Cognitive salience of agreement features modulates language comprehension

Manuscript Pre-Print

R. Muralikrishnan[1], & Ali Idrissi[2]

[1] Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany
[2] Qatar University, Doha, Qatar

Contents

1 Introduction 3
1.1 Background 3
1.2 Previous studies 4
1.3 Arabic as a test case 8
1.4 The present study 9
2 Methods 10
2.1 Participants 10
2.2 Experimental Design 10
2.3 Materials 12
2.4 Tasks 12
2.5 Procedure 13
2.6 EEG recording, pre-processing and statistical analysis . . 15
3 Results 16
3.1 Behavioural data 16
3.2 ERP data 19
3.3 Summary of Results 38
4 Discussion 40
4.1 Hierarchical Modulation of Effects 41
4.2 Salience-weighted Feature Hierarchy 43
4.3 Language-specific Feature Weightings 45
4.4 Implications and Outlook 47
4.5 Conclusion 50

Abstract

The brain establishes relations between elements of an unfolding sentence in order to incrementally build a representation of who is doing what based on various linguistic cues. Many languages systematically mark the verb and/or its arguments to imply the manner in which they are related. A common mechanism to this end is subject-verb agreement, whereby the marking on the verb covaries with one or more of the features such as person, number and gender of the more agent-like argument in a sentence. The cross-linguistic variability of these features would suggest that they may modulate language comprehension differentially based on their relative weightings in a given language. To test this, we investigated the processing of verb agreement in simple intransitive Arabic sentences in a visual event-related brain potential (ERP) study. Specifically, we examined the differences, if any, that ensue in the processing of person, number and gender features during online comprehension, employing sentences in which the verb either showed full agreement with the subject noun (singular or plural) or did not agree in one of the features. ERP responses were measured at the post-nominal verb. Results showed a biphasic negativity – late-positivity effect when the verb did not agree with its subject noun in either of the features, in line with similar findings from other languages. Crucially however, the biphasic effect for agreement violations was systematically graded based on the feature that was violated, which is a novel finding in view of results from other languages. Furthermore, this graded effect was qualitatively different for singular and plural subjects based on the differing salience of the features for each subject-type. These results suggest that agreement features, varying in their cognitive salience due to their language-specific weightings, differentially modulate language comprehension.

1 Introduction

1.1 Background

An important evolutionary function of human language is to convey ecologically relevant information, such as the state of affairs of the entities in the immediate environment and event descriptions about who is doing what and to whom, so as to elicit an appropriate response in the given context. In order to comprehend the message, the brain thus has to be able to decipher these details incrementally by establishing the relations between the various elements in the unfolding utterance. Furthermore, these relationships must be constructed from the linguistic input even in the presence of intervening material separating the related elements.

Languages differ considerably in the mechanisms they employ in order to establish relations between an event being described in an utterance and the participant(s) of the event. One frequent device attested in many languages is to overtly mark such relations on certain elements morphosyntactically (Nichols & Bickel, 2013), such that there is a 'systematic covariance between a semantic or formal property of one element and a formal property of another (Steele, 1978, p. 610)', commonly referred to as agreement. Agreement between a verb and its arguments may be based on some or all of the properties or features of the arguments concerned, such as person, number and gender of the nouns, collectively called agreement features (Wechsler, 2009) or phi-features. Greenberg (1963) formulated a number of generalisations based on how these features show patterns of dependency and their frequency of occurrence relative to each other across a number of languages of the world. One such generalisation is the Feature Hierarchy shown in (1) below, which has been an important explanatory concept to account for the cross-linguistic diversity in how these features co-occur.

(1) Feature Hierarchy: Person > Number > Gender

Further sub-hierarchies of these features have been proposed in accounting for how they influence certain linguistic phenomena (Silverstein, 1976; Shlonsky, 1989; Corbett, 2000a,b) in a variety of typologically diverse languages. The question of course is, whether such a hierarchy of agreement features identified on a cross-linguistic basis plays a role in online language comprehension. Indeed, such a systematic organisation of features is said to be a grammaticalised representation of fundamental cognitive categories (Harley & Ritter, 2002), with the relative position of a feature in the hierarchy reflecting the cognitive salience of the feature in relation to the other features, which in turn is said to correlate with dissociations in their online processing (De Vincenzi, 1999; Carminati, 2005; Acuña-Fariña, 2009). For instance, Carminati (2005, p. 263) proposes a Feature Strength Hypothesis mirroring the hierarchy in (1), suggesting that the cognitive significance of a feature should directly correlate with its relative hierarchical importance in language processing. This would imply that a differential contribution, if any, of the agreement features must be observable in their respective neural processing correlates.

1.2 Previous studies

A number of studies have examined the processing of agreement features in different languages. In the behavioural domain, many studies employed a completion task involving sentences with agreement errors in the context of complex phrases, such as '*The key to the cabinets are on the table.*' (see for instance Bock & Miller, 1991; Vigliocco, Butterworth, & Semenza, 1995; Hartsuiker, Antón-Méndez, & van Zee, 2001; Eberhard, Cutting, & Bock, 2005; Haskell, Thornton, & MacDonald, 2010, among others), whilst some involved a self-paced reading task (Wagers, Lau, & Phillips, 2009; Lago, Shalom, Sigman, Lau, & Phillips, 2015; Tucker, Idrissi, & Almeida, 2015), in some cases with eye-tracking (Pearlmutter, Garnsey, & Bock, 1999; Dillon, Mishler, Sloggett, & Phillips, 2013). An overarching finding from this line of research is that, agreement processes in sentence comprehension and production are complex, and involve factors that are 'not only syntactic, not only semantic, and not only pragmatic, but all of these at once' (Eberhard et al., 2005, p. 531). The cross-linguistic importance of agreement as a cue to sentence interpretation was demonstrated in a range of offline experiments conducted within the scope of the Competition Model (e.g., MacWhinney, Bates, & Kliegl, 1984; Bates & MacWhinney, 1989; Bates, 1999). These studies further showed that the strength of agreement as a cue to interpretation varies across different languages (for an overview, see Bates, McNew, Devescovi, & Wulfeck, 2001).

In order to gain insights into the neurocognitive mechanisms underlying the processing of agreement features, several previous studies have employed the ERP technique, which is particularly well-suited for studying language comprehension in real time thanks to its excellent temporal resolution. ERP studies investigating subject-verb agreement typically employ a violation paradigm, whereby the verb does not agree with its subject in the feature of interest. When ERPs measured at the anomalous verb are compared with those measured at the verb that shows correct agreement, this then is said to shed light on the neural correlates of processing the agreement feature under investigation.

Kutas & Hillyard (1983) reported one of the earliest ERP studies on subject-verb agreement in English and found that agreement violations elicited ERP effects that are different in scalp distribution to those elicited by semantic anomalies. Ever since, several ERP studies on the processing of agreement features in a number of languages have been reported (see Molinaro, Barber, & Carreiras, 2011, for a detailed review), the vast majority of which are on Indo-European languages. Results from these studies show that agreement violations of various types of dependencies (i.e., subject-verb, adjective-noun, article-noun etc.) generally elicit a left-anterior negativity (LAN) effect around 300 to 500 ms after the onset of the violation followed by a late-positivity (P600) effect in the 500 to 700 ms window and/or later. This is true for the number feature in Dutch (Hagoort, Brown, & Groothusen, 1993; Kaan, 2002), English (Osterhout & Mobley, 1995; Coulson, King, & Kutas, 1998), Finnish (Palolahti, Leino, Jokela, Kopra, & Paavilainen, 2005; Leinonen, Brattico, Järvenpää, & Krause, 2008), German (Roehm, Bornkessel, Haider, & Schlesewsky, 2005), Italian (Angrilli et al., 2002; De Vincenzi et al., 2003) and Spanish (Barber & Carreiras, 2005; Mancini, Molinaro, Rizzi, & Carreiras, 2011a); for the gender feature in German (Gunter, Friederici, & Schriefers, 2000), Italian (Molinaro, Vespignani, & Joba, 2008) and Spanish (Barber & Carreiras, 2005; Alemán Bañón & Rothman, 2016); for the person feature in German (Rossi, Gugler, Hahne, & Friederici, 2005) and Spanish (Silva-Pereyra & Carreiras, 2007). However, there are a few studies in which an N400 effect ensued instead of a LAN, or no negativity effect ensued at all. These include studies that examined number violations in Basque (Díaz, Sebastián-Gallés, Erdocia, Mueller, & Laka, 2011; Zawiswewski, Santestegan, & Laka, 2016, no negativity in the former, N400-P600 in the latter) and English (Kaan, Harris, Gibson, & Holcomb, 2000, no negativity); gender violations in Dutch (Hagoort et al., 1993, N400-P600), French (Frenck-Mestre, Osterhout, McLaughlin, & Foucart, 2008, no negativity), Hebrew (Deutsch & Bentin, 2001, N400-P600, with an animacy interaction), Spanish (Guajardo & Wicha, 2014); gender, number, and combined number-gender and person-gender violations in Hindi (Nevins, Dillon, Malhotra, & Phillips, 2007, no negativity), and person violations in Basque (Zawiszewski & Friederici, 2009; Zawiswewski et al., 2016, N400-P600) and Spanish (Mancini et al., 2011a, N400-P600).

Furthermore, a handful of studies on subject-verb agreement compared a pair of feature violations with each other or with a combined violation of more than one feature, reporting a modulation of effects (i.e., a quantitative ERP difference) based on the violating feature(s). These include the study on Basque cited above, in which the late-positivity was larger for person and combined person-number violations compared to number violations (Zawiswewski et al., 2016), the study on Hindi mentioned above, in which the late-positivity was found to be larger for combined person-gender violations versus number or gender or combined number-gender violations (Nevins et al., 2007), a study on Spanish that reported a larger late-positivity for combined person-number violations versus person or number violations (Silva-Pereyra & Carreiras, 2007), with the latter concluding that these results do not support the Feature Strength Hypothesis mentioned earlier. By contrast, in another Spanish study, Mancini et al. (2011a) reported a different distribution of the late-positivity in addition to a qualitative difference in the negativities, namely a LAN effect for the number feature and an N400 effect for the person feature.

To briefly summarise these findings, whilst the LAN-P600 biphasic ERP effect is more common for agreement violations of many types, it is by no means universal, especially for subject-verb agreement. An N400 is elicited instead of a LAN in many cases, and no negativity at all ensues in some cases. Thus, neurophysiological evidence or counter-evidence for a hierarchy of features based on differences in their cognitive salience is inconclusive, with some studies reporting processing differences between the agreement features and others countering this, sometimes within the same language (compare for instance Alemán Bañón & Rothman, 2016; Silva-Pereyra & Carreiras, 2007). Further, a crucial aspect to note in this regard is that, although each feature has been studied in detail in a number of languages, and in some cases combinations of features compared with each other, to our knowledge, none of the studies reported to date compared the processing of person, number and gender agreement systematically within a single experiment. However, such a comparison would be imperative in order to shed light on whether the hierarchy of features postulated based on differences in their cognitive salience and distribution across languages has a neural equivalent to it.

Does the hierarchy of agreement features indeed have a neural equivalent to it? If so, what is the nature of such a dissociation in neural correlates: is it a quantitative ERP difference (i.e., graded amplitude modulation of effects), or a qualitative one (i.e., different ERP effects)? Do the specific syntactic properties of a language interact with such a dissociation? In other words, do language-specific weightings of the features matter? In the following section, we argue that Arabic provides an ideal testing ground for investigating this.

1.3 Arabic as a test case

In order to investigate whether the hierarchy of agreement features postulated based on differences in their cognitive salience has a neural equivalent to it, certain properties appear to be crucial in determining the language of choice. A first such property is that the language in question should be morphologically rich such that all three agreement features, namely person, number, and gender are overtly marked on the verb. A second and related property that becomes relevant is that these features should be expressed independently of each other, i.e., they are not fused together nor underspecified for a certain feature in most cases such that the violation of each feature can be studied without simultaneously violating other features. For instance, commonly studied morphologically rich languages such as German or Spanish underspecify gender in verb agreement. A third property that is desirable for present purposes is that the agreement features are not merely formal in the language in question, i.e., the realisation of one or more features is sometimes dependent upon certain structural properties of an utterance, such that changes in that property give rise to a concomitant change in the feature marking on the verb despite the subject being identical in both instances.

