accepted ungrammatical sentences more often than monolinguals, as shown by a main effect of GROUP ($\hat{\beta}$ = 1.22; *SE* = 0.48; *z* = 2.52; *p* = .012). However, their attraction rates did not differ: acceptability differences in ungrammatical sentences with and without plural attractors did not differ statistically between monolinguals and heritage speakers (ns. interaction between GROUP and ATTRACTOR NUMBER: β = -0.77; *SE* = 0.55; *z* = -1.40; *p* = .161). In brief, heritage speakers accepted ungrammatical sentences more often than monolinguals but displayed similar attraction rates.

UNP task

As in the bilingual whole-group analyses, heritage speakers showed an interaction between SUBJECT NUMBER and AGREEMENT: With singular subjects, they showed a canonical agreement effect by preferring singular to plural verbs. But with plural subjects, plural and unmarked verbs were accepted to similar extents.[3]

Although the statistical patterns replicated those of the bilingual group reported in the paper, the comparison of FIGURE 2 (in the paper) and FIGURE S1 shows that, compared to the bilingual group as a whole, the subgroup of heritage speakers rejected unmarked verbs with plural subjects to a greater extent. This pattern is reflected in the German age of acquisition effect reported in the paper, which showed that bilinguals who had acquired German earlier were less likely to accept unmarked verbs compared to bilinguals who had acquired German later. Thus, the supplementary analyses support the claim that the absence of the UNP in bilinguals was modulated by their age of acquisition of German.

In order to statistically assess between-group differences, judgments in the plural subject conditions were compared using as predictors GROUP (*monolingual/heritage*) and AGREEMENT (*agree/disagree*) as well as their interaction. The difference between the two groups was supported by a GROUP × AGREEMENT interaction ($\hat{\beta}$ = -6.55; *SE* = 1.68; *z* = -3.91; *p* = .000). With plural subjects, monolinguals showed the UNP and preferred unmarked to plural-marked verbs, whereas heritage speakers accepted both types of verbs to similar extents.

Table S2. Model results of the supplementary analyses on a subset of heritage speakers. Negative coefficients index a reduction in the proportion of "*acceptable*" responses. Significant effects at the α = .05 level are bolded.

	Heritage speakers			
	$\hat{\beta}$	SE	z	p
Attraction task				
Grammaticality	**-2.85**	**0.40**	**-7.22**	**.000**
Attractor number	0.41	0.25	1.62	.105
Grammaticality × Attractor number	**1.15**	**0.47**	**2.46**	**.014**
Attractor number: Ungram conditions	**1.00**	**0.31**	**3.26**	**.001**
Attractor number: Gram conditions	-0.44	0.45	-0.98	.328
UNP task				
Subject number	1.56	0.97	1.61	.107
Agreement	**-7.13**	**1.57**	**-4.55**	**.000**
Subject number × Agreement	**9.45**	**1.69**	**5.60**	**.000**
Agreement: SG subject conditions	**8.72**	**1.56**	**-5.60**	**.000**
Agreement: PL subject conditions	-4.01	2.92	-1.38	.169

---

[3] Due to non-convergence, the SUBJECT NUMBER by-item slope was removed in the within-group analyses. In the between-group analysis, the AGREEMENT by-participant slope was removed.