Arabic is a case in point in this regard. A Semitic language with upwards of 290 million speakers (Simons & Fenning, 2017) in countries spread around a great part of West Asia and North Africa, Arabic exhibits a typical example of Diglossia (Ferguson, 1959), whereby a standard written variety for formal purposes and a number of colloquially spoken dialects exist in parallel. Modern Standard Arabic (Ryding, 2005) is the written standard taught in schools and universities (and the variety used for the stimuli here). Arabic is ideally suited to investigate possible processing differences between the agreement features in the context of subject-verb agreement mainly due to the fact that all three agreement features are marked on the verb (with a few exceptions: gender is not marked in the first person). Further, verb agreement is not merely formal in Arabic. Thus, agreement is always full in the subject-verb (SV) order, i.e., the verb agrees with the subject noun in all three features, whereas in the VS order, if the subject noun is third-person plural and overt, number marking on the verb defaults to the singular form, but if the plural subject were to be dropped from the utterance, the sentence-initial verb is marked plural. That is, number marking on the verb is idiosyncratic depending upon the structure in which agreement is computed between a verb and its subject noun and the overtness of the latter. These properties of Arabic allow investigating potential differences between the agreement features in language processing not possible in quite a comparable manner in many other widely-spoken languages. Furthermore, neurolinguistic investigations of diverse languages are crucial for gaining a broader understanding of the neural underpinnings of language (Bornkessel-Schlesewsky & Schlesewsky, 2016), and as such, neurophysiological studies on Arabic are an important contribution from a widely-spoken but extremely understudied language.

1.4 The present study

We report a visual ERP study here, in which we investigated the processing of subject-verb agreement in simple intransitive SV sentences in Arabic. In this experiment, ERP responses were measured at the post-nominal verb as participants read the stimuli in a rapid serial visual presentation. We specifically examined the differences, if any, that ensue in the processing of person, number and gender features during online comprehension, employing sentences in which the verb either showed full agreement with the subject noun (singular or plural) or did not agree in one of the features. In the following section, we describe the methods, materials and procedure, and report results from the study at the critical position, namely, at the verb and discuss the findings in further sections below.

2 Methods

2.1 Participants

Thirty-four persons, most of them students at the United Arab Emirates University in Al Ain, participated in the experiment after giving informed consent, and received monetary compensation for their participation. All participants were right-handed native Arabic speakers, with normal or corrected- to-normal vision and normal hearing. Three further participants had to be excluded from the final data analyses on the basis of excessive EEG artefacts and/or too many errors in the behavioural control task.

2.2 Experimental Design

We employed an experimental design constituting Arabic sentences in the SV order with singular and plural subject nouns. All critical sentences were of the form adverb - subject - verb - prepositional phrase. The adverb was always *'yesterday'*, subject nouns were always animate and human, and the verbs were in the simple past tense. The verb in each sentence was either in correct agreement with its subject noun, or it did not agree in either person, or number, or gender with the subject (see Section 2.3 for details about the materials). Note that any violation was always with one feature only. This yielded a design consisting of eight critical conditions that differed based on whether the subject-type (ST) was singular or plural; and whether the condition-type (CT) was acceptable or person violation or number violation or gender violation. Table 1 provides an overview of the factors and their levels, with the condition codes relevant to each level, as well as examples pertaining to each condition. Such a design would enable observing the ERP effects at the position of the verb when information about the subject noun is already available. Any differences in ERP effects at the verb would reflect the differences in the processing of the different agreement features.

Table 1: Critical factors and their individual levels, with condition codes and a complete set of corresponding example stimuli.

Factor	Level	Conditions
ST : Subject-Type	S : Singular noun	SACP, SGEN, SNUM, SPER have a human singular subject, half of them feminine and half masculine
	P : Plural noun	PACP, PGEN, PNUM, PPER have a human plural subject, half of them feminine and half masculine
CT : Condition-Type	ACP : Acceptable	SACP, PACP have a verb that shows full agreement with its subject noun
	GEN : Violation of Gender Agreement	SGEN, PGEN have a verb that violates gender agreement
	NUM : Violation of Number Agreement	SNUM, PNUM have a verb that violates number agreement
	PER : Violation of Person Agreement	SPER, PPER have a verb that violates person agreement

SACP
بالأمس الممرضة حضَرَتْ إلى المكان.
ˀl-makaan ˀilaa ħadˤar-at ˀl-mumarridˤa bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-3sg.F}}$ [DEF-nurse]$_{\text{3sg.F}}$ [yesterday]

*** SGEN**
بالأمس الممرضة حضَرَ إلى المكان.
ˀl-makaan ˀilaa ħadˤar-a ˀl-mumarridˤa bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-3sg.M}}$ [DEF-nurse]$_{\text{3sg.F}}$ [yesterday]

*** SNUM**
بالأمس الممرضة حضَرْنَ إلى المكان.
ˀl-makaan ˀilaa ħadˤar-na ˀl-mumarridˤa bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-3pl.F}}$ [DEF-nurse]$_{\text{3sg.F}}$ [yesterday]

*** SPER**
بالأمس الممرضة حضَرْتُ إلى المكان.
ˀl-makaan ˀilaa ħadˤar-tu ˀl-mumarridˤa bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-1sg}}$ [DEF-nurse]$_{\text{3sg.F}}$ [yesterday]

Singular Conditions. Right to left Arabic text with corresponding glosses.
Intended meaning: *'Yesterday the female nurse came to the place.'*
A * before condition labels indicates ungrammaticality.

PACP
بالأمس الممرضات حضَرْنَ إلى المكان.
ˀl-makaan ˀilaa ħadˤar-na ˀl-mumarridˤaat bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-3pl.F}}$ [DEF-nurses]$_{\text{3pl.F}}$ [yesterday]

*** PGEN**
بالأمس الممرضات حضَرُوا إلى المكان.
ˀl-makaan ˀilaa ħadˤa-ruu ˀl-mumarridˤaat bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-3pl.M}}$ [DEF-nurses]$_{\text{3pl.F}}$ [yesterday]

*** PNUM**
بالأمس الممرضات حَضَرَتْ إلى المكان.
ˀl-makaan ˀilaa ħadˤar-at ˀl-mumarridˤaat bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-3sg.F}}$ [DEF-nurses]$_{\text{3pl.F}}$ [yesterday]

*** PPER**
بالأمس الممرضات حضَرْنا إلى المكان.
ˀl-makaan ˀilaa ħadˤar-naa ˀl-mumarridˤaat bil-ˀams
[DEF-place] [to] [come]$_{\text{Past-1pl}}$ [DEF-nurses]$_{\text{3pl.F}}$ [yesterday]

Plural Conditions. Right to left Arabic text with corresponding glosses.
Intended meaning: *'Yesterday the female nurses came to the place.'*
A * before condition labels indicates ungrammaticality.

2.3 Materials

As a first step, 120 intransitive verbs were used to construct acceptable sentences with singular nouns, half of which were masculine and the other half feminine. The plural acceptable sentences were then generated from these such that the subject nouns and verbs now marked plurality. Finally, three violation conditions were generated from each of the 240 acceptable sentences such that the agreement marking on the verb either violated person or number or gender in each instance (i.e., never was a combination of these features violated). This resulted in a total of 120 sets of sentences in eight critical conditions, thus 960 critical sentences.

Fillers were constructed to ensure that the sentence structure is not predictable based on the first word alone and that the overall number of acceptable versus violation sentences is counterbalanced. Thus there were sentences involving subject and object relative clauses, verb-initial orders and a few semantic anomalies. The 960 target sentences were distributed into five unique sets according to a latin square, such that each list contained 36 sentences per critical condition. Of the resulting 288 critical sentences in a given list, 72 were acceptable and the rest were violations. Fillers were interspersed with these lists such that, overall, each stimulus list ended up with an equal number of acceptable and violation sentences, with equal number of sentences with a masculine or feminine, singular or plural noun. This resulted in a total of 600 sentences per stimulus list (288 critical sentences and 312 fillers). These were each pseudo-randomised to obtain five stimulus lists, one of which was used for every participant. The presentation of the randomised lists was counterbalanced across participants.

2.4 Tasks

Given the use of a violation paradigm, an acceptability judgement task followed the presentation of each stimulus sentence, which required 'yes' or 'no' as answers. In addition, in order to ensure that participants would process the sentences attentively, a probe word detection task followed the acceptability judgement task. The probe task was constructed in such a way that an equal number of trials requred 'yes' or 'no' as answers. If the probe word was one of the words that occurred in the preceding stimulus, this required a 'yes', whereas if it was novel, it required a 'no'.awk workshop Crucially, the word position from which the probe word was chosen was equiprobable across the experiment as well as within each condition, which meant that participants had to be very attentive throughout stimulus presentation so as to perform the task correctly.

2.5 Procedure

The experiment was performed in the EEG laboratory of the Department of Linguistics at the United Arab Emirates University in Al Ain. The methods and procedure employed in the experiment were in accordance with local regulations, and followed the guidelines of the Helsinki declaration. Participants filled an Edinburgh-Handedness questionnaire in Arabic, and dominant right-handers alone were accepted for participation. They received printed instructions about the experiment and the task they had to perform.

Stimuli were presented using the Presentation software (www.neurobs.com) that recorded, among other things, the trial number, reaction time and the button responses. The brightness and contrast settings of the monitor were maintained the same for all the participants.

After setting up the electrode cap, the participant moved to a sound-proof chamber, where they were seated on a comfortable chair and were requested to avoid abrupt and drastic movements, especially of the head. Then the so-called 'resting EEG' was recorded for possible frequency-based EEG analyses later, where the participant had to sit still for two minutes with no specific task to perform. Two more minutes of resting EEG was recorded, but now the participant had to close their eyes. After a short pause, the experimental session commenced, which consisted of a short practice followed by the actual experiment.

The structure of each trial in the experiment was as follows. The flat-screen LCD monitor was clear before the trial commenced. A fixation asterisk was shown in the centre of the screen for 500 ms, after which the screen became blank for 100 ms. Then the rapid serial visual presentation of the stimulus sentence started. Each word or phrase appeared in the centre of the screen and remained for 600 ms, after which the screen became blank for 100 ms before the appearance of the next word. Phrases containing two words were presented for 750 ms. After the last word of the stimulus sentence was presented, the screen was blank for 500 ms. Following this, a pair of smileys appeared on screen, which prompted the participant to judge the acceptability of the sentence that just preceded. After a maximum of 2000 ms or after a button press, whichever was earlier, the screen became blank again for 500 ms. A time-out was registered when no button was pressed within 2000 ms. Then, a probe word appeared in the middle of the screen for a maximum of 2000 ms, within which the participant had to detect whether the word was present in the preceding stimulus sentence or not. When no button was pressed within 2000 ms, a time-out was registered. At the end of the trial, the screen became blank for 1500 ms (inter-stimulus interval) before the next trial started.

Before the actual experiment commenced, there was a short practice consisting of twelve trials, which helped participants to get used to the task and to feel comfortable about the pace of the trials and the blinking regime. For a given participant, none of the experimental stimuli occurred in their practice phase. The task was identical to that of the experiment phase. The EEG of the participants was not recorded in this phase.

In the main phase of the experiment, one of the five sets of materials as mentioned above was chosen to be presented in 12 blocks of 50 trials each. There were equal number of probe words that required 'Yes' or 'No' as answers in each block. Half the number of participants had the 'Yes' button on the right side and the other half had it on the left side so as to counterbalance for any right-dominance effects. The 'Yes' button being on the right or left was also counterbalanced across the stimuli sets. There was a short pause between blocks. Resting EEG was again recored at the end of the experimental session.

2.6 EEG recording, pre-processing and statistical analysis

The EEG was recorded by means of 25 AgAgCl active electrodes fixed at the scalp by means of an elastic cap (Easycap GmbH, Herrsching, Germany). AFZ served as the ground electrode. Recordings were referenced to the left mastoid, but re-referenced to the average of linked mastoids offline. The electrooculogram (EOG) was monitored by means of elecrodes placed at the outer canthus of each eye for the horizontal EOG and above and below the participant's right eye for the vertical EOG. Electrode impedances were kept below appropriate levels such as to ensure a good quality signal with minimal noise. All EEG and EOG channels were amplified using a BrainAmp amplifier (Brain Products GmbH, Gilching, Germany) and recorded with a digitisation rate of 250 Hz. The EEG data thus collected was pre-processed for further analysis using a $0.3-20$Hz bandpass filter in order to remove slow signal drifts. The statistical analyses were performed on this data, but an 8.5 Hz low-pass filter was further applied on the data for achieving smoother ERP plots.

ERPs were calculated for each participant from 200 ms before the onset of the verb until 1200 ms after onset (so -200 ms to 1200 ms). These were averaged across items per condition per participant before computing the grand-average ERPs across participants per condition. Repeated-measures analyses of variance (ANOVAs) were computed for the statistical analysis of the ERP data, involving the within-participants factors subject-type (ST) and condition-type (CT) for mean amplitude values per time-window per condition in 4 lateral Regions of Interest (ROIs) and 6 midline ROIs. The lateral ROIs were defined as follows: LA, comprised of the left-anterior electrodes F7, F3, FC5 and FC1; LP, comprised of the left-posterior electrodes P7, P3, CP5 and CP1; RA, comprised of the right-anterior electrodes F8, F4, FC6 and FC2; and RP, comprised of the right-posterior electrodes P8, P4, CP6 and CP2. Each of the 6 midline electrodes FZ, FCZ, CZ, CPZ, PZ and POZ constituted an individual midline ROI of the same name respectively.

The statistical analysis of the ERP data was carried out in a hierarchical manner, that is to say, only interactions that are at least marginally significant were resolved. To avoid excessive type 1 errors due to violations of sphericity, the correction of Huynh & Feldt (1970) was applied when the analysis involved factors with more than one degree of freedom in the numerator. Factors with more than two levels, which resulted in a significant effect, were further resolved by comparing their individual levels pairwise. An effect resulting from such an individual pairwise comparison would be reported as significant only if it was still significant after applying the modified Bonferroni correction (Keppel, 1991). Further, given a resolvable effect was significant both with and without a ROI interaction in a certain analysis, only the interaction involving ROI was resolved further.

3 Results

3.1 Behavioural data

The mean acceptability ratings for the stimuli, as well as the probe detection accuracy for the critical conditions, shown in Table 2, were calculated using the behavioural data collected during the experiment. Only those trials in which the acceptability judgement task following each trial was performed (i.e., not timed out) were considered for the analysis. Further, the acceptability data presented here pertain only to those trials in which the participants performed the probe detection task correctly (see Section 2.4 for details about the tasks). Acceptability was highest for the conditions with no violations, whereas it was the lowest for conditions with number violations. Across conditions, acceptability was relatively slightly higher for conditions with a plural subject as compared to the corresponding conditions with a singular subject. The overall accuracy was very high across all conditions. Owing to the fact that the reaction time data are not time-locked to the critical manipulation in the stimulus sentences, they are not reported here (but are available from the corresponding author on request).

Table 2: Acceptability and accuracy

Condition	Acceptability %	SD	Accuracy %	SD
SACP	90.51	11.86	95.74	4.54
SGEN	23.68	29.67	94.48	4.10
SNUM	25.55	28.76	94.16	5.63
SPER	47.62	22.53	95.58	3.93
PACP	91.68	7.21	95.66	4.44
PGEN	67.00	22.99	94.42	4.89
PNUM	33.91	30.82	94.91	4.45
PPER	51.83	30.02	93.57	5.22

The statistical analysis of the behavioural data was performed by means of ANOVAs involving the within-subjects factors subject-type (ST) and condition-type (CT), and the random factors participants (F1) and items (F2). The statistical analysis was carried out in a hierarchical manner, that is to say, only interactions that are at least marginally significant were resolved. To avoid excessive type 1 errors due to violations of sphericity, the correction of Huynh & Feldt (1970) was applied when the analysis involved factors with more than one degree of freedom in the numerator. Factors with more than two levels, which resulted in a significant effect, were further resolved by comparing their individual levels pairwise. An effect resulting from such an individual pairwise comparison would be reported as significant only if it was still significant after applying the modified Bonferroni correction (Keppel, 1991). Further, given a resolvable effect was significant both with and without a ROI interaction in a certain analysis, only the interaction involving ROI was resolved further. Table 3 shows a summary of effects on the behavioural data collected during the experiment.

Table 3: ANOVA: Behavioural Data

Effects on Acceptability	DF	F1:Participants		DF	F2:Items	
◆ ST	1,33	56.39	⋆⋆⋆	1,59	246.80	⋆⋆⋆
◆ CT	3,99	105.60	⋆⋆⋆	3,177	642.50	⋆⋆⋆
└CT = ACP+PER ◆ CT	1,33	100.70	⋆⋆⋆	1,59	747.10	⋆⋆⋆
└CT = ACP+NUM ◆ CT	1,33	138.20	⋆⋆⋆	1,59	2503.00	⋆⋆⋆
└CT = NUM+PER ◆ CT	1,33	54.05	⋆⋆⋆	1,59	115.30	⋆⋆⋆
└CT = ACP+GEN ◆ CT	1,33	118.10	⋆⋆⋆	1,59	1298.00	⋆⋆⋆
└CT = GEN+PER ◆ CT	1,33	5.40	⋆	1,59	8.99	⋆⋆
└CT = NUM+GEN ◆ CT	1,33	48.57	⋆⋆⋆	1,59	130.90	⋆⋆⋆
◆ ST x CT	3,99	30.96	⋆⋆⋆	3,177	142.60	⋆⋆⋆
└ST = S ◆ CT	3,99	122.40	⋆⋆⋆	3,177	613.30	⋆⋆⋆
└ST = S └CT = ACP+PER ◆ CT	1,33	100.90	⋆⋆⋆	1,59	483.80	⋆⋆⋆
└ST = S └CT = ACP+NUM ◆ CT	1,33	141.10	⋆⋆⋆	1,59	1512.00	⋆⋆⋆
└ST = S └CT = NUM+PER ◆ CT	1,33	80.24	⋆⋆⋆	1,59	117.50	⋆⋆⋆
└ST = S └CT = ACP+GEN ◆ CT	1,33	138.30	⋆⋆⋆	1,59	1967.00	⋆⋆⋆
└ST = S └CT = GEN+PER ◆ CT	1,33	108.10	⋆⋆⋆	1,59	151.00	⋆⋆⋆
└ST = P ◆ CT	3,99	52.95	⋆⋆⋆	3,177	315.50	⋆⋆⋆
└ST = P └CT = ACP+PER ◆ CT	1,33	66.45	⋆⋆⋆	1,59	487.80	⋆⋆⋆
└ST = P └CT = ACP+NUM ◆ CT	1,33	102.80	⋆⋆⋆	1,59	1077.00	⋆⋆⋆
└ST = P └CT = NUM+PER ◆ CT	1,33	11.68	⋆⋆	1,59	53.87	⋆⋆⋆
└ST = P └CT = ACP+GEN ◆ CT	1,33	44.88	⋆⋆⋆	1,59	188.70	⋆⋆⋆
└ST = P └CT = GEN+PER ◆ CT	1,33	14.56	⋆⋆⋆	1,59	71.44	⋆⋆⋆
└ST = P └CT = NUM+GEN ◆ CT	1,33	48.92	⋆⋆⋆	1,59	253.80	⋆⋆⋆

There were main effects of subject-type and condition-type on the acceptability in the analysis by participants as well as the analysis by items. Resolving the effect of condition-type by comparing the condition-types pairwise showed a significant simple effect of condition-type for all possible comparisons in both analyses, with the largest effect for the ACP + NUM comparison. The interaction subject-type x condition-type was significant in both the analyses, which when resolved for subject-type showed an effect of condition-type for both subject-types. This was further resolved by comparing the condition-types pairwise, which showed a simple effect of condition-type for all comparisons except NUM + GEN in both analyses when the subject was singular. When the subject was plural, there was a simple effect of condition-type for all comparisons in both analyses. There were no effects on the probe detection accuracy.

3.2 ERP data

The ERPs at the verb are shown in Figure 1 for the singular subject conditions, and in Figure 2 for the plural subject conditions. Figure 3 shows the topographic map of the ERPs at the position of the verb in the 300 − 500 ms, 400 − 600ms and 700 − 850 ms time-windows for the two subject-types in the three violation conditions, after the effects for the acceptable condition in each case has been subtracted. Four time-windows were chosen for analysis based on ERP components that have been known to be relevant for processing agreement (See Section 2.6 for details regarding the statistical analysis).

3.2.1 Time-window 300–500 ms

The predominant effect in this time-window is the graded negativity for the violation conditions as opposed to the acceptable conditions, with the gradedness differing based on whether the subject was singular or plural. Table 4 shows a summary of all the effects that reached at least marginal significance at the position of the verb in the 300 − 500 ms time-window.

There were no main effects. The interaction ROI x subject-type was significant in the midline regions, which when resolved for the individual ROIs showed a simple effect of subject-type in the fronto-central midline region. The interaction subject-type x condition-type was significant in the lateral and midline regions. This effect was resolved in the lateral regions alone (due to a ROI interaction in the midline regions), which showed an effect of condition-type when the subject was singular. This was resolved further by comparing the condition-types pairwise, which showed a simple effect of condition-type for the comparison PER + NUM alone.

The three-way interaction ROI x subject-type x condition-type was significant in the midline regions alone, which was resolved for the individual ROIs. This showed a marginal effect of subject-type x condition-type in the fronto-central midline region, and the effect was significant in the central, centro-parietal, parietal and parieto-occipital midline regions, which was resolved for subject-type in each region. When the subject-type was singular, there was an effect of condition-type in the central, centro-parietal, parietal and parieto-occipital midline regions. Resolving this effect by comparing the condition-types pairwise revealed the following. There was a simple effect of condition-type in the central, centro-parietal and parieto-occipital midline regions for the comparison ACP + NUM, and this effect was marginal in the parietal midline region. Comparing PER + NUM, there was a simple effect of condition-type in the centro-parietal and parieto-occipital midline regions, with this effect being marginal in the central midline region. Comparing ACP + GEN, there was simple effect of condition-type in the central, parietal and parieto-occipital midline regions, with this effect being marginal in the centro-parietal region. Comparing PER + GEN, there was a simple effect of condition-type in the parieto-occipital midline region, with this effect being marginal in central, centro-parietal and parietal midline regions. When the subject-type was plural, there was an effect of condition-type in the fronto-central, central and centro-parietal midline regions. Resolving this effect by comparing the condition-types pairwise revealed the following. There was a simple effect of condition-type in the fronto-central, central and centro-parietal midline regions for the comparison PER + GEN. For the comparison NUM + GEN, there was a marginal effect of condition-type in the central and centro-parietal midline regions.

Figure 1: ERPs at the verb: Singular subject conditions.

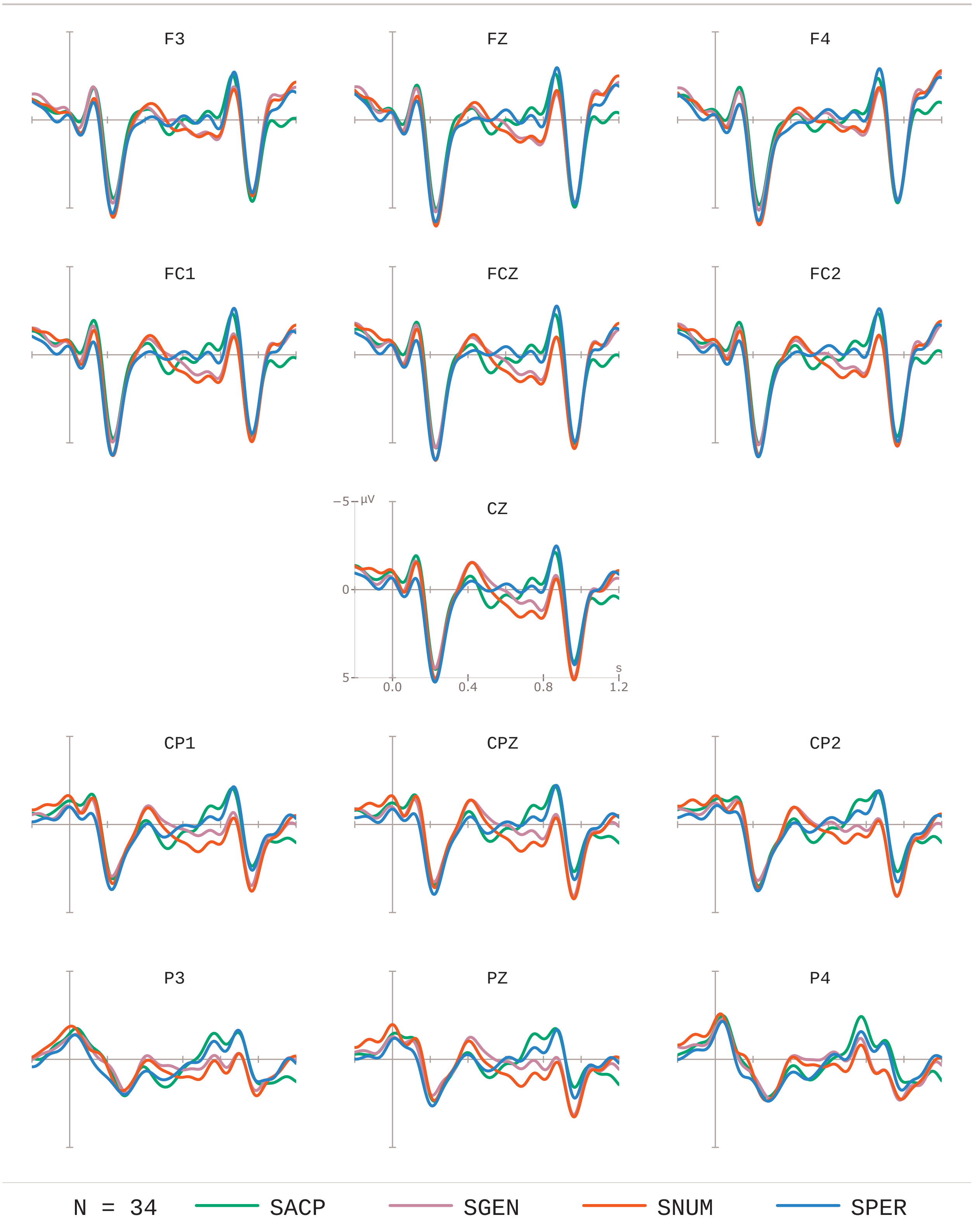

Figure 2: ERPs at the verb: Plural subject conditions.

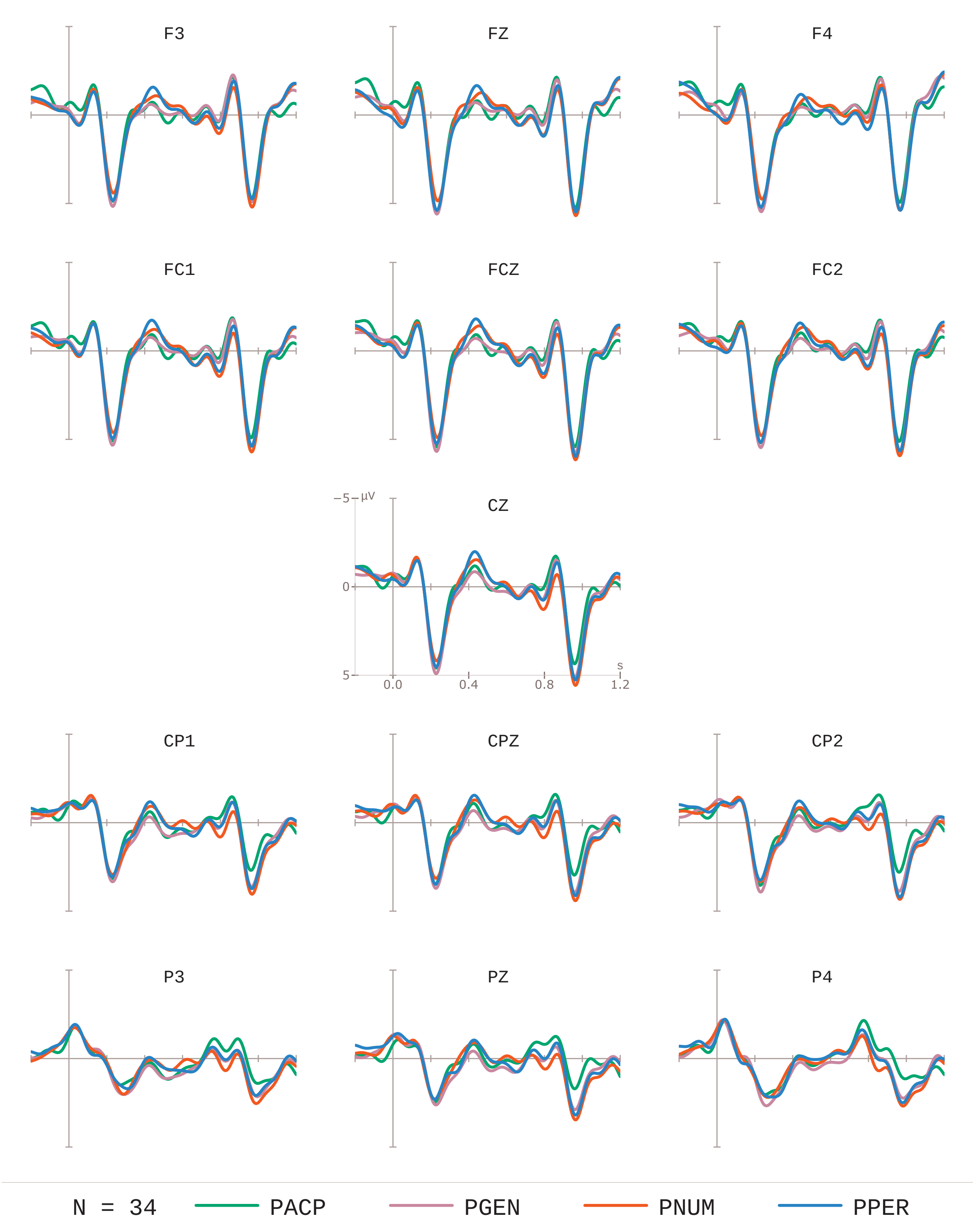

Figure 3: Topography of effects at the verb.

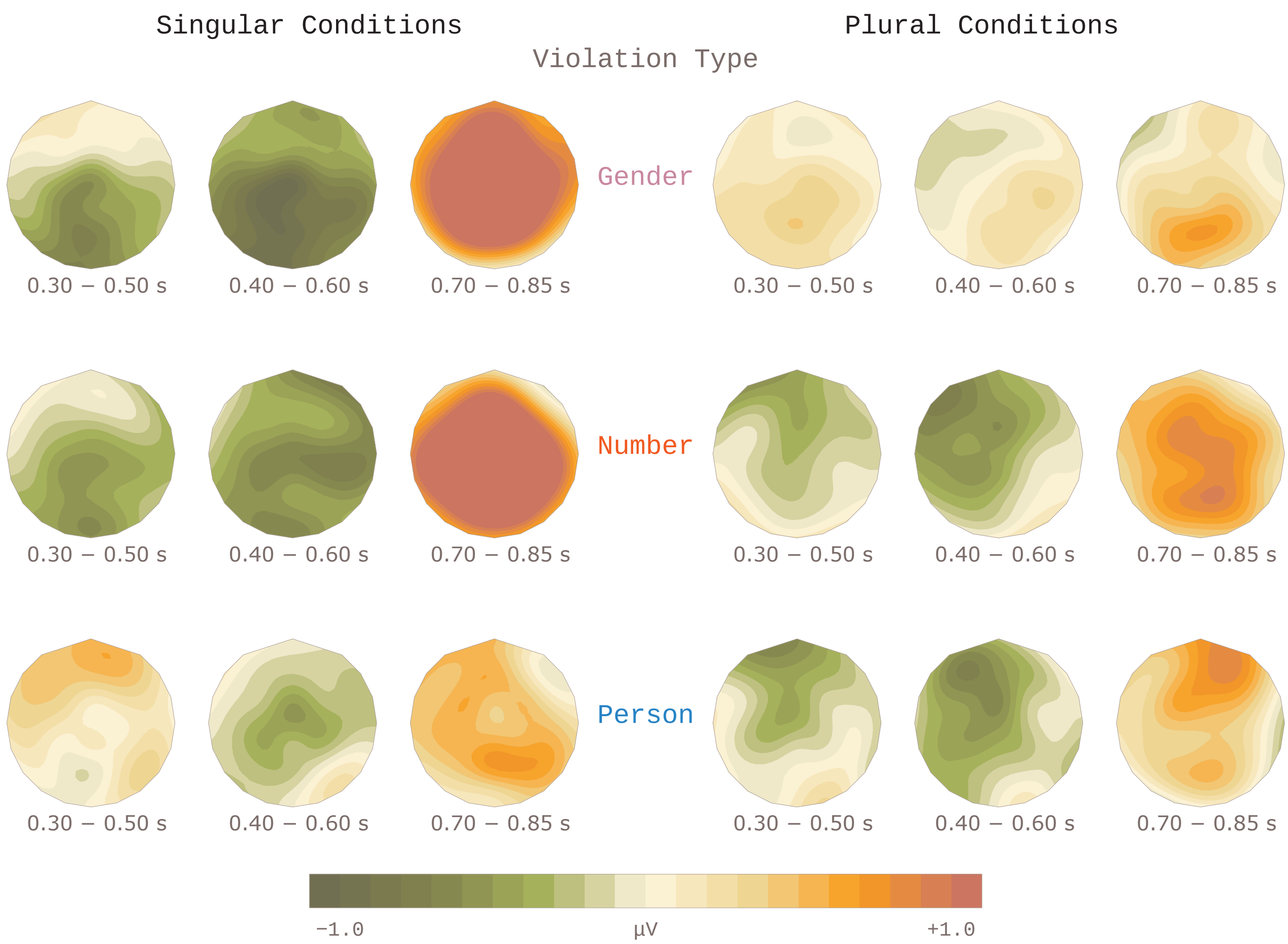

Table 4: ANOVA: ERPs at the Verb : 300 − 500 ms

Effects in time-window 300−500 ms				DF	Lateral Regions		DF	Midline Regions	
◆ ROI x ST							5,165	4.56	★
↳ROI = FCZ							1,33	4.47	★
◆ ST x CT				3,99	2.99	★	3,99	3.70	★★
↳ST = S	◆ CT			3,99	3.02	★			
	↳CT = PER+NUM	◆ CT		1,33	6.77	★			
◆ ROI x ST x CT							15,495	3.03	★★
↳ROI = FCZ	◆ ST x CT						3,99	2.62	☆
	↳ST = P	◆ CT					3,99	2.59	☆
		↳CT = PER+GEN	◆ CT				1,33	5.96	★
↳ROI = CZ	◆ ST x CT						3,99	4.79	★★
	↳ST = S	◆ CT					3,99	3.89	★★
		↳CT = ACP+NUM	◆ CT				1,33	7.47	★
		↳CT = PER+NUM	◆ CT				1,33	4.88	☆
		↳CT = ACP+GEN	◆ CT				1,33	5.86	★
		↳CT = PER+GEN	◆ CT				1,33	5.11	☆
	↳ST = P	◆ CT					3,99	3.11	★
		↳CT = PER+GEN	◆ CT				1,33	8.46	★
		↳CT = NUM+GEN	◆ CT				1,33	4.97	☆
↳ROI = CPZ	◆ ST x CT						3,99	4.27	★★
	↳ST = S	◆ CT					3,99	4.38	★★
		↳CT = ACP+NUM	◆ CT				1,33	6.59	★
		↳CT = PER+NUM	◆ CT				1,33	7.38	★
		↳CT = ACP+GEN	◆ CT				1,33	4.78	☆
		↳CT = PER+GEN	◆ CT				1,33	5.35	☆
	↳ST = P	◆ CT					3,99	2.43	☆
		↳CT = PER+GEN	◆ CT				1,33	5.85	★
		↳CT = NUM+GEN	◆ CT				1,33	5.41	☆
↳ROI = PZ	◆ ST x CT						3,99	4.39	★★
	↳ST = S	◆ CT					3,99	3.92	★★
		↳CT = ACP+NUM	◆ CT				1,33	5.11	☆
		↳CT = ACP+GEN	◆ CT				1,33	7.04	★
		↳CT = PER+GEN	◆ CT				1,33	5.23	☆

Continued on next page ...

…Table 4 continued

Effects in time-window 300−500 ms				DF	Lateral Regions	DF	Midline Regions	
└ROI = POZ	◆ ST x CT					3,99	3.72	⋆⋆
	└ST = S	◆ CT				3,99	5.89	⋆⋆⋆
		└CT = ACP+NUM	◆ CT			1,33	8.94	⋆⋆
		└CT = PER+NUM	◆ CT			1,33	6.09	⋆
		└CT = ACP+GEN	◆ CT			1,33	10.51	⋆⋆
		└CT = PER+GEN	◆ CT			1,33	7.52	⋆

3.2.2 Time-window 400−600 ms

As in the earlier time-window, the graded negativity effect engendered by the violation conditions as opposed to the acceptable conditions is significant in this time-window, with the differences in effects between the two subject-types becoming much more clear. Number violations engendered a large broadly distributed negativity effect regardless of subject-type. The modulation of negativities evoked by person and gender violations differed based on whether the subject-type was singular or plural, with gender violations evoking a larger effect than person violations when the subject was singular, and person violations evoking a larger effect than gender violations when the subject was plural. Similarly, the topographic distribution of these negativities also varied, with the effect for person violations showing a central distribution, and that for gender violations showing a centro-parietal distribution. Table 5 shows a summary of all the effects that reached at least marginal significance at the position of the verb in the 400 − 600 ms time-window.

There was a main effect of subject-type in the lateral and midline regions. The interaction ROI x subject-type was significant in the midline regions, which when resolved for the individual ROIs showed a simple effect of subject-type in the frontal, fronto-central and central midline regions. There was a main effect of condition-type in the lateral and midline regions. This effect was resolved in the lateral regions alone (due to a ROI interaction in the midline regions) by comparing the condition-types pairwise, which showed a simple effect of condition-type for the comparison ACP + NUM. The interaction ROI x condition-type was significant in the midline regions alone, which when resolved for the the individual ROIs showed an effect of condition-type in the fronto-central, central, centro-parietal and parieto-occipital midline regions. In these regions, the condition-types were compared pairwise, which showed a simple effect of condition-type for the comparison ACP + PER in the fronto-central and central midline regions. Comparing ACP + NUM revealed a simple effect of condition-type in all four midline regions concerned, the effect being marginal in the parieto-occipital midline region. Comparing ACP + GEN showed a simple effect of condition-type in the central and parieto-occipital midline regions, with the effect being marginal in the centro-parietal midline region.

The interaction subject-type x condition-type was significant in the lateral and midline regions. The three-way interaction ROI x subject-type x condition-type was likewise significant in the lateral and midline regions. Resolving the interaction for the individual ROIs revealed that, there was an effect of the interaction subject-type x condition-type in the left-posterior and right-posterior regions as well as in the central, centro-parietal, parietal and parieto-occipital midline regions. This effect was resolved for subject-type in each concerned region. In the left-posterior and right-posterior regions, there was an effect of condition-type when the subject was singular. Resolving this by comparing the condition-type pairwise revealed that there was a simple effect of condition-type for the comparisons ACP + NUM, ACP + GEN and PER + GEN in both left-posterior as well as right-posterior regions. This effect reached significance for the comparison PER + NUM in the right-posterior region alone. In the left-posterior region, there was an effect of condition-type when the subject was plural, which when resolved by comparing the condition-types pairwise showed a simple effect of condition-type for the comparison NUM + GEN and a marginal effect of condition-type for the comparison PER + GEN. In all the midline regions concerned, there was an effect of condition-type when the subject was singular. Resolving this by comparing the condition-type pairwise revealed that there was a simple effect of condition-type for the comparison ACP + NUM in the central and parieto-occipital midline regions, with this effect being marginal in the centro-parietal midline region. Comparing ACP + GEN revealed an effect of condition-type in all four midline regions concerned. There was a marginal effect of condition-type for the comparison PER + NUM in the parieto-occipital midline region. For the comparison PER + GEN, there was a marginal effect of condition-type in the centro-parietal midline region and a simple effect of condition-type in the parieto-occipital midline region. In the central and centro-parietal midline regions, there was an effect of condition-type when the subject was plural. Resolving this by comparing the condition-types pairwise revealed that, there was a simple effect of condition-type for the comparisons PER + GEN and NUM + GEN in both midline regions concerned.

Table 5: ANOVA: ERPs at the Verb : 400 − 600 ms

Effects in time-window 400−600 ms				DF	Lateral Regions		DF	Midline Regions	
◆ ST				1,33	5.55	★	1,33	6.47	★★
◆ ROI x ST							5,165	5.19	★★
└ROI = FZ	◆ ST						1,33	10.09	★★
└ROI = FCZ	◆ ST						1,33	12.42	★★★
└ROI = CZ	◆ ST						1,33	5.33	★
◆ CT				3,99	3.24	★	3,99	3.09	★
└CT = ACP+NUM	◆ CT			1,33	9.25	★★			
◆ ROI x CT							15,495	2.37	★
└ROI = FCZ	◆ CT						3,99	3.62	★★
	└CT = ACP+PER						1,33	9.19	★★
	└CT = ACP+NUM						1,33	6.33	★
└ROI = CZ	◆ CT						3,99	4.01	★★
	└CT = ACP+PER						1,33	7.66	★
	└CT = ACP+NUM						1,33	8.19	★
	└CT = ACP+GEN						1,33	8.17	★
└ROI = CPZ	◆ CT						3,99	2.94	★
	└CT = ACP+NUM						1,33	7.13	★
	└CT = ACP+GEN						1,33	5.08	☆
└ROI = POZ	◆ CT						3,99	2.75	★
	└CT = ACP+NUM						1,33	4.72	☆
	└CT = ACP+GEN						1,33	5.91	★
◆ ST x CT				3,99	3.03	★	3,99	3.68	★★
◆ ROI x ST x CT				9,297	2.37	★	15,495	2.13	★
└ROI = LP	◆ ST x CT			3,99	4.54	★★			
	└ST = S	◆ CT		3,99	6.31	★★★			
		└CT = ACP+NUM	◆ CT	1,33	5.76	★			
		└CT = ACP+GEN	◆ CT	1,33	11.95	★★			
		└CT = PER+GEN	◆ CT	1,33	9.70	★★			
	└ST = P	◆ CT		3,99	2.74	★			
		└CT = PER+GEN	◆ CT	1,33	5.38	☆			
		└CT = NUM+GEN	◆ CT	1,33	8.87	★★			
└ROI = CZ	◆ ST x CT						3,99	4.22	★★

Continued on next page ...

…Table 5 continued

Effects in time-window 400−600 ms			DF	Lateral Regions		DF	Midline Regions	
	└ST = S ◆ CT					3,99	5.16	★★
		└CT = ACP+NUM ◆ CT				1,33	7.67	★
		└CT = ACP+GEN ◆ CT				1,33	12.79	★★
	└ST = P ◆ CT					3,99	3.10	★
		└CT = PER+GEN ◆ CT				1,33	7.69	★
		└CT = NUM+GEN ◆ CT				1,33	6.05	★
└ROI = RP ¦¦ CPZ ◆ ST x CT			3,99	6.17	★★★	3,99	4.07	★★
	└ST = S ◆ CT		3,99	7.04	★★★	3,99	4.47	★★
		└CT = ACP+NUM ◆ CT	1,33	10.25	★★	1,33	4.48	☆
		└CT = PER+NUM ◆ CT	1,33	6.25	★			
		└CT = ACP+GEN ◆ CT	1,33	11.48	★★	1,33	9.97	★★
		└CT = PER+GEN ◆ CT	1,33	10.56	★★	1,33	5.38	☆
	└ST = P ◆ CT					3,99	2.64	★
		└CT = PER+GEN ◆ CT				1,33	5.78	★
		└CT = NUM+GEN ◆ CT				1,33	6.45	★
└ROI = PZ ◆ ST x CT						3,99	5.07	★★
	└ST = S ◆ CT					3,99	4.94	★★
		└CT = ACP+GEN ◆ CT				1,33	9.96	★★
		└CT = PER+GEN ◆ CT				1,33	7.18	★
		└CT = NUM+GEN ◆ CT				1,33	4.82	☆
└ROI = POZ ◆ ST x CT						3,99	5.13	★★
	└ST = S ◆ CT					3,99	7.07	★★★
		└CT = ACP+NUM ◆ CT				1,33	5.79	★
		└CT = PER+NUM ◆ CT				1,33	4.59	☆
		└CT = ACP+GEN ◆ CT				1,33	13.21	★★
		└CT = PER+GEN ◆ CT				1,33	11.94	★★

3.2.3 Time-window 500−700 ms

The ERPs in the time-window begin to diverge such that they are more positive-going for the violation conditions as opposed to the acceptable conditions. Whilst differences in the positivity effect can be observed for sentences involving singular subjects, no such clear pattern is apparent for sentences involving plural subjects. However, a general difference between the subject-types is beginning to emerge, especially in the frontal regions. Table 6 shows a summary of all the effects that reached at least marginal significance at the position of the verb in the 500 − 700 ms time-window.

There was a main effect of subject-type in the lateral and midline regions. The interaction ROI x subject-type was significant in the midline regions, which when resolved for the individual ROIs showed an effect of subject-type in the frontal and fronto-central regions. The interaction ROI x condition-type reached significance in the midline regions, but did not resolve further.

The interaction subject-type x condition-type was significant in the midline regions alone. The three-way interaction ROI x subject-type x condition-type was significant in the lateral as well as midline regions, which when resolved for the individual ROIs showed the interaction subject-type x condition-type to be marginal in the left-posterior region, significant in the right-posterior region, marginal in the fronto-central and central midline regions, and significant in the centro-parietal, parietal and parieto-occipital regions. Resolving the interaction in the lateral ROIs revealed an effect of condition-type in the right-posterior region alone when the subject was singular. This was resolved further by comparing the condition-types pairwise, which showed a marginal effect of condition-type for the comparisons ACP + GEN and NUM + GEN. Resolving the interaction in the midline regions showed an effect of condition-type in the fronto-central, central, centro-parietal and parietal midline regions when the subject was singular. This was further resolved by comparing the condition-types pairwise, revealing a simple effect of condition-type for the comparison ACP + PER in the fonto-central and central midline regions. Comparing PER + NUM showed an effect of condition-type in all four midline regions concerned. The comparison NUM + GEN revealed a simple effect of condition-type in the central, centro-parietal and parietal regions.

Table 6: ANOVA: ERPs at the Verb : 500 − 700 ms

Effects in time-window 500−700 ms	DF	Lateral Regions		DF	Midline Regions	
◆ ST	1,33	5.53	⋆	1,33	5.82	⋆
◆ ROI x ST				5,165	4.76	⋆
└ROI = FZ ◆ ST				1,33	14.86	⋆⋆⋆
└ROI = FCZ ◆ ST				1,33	10.40	⋆⋆
◆ ROI x CT				15,495	2.44	⋆
◆ ST x CT				3,99	2.68	⋆
◆ ROI x ST x CT	9,297	3.56	⋆	15,495	2.89	⋆⋆
└ROI = LP ¦ FCZ ◆ ST x CT	3,99	2.56	☆	3,99	2.40	☆
└ST = S ◆ CT				3,99	3.94	⋆⋆
└CT = ACP+PER ◆ CT				1,33	7.02	⋆
└CT = PER+NUM ◆ CT				1,33	11.26	⋆⋆
└ROI = CZ ◆ ST x CT				3,99	2.53	☆
└ST = S ◆ CT				3,99	3.70	⋆⋆
└CT = ACP+PER ◆ CT				1,33	6.37	⋆
└CT = PER+NUM ◆ CT				1,33	8.42	⋆
└CT = NUM+GEN ◆ CT				1,33	6.23	⋆
└ROI = RP ¦ CPZ ◆ ST x CT	3,99	3.27	⋆	3,99	3.06	⋆
└ST = S ◆ CT	3,99	2.43	☆	3,99	2.67	⋆
└CT = PER+NUM ◆ CT				1,33	5.69	⋆
└CT = ACP+GEN ◆ CT	1,33	4.65	☆			
└CT = NUM+GEN ◆ CT	1,33	4.71	☆	1,33	7.09	⋆
└ROI = PZ ◆ ST x CT				3,99	4.01	⋆⋆
└ST = S ◆ CT				3,99	3.18	⋆
└CT = PER+NUM ◆ CT				1,33	7.57	⋆
└CT = NUM+GEN ◆ CT				1,33	9.00	⋆⋆
└ROI = POZ ◆ ST x CT				3,99	2.90	⋆

3.2.4 Time-window 700−850 ms

The predominant effect in this time-window is the late-positivity effect for violation conditions as opposed to acceptable conditions. Crucially, this effect is not only graded based on the violating feature, but its modulation is also different based on the subject-type. The effect thus shows a four-way gradation when the subject is singular, with number violations evoking the largest positivity effect as opposed to acceptable sentences, followed by gender violations and then by person violations. By contrast, only a three-way difference ensues for plural subjects, with the largest positivity effect for number violations followed by smaller but virtually identical positivities for person violations and gender violations as opposed to acceptable sentences. Table 7 shows a summary of all the effects that reached at least marginal significance at the position of the verb in the 700 − 850 ms time-window.

There was a main effect of condition-type in the lateral and midline regions. This effect was resolved by comparing the condition-types pairwise, which showed a significant effect of condition-type for the comparisons ACP + PER, ACP + NUM, ACP + GEN and PER + NUM in all concerned regions, with the effect being marginal for the comparison NUM + GEN in the midline regions alone.

The interaction subject-type x condition-type was significant in the lateral as well as midline regions. The three-way interaction ROI x subject-type x condition-type was similarly significant in the lateral and midline regions. Resolving this interaction for the individual ROIs showed that the interaction subject-type x condition-type was significant in the left-anterior, left-posterior and right-anterior regions, and significant in all the midline regions except in the parieto-occipital midline region, in which it was marginal. This interaction was resolved for subject-type in all regions concerned. In the lateral regions, when the subject was singular, there was an effect of condition-type in the left-anterior, left-posterior and right-anterior regions, which was resolved by comparing condition-types pairwise in these regions. This showed a simple effect of condition-type for the comparisons ACP + NUM, ACP + GEN and PER + GEN in all three regions concerned. There was an effect of condition-type for the comparison PER + NUM in the left-posterior region, and the effect was marginal in the right-anterior region. In the lateral regions, when the subject was plural, there was an effect of condition-type only in the left-anterior region, which when resolved further by comparing the condition-types pairwise revealed a simple effect of condition-type for the comparisons ACP + NUM and NUM + GEN. In the midline regions, when the subject was singular, there was an effect of condition-type in all regions, which was resolved by comparing condition-types pairwise, showing the following. There was a simple effect of condition-type for the comparisons ACP + NUM, ACP + GEN, PER + NUM and PER + GEN in all the midline regions, with the effect being marginal in the parieto-occipital midline region for the comparison PER + GEN. For the comparison ACP + PER, there was an effect of condition-type in the parietal midline region alone. Comparing NUM + GEN showed a simple effect of condition-type in the parietal midline region, and this effect was marginal in the parieto-occipital midline region. In the midline regions, when the subject was plural, there was an effect of condition-type in all regions except the centro-parietal midline region, in which it was marginal. Resolving the effect by comparing condition-types pairwise showed a simple effect of condition-type for the comparison ACP + NUM in all regions. Comparing NUM + GEN revealed a simple effect of condition-type in the frontal and fronto-central midline regions, and this effect was marginal in the central midline region. For the comparison ACP + PER, there was an effect of condition-type in the frontal midline region. Comparing ACP + GEN showed a simple effect of condition-type in the parietal and parieto-occipital midline regions.

Table 7: ANOVA: ERPs at the Verb : 700 − 850 ms

Effects in time-window 700–850 ms				DF	Lateral Regions		DF	Midline Regions	
◆ CT				3,99	10.82	★★★	3,99	13.64	★★★
└CT = ACP+PER	◆ CT			1,33	6.44	★	1,33	5.69	★
└CT = ACP+NUM	◆ CT			1,33	30.66	★★★	1,33	30.63	★★★
└CT = PER+NUM	◆ CT			1,33	6.67	★	1,33	11.14	★★
└CT = ACP+GEN	◆ CT			1,33	15.24	★★★	1,33	22.45	★★★
└CT = NUM+GEN	◆ CT						1,33	4.41	☆
◆ ST x CT				3,99	3.19	★	3,99	4.85	★★
◆ ROI x ST x CT				9,297	2.66	★	15,495	3.40	★★
└ROI = LA ¦ FZ	◆ ST x CT			3,99	3.51	★	3,99	5.98	★★★
	└ST = S	◆ CT		3,99	6.40	★★★	3,99	10.75	★★★
		└CT = ACP+NUM	◆ CT	1,33	12.74	★★	1,33	18.18	★★★
		└CT = PER+NUM	◆ CT				1,33	7.06	★
		└CT = ACP+GEN	◆ CT	1,33	12.72	★★	1,33	22.46	★★★
		└CT = PER+GEN	◆ CT	1,33	5.53	★	1,33	10.61	★★
	└ST = P	◆ CT		3,99	3.74	★★	3,99	3.40	★
		└CT = ACP+PER	◆ CT				1,33	5.82	★
		└CT = ACP+NUM	◆ CT	1,33	9.48	★★	1,33	6.57	★
		└CT = NUM+GEN	◆ CT	1,33	8.27	★	1,33	5.51	★
└ROI = LP ¦ FCZ	◆ ST x CT			3,99	2.96	★	3,99	5.55	★★★
	└ST = S	◆ CT		3,99	10.89	★★★	3,99	11.11	★★★
		└CT = ACP+NUM	◆ CT	1,33	18.51	★★★	1,33	18.34	★★★
		└CT = PER+NUM	◆ CT	1,33	11.08	★	1,33	10.10	★★
		└CT = ACP+GEN	◆ CT	1,33	14.93	★★★	1,33	22.08	★★★
		└CT = PER+GEN	◆ CT	1,33	8.18	★	1,33	13.40	★★
	└ST = P	◆ CT					3,99	3.16	★
		└CT = ACP+NUM	◆ CT				1,33	7.82	★
		└CT = NUM+GEN	◆ CT				1,33	6.07	★
└ROI = RA ¦ CZ	◆ ST x CT			3,99	3.64	★	3,99	3.84	★★
	└ST = S	◆ CT		3,99	6.55	★★★	3,99	11.06	★★★
		└CT = ACP+NUM	◆ CT	1,33	7.00	★	1,33	17.33	★★★
		└CT = PER+NUM	◆ CT	1,33	4.48	☆	1,33	12.27	★★
		└CT = ACP+GEN	◆ CT	1,33	14.01	★★	1,33	15.35	★★★

Continued on next page ...

…Table 7 continued

Effects in time-window 700–850 ms				DF	Lateral Regions		DF	Midline Regions	
		↳CT = PER+GEN	◆ CT	1,33	9.21	★★	1,33	10.35	★★
	↳ST = P	◆ CT					3,99	3.18	★
		↳CT = ACP+NUM	◆ CT				1,33	10.74	★★
		↳CT = NUM+GEN	◆ CT				1,33	5.43	☆
↳ROI = CPZ	◆ ST x CT						3,99	4.60	★★
	↳ST = S	◆ CT					3,99	13.72	★★★
		↳CT = ACP+NUM	◆ CT				1,33	21.72	★★★
		↳CT = PER+NUM	◆ CT				1,33	13.42	★★
		↳CT = ACP+GEN	◆ CT				1,33	19.75	★★★
		↳CT = PER+GEN	◆ CT				1,33	10.32	★★
	↳ST = P	◆ CT					3,99	2.52	☆
		↳CT = ACP+NUM	◆ CT				1,33	9.09	★★
↳ROI = PZ	◆ ST x CT						3,99	4.37	★★
	↳ST = S	◆ CT					3,99	17.01	★★★
		↳CT = ACP+PER	◆ CT				1,33	5.49	★
		↳CT = ACP+NUM	◆ CT				1,33	21.12	★★★
		↳CT = PER+NUM	◆ CT				1,33	18.88	★★★
		↳CT = ACP+GEN	◆ CT				1,33	20.51	★★★
		↳CT = PER+GEN	◆ CT				1,33	9.79	★★
		↳CT = NUM+GEN	◆ CT				1,33	5.94	★
	↳ST = P	◆ CT					3,99	3.58	★
		↳CT = ACP+NUM	◆ CT				1,33	9.79	★★
		↳CT = ACP+GEN	◆ CT				1,33	9.01	★★
↳ROI = POZ	◆ ST x CT						3,99	2.42	☆
	↳ST = S	◆ CT					3,99	9.76	★★★
		↳CT = ACP+NUM	◆ CT				1,33	15.80	★★★
		↳CT = PER+NUM	◆ CT				1,33	13.36	★★
		↳CT = ACP+GEN	◆ CT				1,33	11.66	★★
		↳CT = PER+GEN	◆ CT				1,33	4.99	☆
		↳CT = NUM+GEN	◆ CT				1,33	4.71	☆
	↳ST = P	◆ CT					3,99	2.99	★
		↳CT = ACP+NUM	◆ CT				1,33	9.22	★★
		↳CT = ACP+GEN	◆ CT				1,33	9.90	★★

3.3 Summary of Results

The ERP results from our study can be summarised as follows.

- Violation of either of the agreement features generally elicited a biphasic negativity – late-positivity effect as opposed to acceptable sentences.
- However, this biphasic effect was graded based on the feature that was violated as well as the subject-type. Whilst number violations evoked the largest effects overall, the effects for person and gender violations were smaller in comparison, and showed a qualitative difference based on whether the subject-type was singular or plural.
- In the early negativity time-window, number and gender violations evoked a negativity effect with a posterior distribution as opposed to person violations and acceptable conditions, when the subject was singular. When the subject was plural, number and person violations evoked a negativity effect as opposed gender violations and acceptable conditions. This effect with an anterior distribution showed a slight left-lateral trend visually, which was however not significant statistically.
- In the negativity time-window, a three-way distinction ensued that was different for the two subject-types. Number violations evoked a large broadly distributed negativity effect for both subject-types. Gender violations evoked a larger negativity effect as opposed to person violations when the subject was singular, with a centro-parietal maximum. Person violations evoked a larger negativity effect as opposed to gender violations as well as acceptable conditions when the subject was plural, with a fronto-central distribution.
- In the late-positivity time-window, there was a four-way gradation of the positivity effect when the subject was singular, whereas it was a three-way gradation when the subject was plural. Number violations evoked the largest positivity effect in both subject-types. Gender violations

evoked a larger positivity effect compared to person violations when the subject was singular, but no such difference ensued when the subject was plural.

Based on the latency and topography of effects as well as the experimental conditions, the negativity effects in our study can be plausibly interpreted as instances of an N400 effect and the late-positivities as instances of a P600 effect. Agreement violations of the sort we employed have been traditionally viewed as formal morphosyntactic rule violations, which are associated with concomitant LAN effects that are interpreted as indicative of the detection of the morphosyntactic violation (Münte, Szentkuti, Wieringa, Matzke, & Johannes, 1997; Münte, Matzke, & Johannes, 1997; Friederici, 2002; Bornkessel & Schlesewsky, 2006). Further, the likelihood of observing a LAN has been thought to be in direct proportion to the morphological richness of a language (Friederici & Weissenborn, 2007), which nevertheless is said to depend upon whether morphological marking is crucial for assigning syntactic roles in the given language (Friederici, 2011). However, Molinaro, Barber, & Carreiras (2011) draw attention to the fact that the nature of the complexities involved in morphological decomposition for feature identification may dissociate whether a LAN or, alternatively, an N400 ensues. Indeed, Choudhary, Schlesewsky, Roehm, & Bornkessel-Schlesewsky (2009) provide converging evidence from Hindi that this dissociation results from whether an interpretively relevant cue is violated (in which case an N400 ensues), or alternatively the violation involves a cue that is irrelevant for interpretation. In view of this, the absence of a LAN effect and the presence of an N400 in our study is not surprising, given that agreement computation in Arabic is not simply formal, but is dependent upon specific syntactic properties of the construction involved, such as word-order and whether or not the subject is overt, as well as properties at the syntax-semantic interface such as humanness (see below for a detailed discussion). The P600 effects in our study can be plausibly interpreted as reflecting repair or reanalysis processes associated with agreement violations (Friederici, 2002, 2011; Bornkessel

& Schlesewsky, 2006), thought to be triggered by domain-general conflict monitoring processes (Van de Meerendonk, Chwilla, & Kolk, 2013).

Before discussing our findings further below, we briefly turn to an ongoing debate in the ERP literature on morphosyntactic processing, which relates to inter-participant variation of ERP effects. The claim is that most of the LAN effects reported for agreement processing in the literature may simply be artefactual resulting from components that partially overlap temporally and spatially, which happen to be aggregated together due to the grand-averaging of individual ERPs (Tanner, 2018, 2015; Tanner & Van Hell, 2014). Whilst we did not find a LAN effect in our study, it is important to consider whether individual variation amongst our participants might be able to explain some of our results. However, several points speak against such a possibility. Firstly, individual variation, if any, would have been systematic across conditions for a given participant. Secondly, in view of our within-participants design, any such variation would have equally contributed to all the critical conditions since all of them consisted of the same type of violation, namely a violation of an agreement feature between the subject and the verb. Thirdly, as Grey, Tanner, & Van Hell (2017) have reported, variation in ERP effects engendered during morphosyntactic processing across individuals is said to be minimal in case of dominant right-handers (as opposed to left-handers). All our participants were dominantly right-handed individuals (see Section 2.5). Crucially therefore, given that any potential artefactual outcome resulting from the grand-averaging procedure would be equally applicable to all our critical conditions, individual variation may not fully account for, nor nullify, the subject-type specific differences in the modulation of effects we found in our study.

4 Discussion

We investigated the processing of subject-verb agreement in simple intransitive Arabic sentences in this ERP experiment, and present here the findings from one of the first neurocognitive studies examining sentence level processes in

Arabic, thus providing first insights into the online processing of the language.

4.1 Hierarchical Modulation of Effects

As illustrated by the topographic map of effects in Figure 3, our results show that violations of the features not only modulate the biphasic pattern of ERPs, but they also show a qualitative difference in their modulation based on whether the subject was singular or plural. The modulation of effects for the violation conditions as opposed to the acceptable conditions in our study can be represented as a hierarchy as follows:

(2) Modulation of N400
Singular subject conditions : Number / Gender > Person
Plural subject conditions : Number / Person > Gender

(3) Modulation of P600
Singular subject conditions : Number > Gender > Person
Plural subject conditions : Number > Person / Gender

A handful of previous studies investigating processing differences between agreement features (in the subject-verb context or otherwise) have indeed reported a modulation of late-positive P600 effects, albeit for across subject-type comparisons (Kaan et al., 2000; Deutsch & Bentin, 2001) or for combined versus single feature violations (Nevins et al., 2007; Zawiswewski et al., 2016; Alemán Bañón & Rothman, 2016).[1] However, to our knowledge, a modulation

1. Since the focus in our study is on the differences in effects *within* a subject-type, it is not straightforward to compare our results with these studies. Nevertheless, restricting to subject-verb agreement and single feature violations, the P600 for the two subject-types in our results may be compared to at least two studies. The P600 difference in our results for the number feature between the two subject-types would be in line with a similar finding in Dutch that Kaan et al. (2000) reported. On the other hand, we found P600 effects for gender violations in both subject-types, whereas Deutsch & Bentin (2001) reported a P600 effect only when the subject was plural in Hebrew.

of negativities (LAN or N400) has not been reported to date in the context of agreement violations[2], and ours may be the first study on agreement processing to have observed such a hierarchical modulation of the N400 based on the agreement feature involved.

The modulation of effects in our study, especially of the negativity, cannot be satisfactorily explained simply based on differences in the complexity between the conditions, nor the frequency of a certain type of violation, nor orthographic salience, nor even the morphosyntactic markedness per se. Rather, we argue that these results suggest that the processing system takes into account the relative cognitive salience of the features in the given language. If so, this would be in line with the view put forward by Molinaro, Barber, & Carreiras (2011) in their exhaustive review of cross-linguistic findings on processing agreement, namely that the processing of agreement dependencies is sensitive to the feature involved as well as the constituents expressing the dependency, and that this can result in modulations of negativity effects such as LAN and/or N400 as well as of the late-positivity[3]. Specifically, the processing of subject-verb agreement is said to involve semantic factors in addition to morphosyntactic information (Molinaro, Vespignani, Zamparelli, & Job, 2011), and the extent to which lexical-semantic factors play a role in computing subject-verb agreement in a given language is said to have a direct bearing on whether a LAN or an N400 effect is engendered (Molinaro, Barber, Caffarra, & Carreiras, 2015).

2. In their Basque study, Zawiswewski et al. (2016) found a significant difference in the 300-500 time-window, which they had initially interpreted as an N400 modulation for number violations as opposed to the other conditions; however, the statistics in their study as well as the topography of this effect with a posterior maximum in the P3 electrode region suggests that it is rather a P300 difference that they observe, as has been correctly pointed out by an anonymous reviewer of their manuscript. Their behavioural data also support this interpretation rather than their initial conjecture.
3. Nevertheless, we do not assume that these effects are strictly following each other in stages.

4.2 Salience-weighted Feature Hierarchy

In view of the agreement feature hierarchy in (1) originally proposed based on cross-linguistic distribution of these features (Greenberg, 1963), the graded effects in the present study are not surprising. Rather, our results provide neurophysiological evidence for the idea that the relative position of a feature in the hierarchy directly reflects its relative cognitive salience (Harley & Ritter, 2002, but see below) and thereby its relative importance in language processing (Carminati, 2005) in comparison to other features. Nevertheless, the modulation of effects found in our study suggests that the feature hierarchy may not be identical across languages. That is, it appears to differ based on language-specific properties, and may exhibit slight variations even within a given language depending upon the specifics of the construction involved and general language-use in the speech community. We argue below that this variation is neither an aberration nor arbitrary. Rather, it is systematic based on the relative weightings of the salience of a given feature in a language.

We postulate a hierarchy of agreement features that is language-specific rather than universal. The properties of a given language as well as the relative salience of the agreement features determine the hierarchy and thereby differentially modulate language comprehension. If the hierarchy is flat such that there is minimal difference in cognitive salience between the features, then no modulation of ERP effects should be expected for single feature violations in the given language. If there is a hierarchical difference in salience between the features, then there would be a concomitant modulation of ERP effects, such that violating a highly salient agreement feature evokes a larger effect compared to violations of a less salient feature.

The idea that some properties in a language may be relatively more salient than others for language comprehension is not new. This was first put forward in the context of the Competition Model (e.g., MacWhinney et al., 1984; Bates & MacWhinney, 1989; Bates, 1999), in which such linguistic properties or cues with language-specific weightings interact during language comprehension. Agreement is one such cue, the strength of which has been shown to vary across languages (for an overview, see Bates et al., 2001). Within the domain of online language processing, Bornkessel-Schlesewsky & Schlesewsky (2009) have proposed that information types such as case-marking, word-order, animacy etc., called prominence information, all interact during online language comprehension, and that their relative prominence depends upon their language-specific weightings. For instance, word-order has higher weighting in English than, say, animacy, and determines the argument roles, whereas in a language like Turkish, case is more prominent, and so on. Furthermore, in a recent article, Bornkessel-Schlesewsky & Schlesewsky (2019) have argued for a neurobiologically plausible model, whereby they posit that all the language-related negativities form a family of functionally related rather than distinct negativities, and that amplitude modulations of negativities are said to reflect 'precision-weighted predictive coding errors, with precision (the inverse of variance) reflecting the relevance of a particular stimulus feature in a given language (Bornkessel-Schlesewsky & Schlesewsky, 2019, p. 11)'.

We extend the idea of weightings to the domain of agreement features, such that they are organised on a salience-weighted hierarchy, with differences in their salience due to specific properties of the language determining their position on the hierarchy. Further, this hierarchy interacts with other prominence scales and discourse context such that the modulation of effects may vary within the language based on the specific type of construction (Mancini et al., 2011a). This proposal would parsimoniously account for the fact that ERP effects for subject-verb agreement violations show a graded hierarchy in certain languages / constructions, whereas such a pattern does not ensue in others.

4.3 Language-specific Feature Weightings

In support of our proposal that the hierarchy of agreement features may very well be language-specific based on the relative cognitive salience of the features in the given language, we motivate here a language-internal explanation of how these different weightings come about in Arabic.

As mentioned when motivating Arabic as a suitable language for the purpose of investigating our research question, verb agreement is not merely formal in Arabic. Specifically, number and gender agreement in particular involve complex interactions with syntactic and semantic properties of a sentence and its event participants, such that verb agreement with these features is not always determined by the number and gender properties of the subject noun concerned.

The Arabic verb agrees with its subject noun in all three features when the word-order is SV, whereas it shows an idiosyncratic agreement pattern for third-person plural subject nouns in the VS word-order depending upon whether the subject is overt, or alternatively, covert (i.e., dropped) in a sentence. An upcoming overt plural subject would require the verb to show singular agreement, whereas if the plural subject were to be dropped from the utterance, the verb must show plural agreement. That is, number agreement is not only dependent upon the subject's number property, but also upon the relative order of the verb and the subject as well as whether or not the subject is going to be overtly uttered. In other words, under certain circumstances, an identical verb form is sometimes correct and sometimes incorrect based on the above mentioned factors that go beyond the number property of the subject noun per se. Thus, as far as the processing system is concerned, the number feature provides highly salient information in order to process a sentence and construct its intended meaning, because a simple feature-matching between the number property of the subject and the verb would not always lead to the correct interpretation.

Similarly, the gender feature also involves an idiosyncracy, whereby plural masculine nouns require the verb to show full agreement if the noun is human, whereas for non-human (animate and inanimate) masculine plural nouns, the verb must show singular-feminine agreement. In effect, this means that gender interacts with the animacy hierarchy (or rather, humanness to be specific) and plays a crucial role in determining the correct agreement paradigm.

By contrast, the person feature does not exhibit any idiosyncratic variation in Arabic as far as agreement is concerned. In other words, an agreement violation involving the person feature is always reliably a violation, and therefore, the processing system need not consider global properties of the language when evaluating person agreement in Arabic.

Therefore, evaluating number and gender agreement involves factors beyond the number and gender properties of the noun, which co-determine the processing of these features.

In other words, number and gender, in that order, are cognitively more salient in Arabic for processing subject-verb agreement than the person feature. This would explain the hierarchy of effects we observed in our study for singular subjects. However, results for plural subject nouns would suggest that the hierarchy is qualitatively different from that for singular subjects, in that the gender feature may be less salient than the person feature for plural subjects. There may be at least two language-internal reasons for this difference. First, when speaking about or adressing a group of individuals involving both women and men, speakers of Arabic overwhelmingly tend to use masculine agreement. That is, based on language-use, a verb that shows masculine agreement for a group of individuals also involving women is not a violation in Arabic. Second, the idiosyncracy described above, namely that non-human masculine plural nouns require the verb (and adjectives etc.) to show singular feminine agreement across the board, means that roughly half of the utterances involving masculine plural nouns require feminine agreement. That is, a verb that shows feminine agreement in the context of masculine plural subjects (albeit non-human) is not always a violation either. Simply put, both mismatching paradigms, namely, masculine subjects requiring feminine verb agreement and feminine subjects requiring masculine verb agreement, exist in the language. As a consequence, gender agreement mismatches involving plural nouns (masculine or feminine) are not always violations, which explains the fact that the effects for gender violations in our study were smallest for plural subjects.

4.4 Implications and Outlook

Previous results from morphologically rich languages such as Hindi and Basque are of particular relevance here. In Hindi, in which subject-verb agreement is purely based on the features of the nominative subject noun, Nevins et al. (2007) have found that no modulation of effects ensued based on whether number or gender is violated. Although their study did not include a pure person violation, there was a combined person + number violation condition, which elicited a larger P600 effect. Furthermore, the P600 effect for the combined violation of person + number was larger than that for both the combined number + gender violation and individual number and gender violations. A potential caveat for the larger P600 in their study for person + number violations may have been the relative infrequency of this condition in comparsion to the other violation conditions in their stimuli. Remaining ambivalent about whether or not the infrequency of this condition played a role, the authors attributed the larger P600 effect to the increased orthographic salience of person violations in Hindi rather than an additive effect, but added that it may be due to a cross-linguistically privileged status of the person feature. In Basque, Zawiswewski et al. (2016) have reported a larger P600 effect for person violations as opposed to number violations. By contrast, the effect for a combined violation of person + number did not differ from that for a person violation, with these two conditions showing an equally larger P600 effect compared to number violation. This would imply that the person feature is more salient in Basque, and that a combined violation of a more salient and less salient feature would not lead to additive effects. In the first instance, these results may appear to suggest that the person feature is somehow more significant and especially salient than the other features. However, two studies that specifically investigated person agreement are a case in point here. Seemingly anomalous, but alternatively interpretable person agreement (the so-called 'unagreement') in Spanish (Mancini, Molinaro, Rizzi, & Carreiras, 2011b) and Basque (Mancini, Massol, Duñabeitia, Carreiras, & Molinaro, 2019) evoke negativities (albeit with topographic and latency differences) but no P600 in comparison to outright and irresolvable person violations. Taken together, these results and the results from our study do not speak for a universal special status for the person feature, but rather provide converging evidence to our claim that modulation of effects correlate with language-specific differences in the salience of a given feature.

Consequently, we argue that a parsimonious account of apparently contradictory results from Hindi and Basque mentioned above on the one hand and our results from Arabic on the other would be along the lines of what we propose here, namely that the agreement features person, number and gender may show different hierarchies (or no hierarchies at all) depending upon the specific properties of a language, based on their relative salience in the given language, with differences in language-specific weightings resulting from a number of language-internal reasons.

Our proposal here is not entirely incompatible with the idea of orthographic salience contributing to the larger P600 effect for the combined person + number violation in Hindi. However, a more general explanation for the increased salience of the person feature in Hindi may be that, whereas number and gender are marked on the main verb even in the perfective aspect when the agent is in ergative case, person marking is underspecified by contrast, and only available from the auxiliary in the perfective aspect. In the imperfective aspect, when the person feature is exclusively available on the main verb as in the stimuli employed by Nevins et al. (2007), it is more prominent and salient, and a violation of this highly salient feature would have led to a concomitantly larger P600 effect in their study. Similarly, in the Basque study by Zawiswewski et al. (2016), the word-final subject morpheme in the auxiliary verb for the number violation was orthographically minimally different from that for the grammatical condition (*-*zue* versus -*zu*), whereas the morpheme indicating person and person + number violations were quite distinct (*-*t* and *-*gu* versus -*zu*), and therefore much more salient visually and phonetically. This may have contributed to the equally larger P600 for these latter violations as opposed to number violations in that study. The behavioural data that the authors have reported strongly supports this claim, such that the violations involving the distinct morphemes were almost always accurately detected as such, and much quicker, in comparison to number violations, which were detected with a lower accuracy, and at a slower pace.

In sum, existing results from Hindi and Basque can be accounted for based on the increased salience of the person feature (in the Hindi study, due to the orthography and exclusive availability of person in their stimuli; in the Basque study, due to the orthographic distinctness of the morpheme indicating person violation), without having to resort to a universally privileged or significant status for the person feature. In other words, the person feature is more salient in Hindi and Basque, albeit due to very different reasons, leading to a similar pattern of results in these languages. By contrast, the person feature seems to be the least salient feature in Arabic in comparison to number or gender, as explained earlier. Our proposal here of a language-specific salience-weighted feature hierarchy also satisfactorily provides a general and overarching account for cross-linguistic differences in the pattern of results reported for subject-verb agreement violations, such as in Dutch (Kaan et al., 2000) and Hebrew (Deutsch & Bentin, 2001), among others.

Furthermore, the concept of salience-based weightings enables generating specific hypotheses about agreement processing in understudied languages. For instance, whilst we do not yet have any independent evidence from Arabic for whether or not multiple feature violations would lead to additive effects when the features are on a hierarchy based on their salience differences, results from Hindi and Basque point to the fact that an additive effect need not necessarily ensue. Nevertheless, further research is necessary to investigate the question in detail. Furthermore, if agreement is purely formal in a language, and there are no apparent salience differences between the features, as in say a language like Malayalam, then no modulation of effects should be expected, whereas in languages in which the features have different salience weightings either due to orthographic reasons (as in Hindi or Basque), or syntactic or semantic reasons (as in Arabic), then modulations of effects should be expected.

Our proposal for a Salience-weighted Feature Hierarchy is generally in line with the agreement feature hierarchy (Greenberg, 1963) and the Feature Strength Hypothesis (Carminati, 2005), and provides neurophysiological evidence for the idea that agreement processes involve syntactic, semantic and pragmatic factors all at once (Eberhard et al., 2005; Vigliocco et al., 1995). However, our findings do not support a universal special status for specific features (say, person) across languages. Rather, the crucial difference between what we postulate here and previous proposals is that, the hierarchy of features may and do differ based on language-specific characteristics, whereby the relative position of a feature depends upon the language-speicific weighting of the feature. Our proposal parsiomoniously accounts for the fact that ERP effects for agreement violations show a graded hierarchy in certain languages / constructions, whereas such a pattern does not ensue in others.

4.5 Conclusion

In the study presented here, we investigated the processing differences between person, number and gender agreement in Arabic, a widely-spoken but understudied language. One of the first neurophysiological studies examining sentence level processes in Arabic to date, the findings from our study are an important addition to existing cross-linguistic results on online language comprehension.

Thanks to the within-language within-participants systematic comparison of all three agreement features in the context of subject-verb agreement in our study, we could show for the first time that the agreement features differentially modulate language comprehension based on their relative cognitive salience. The salience weightings of features may and do differ across languages depending upon the specific properties of the language concerned, thus giving rise to a hierarchy of agreement features that is language-specific rather than universal. Such a Salience-weighted Feature Hierarchy, we argue, would parsimoniously account for the diversity of existing cross-linguistic neurophysiological results on verb agreement processing, without having to resort to a universal special status for a certain feature, nor having to assume a universally static hierarchy of features that does not take into account the typological properties of individual languages. Furthermore, our proposal enables generating specific testable hypotheses about agreement processing in languages that are as yet unstudied or understudied. In this respect, our study is an important contribution towards understanding how the human brain processes and comprehends vastly diverse languages with equal ease and poise.

Acknowledgements

The research reported here conducted in the context of several research visits by R. Muralikrishnan was partly supported by a Research Institute grant (G1001) from New York University in Abu Dhabi to Prof.Alec Marantz, whom we should like to thank here. Ali Idrissi was at the United Arab Emirates University in Al Ain during the study.

References

Acuña-Fariña, J. C. (2009). The linguistics and psycholinguistics of agreement: A tutorial overview. *Lingua*, *119*(3), 389–424. Retrieved from https://doi.org/10.1016/j.lingua.2008.09.005 doi: 10.1016/j.lingua.2008.09.005 [Cited at p. 4]

Alemán Bañón, J., & Rothman, J. (2016). The role of morphological markedness in the processing of number and gender agreement in spanish: an event-related potential investigation. *Language, Cognition and Neuroscience*, *31*(10), 1273–1298. Retrieved from https://doi.org/10.1080/23273798.2016.1218032 doi: 10.1080/23273798.2016.1218032 [Cited at p. 6, 7, 41]

Angrilli, A., Penolazzi, B., Vespignani, F., Vincenzi, M. D., Job, R., Ciccarelli, L., … Stegagno, L. (2002). Cortical brain responses to semantic incongruity and syntactic violation in italian language: an event-related potential study. *Neuroscience Letters*, *322*(1), 5–8. Retrieved from https://doi.org/10.1016/s0304-3940(01)02528-9 doi: 10.1016/s0304-3940(01)02528-9 [Cited at p. 6]

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137–153. [Cited at p. 6]

Bates, E. (1999). Processing complex sentences: A cross-linguistic study. *Language and Cognitive Processes*, *14*(1), 69–123. Retrieved from https://doi.org/10.1080/016909699386383 doi: 10.1080/016909699386383 [Cited at p. 5, 43]

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–76). New York: Cambridge University Press. [Cited at p. 5, 43]

Bates, E., McNew, S., Devescovi, A., & Wulfeck, B. (2001). Psycholinguistics: A cross-language perspective. *Annual Review of Psychology*, *52*, 369–396. [Cited at p. 5, 44]

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45–93. Retrieved from https://doi.org/10.1016/0010-0285(91)90003-7 doi: 10.1016/0010-0285(91)90003-7 [Cited at p. 4]

Bornkessel, I., & Schlesewsky, M. (2006). The Extended Argument Dependency Model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, *113*(4), 787–821. [Cited at p. 39]

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The Role of Prominence Information in the Real-Time Comprehension of Transitive Constructions: A Cross-Linguistic Approach. *Language and Linguistics Compass*, *3*(1), 19–58. [Cited at p. 44]

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2016). The importance of linguistic typology for the neurobiology of language. *Linguistic Typology*, *20*(3). Retrieved from https://doi.org/10.1515/lingty-2016-0032 doi: 10.1515/lingty-2016-0032 [Cited at p. 9]

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative Event-Related Potentials. *Frontiers in Psychology*, *10*. Retrieved from https://doi.org/10.3389/fpsyg.2019.00298 doi: 10.3389/fpsyg.2019.00298 [Cited at p. 44]

Carminati, M. N. (2005). Processing reflexes of the feature hierarchy (person > number > gender) and implications for linguistic theory. *Lingua*, *115*(3), 259–285. Retrieved from https://doi.org/10.1016/j.lingua.2003.10.006 doi: 10.1016/j.lingua.2003.10.006 [Cited at p. 4, 43, 50]

Choudhary, K. K., Schlesewsky, M., Roehm, D., & Bornkessel-Schlesewsky, I. (2009). The n400 as a correlate of interpretively relevant linguistic rules: Evidence from hindi. *Neuropsychologia*, *47*(13), 3012–3022. Retrieved from https://doi.org/10.1016/j.neuropsychologia.2009.05.009 doi: 10.1016/j.neuropsychologia.2009.05.009 [Cited at p. 39]

Corbett, G. G. (2000a). *Number*. Cambridge: Cambridge University Press. [Cited at p. 4]

Corbett, G. G. (2000b). *Number*. Cambridge: Cambridge University Press. [Cited at p. 4]

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes*, *13*(1), 21–58. [Cited at p. 5]

De Vincenzi, M. (1999). [Cited at p. 4] *Journal of Psycholinguistic Research*, *28*(5), 537–553. Retrieved from https://doi.org/10.1023/a:1023272511427 doi: 10.1023/a:1023272511427

De Vincenzi, M., Job, R., Matteo, R. D., Angrilli, A., Penolazzi, B., Ciccarelli, L., & Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language*, *85*(2), 280–296.

Retrieved from https://doi.org/10.1016/s0093-934x(03)00055-5 doi: 10.1016/s0093-934x(03)00055-5 [Cited at p. 6]

Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*, *45*(2), 200–224. Retrieved from https://doi.org/10.1006/jmla.2000.2768 doi: 10.1006/jmla.2000.2768 [Cited at p. 6, 41, 49]

Díaz, B., Sebastián-Gallés, N., Erdocia, K., Mueller, J. L., & Laka, I. (2011). On the cross-linguistic validity of electrophysiological correlates of morphosyntactic processing: A study of case and agreement violations in basque. *Journal of Neurolinguistics*, *24*(3), 357–373. Retrieved from https://doi.org/10.1016/j.jneuroling.2010.12.003 doi: 10.1016/j.jneuroling.2010.12.003 [Cited at p. 6]

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103. Retrieved from https://doi.org/10.1016/j.jml.2013.04.003 doi: 10.1016/j.jml.2013.04.003 [Cited at p. 4]

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*(3), 531–559. Retrieved from https://doi.org/10.1037/0033-295x.112.3.531 doi: 10.1037/0033-295x.112.3.531 [Cited at p. 4, 5, 50]

Ferguson, C. F. (1959). Diglossia. *Word*, *15*(2), 325–340. [Cited at p. 8]

Frenck-Mestre, C., Osterhout, L., McLaughlin, J., & Foucart, A. (2008). The effect of phonological realization of inflectional morphology on verbal agreement in french: Evidence from ERPs. *Acta Psychologica*, *128*(3), 528–536. Retrieved from https://doi.org/10.1016/j.actpsy.2007.12.007 doi: 10.1016/j.actpsy.2007.12.007 [Cited at p. 6]

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84. [Cited at p. 39]

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357–1392. Retrieved from https://doi.org/10.1152/physrev.00006.2011 doi: 10.1152/physrev.00006.2011 [Cited at p. 39]

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax–semantic interface. *Brain Research*, *1146*, 50–58. Retrieved from https://doi.org/10.1016/j.brainres.2006.08.038 doi: 10.1016/j.brainres.2006.08.038 [Cited at p. 39]

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of human language* (p. 73-113). Cambridge, Mass: MIT Press. [Cited at p. 3, 43, 50]

Grey, S., Tanner, D., & Van Hell, J. G. (2017). How right is left? handedness modulates neural responses during morphosyntactic processing. *Brain Research*, *1669*, 27–43. Retrieved from https://doi.org/10.1016/j.brainres.2017.05.024 doi: 10.1016/j.brainres.2017.05.024 [Cited at p. 40]

Guajardo, L. F., & Wicha, N. Y. (2014). Morphosyntax can modulate the N400 component: Event related potentials to gender-marked post-nominal adjectives. *NeuroImage*, *91*, 262–272. Retrieved from https://doi.org/10.1016/j.neuroimage.2013.09.077 doi: 10.1016/j.neuroimage.2013.09.077 [Cited at p. 6]

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic Gender and Semantic Expectancy: ERPs Reveal Early Autonomy and Late Interaction. *Journal of Cognitive Neuroscience*, *12*(4), 556–568. [Cited at p. 6]

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (sps) as an erp measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439–483. Retrieved from https://doi.org/10.1080/01690969308407585 doi: 10.1080/01690969308407585 [Cited at p. 5, 6]

Harley, H., & Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis. *Language*, *78*(3), 482–526. Retrieved from https://doi.org/10.1353/lan.2002.0158 doi: 10.1353/lan.2002.0158 [Cited at p. 4, 43]

Hartsuiker, R. J., Antón-Méndez, I., & van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, *45*(4), 546–572. Retrieved from https://doi.org/10.1006/jmla.2000.2787 doi: 10.1006/jmla.2000.2787 [Cited at p. 4]

Haskell, T. R., Thornton, R., & MacDonald, M. C. (2010). Experience and grammatical agreement: Statistical learning shapes number agreement production. *Cognition*, *114*(2), 151–164. Retrieved from https://doi.org/10.1016/j.cognition.2009.08.017 doi: 10.1016/j.cognition.2009.08.017 [Cited at p. 4]

Huynh, H., & Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact f-distributions. *Journal of the American Statistical Association*, *65*(332), 1582-1589. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01621459.1970.10481187 [Cited at p. 16, 17]

Kaan, E. (2002). [Cited at p. 5] *Journal of Psycholinguistic Research*, *31*(2), 165–193. Retrieved from https://doi.org/10.1023/a:1014978917769 doi: 10.1023/a:1014978917769

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201. [Cited at p. 6, 41, 49]

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NY: Prentice Hall. [Cited at p. 16, 17]

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*(5), 539–550. Retrieved from https://doi.org/10.3758/bf03196991 doi: 10.3758/bf03196991 [Cited at p. 5]

Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement attraction in spanish comprehension. *Journal of Memory and Language*, *82*, 133–149. Retrieved from https://doi.org/10.1016/j.jml.2015.02.002 doi: 10.1016/j.jml.2015.02.002 [Cited at p. 4]

Leinonen, A., Brattico, P., Järvenpää, M., & Krause, C. M. (2008). Event-related potential (ERP) responses to violations of inflectional and derivational rules of finnish. *Brain Research*, *1218*, 181–193. Retrieved from https://doi.org/10.1016/j.brainres.2008.04.049 doi: 10.1016/j.brainres.2008.04.049 [Cited at p. 5]

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in english, german, and italian. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 127–150. Retrieved from https://doi.org/10.1016/s0022-5371(84)90093-8 doi: 10.1016/s0022-5371(84)90093-8 [Cited at p. 5, 43]

Mancini, S., Massol, S., Duñabeitia, J. A., Carreiras, M., & Molinaro, N. (2019). Agreement and illusion of disagreement: An ERP study on basque. *Cortex*, *116*, 154–167. Retrieved from https://doi.org/10.1016/j.cortex.2018.08.036 doi: 10.1016/j.cortex.2018.08.036 [Cited at p. 48]

Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011a). A person is not a number: Discourse involvement in subject–verb agreement computation. *Brain Research*, *1410*, 64–76. Retrieved from https://doi.org/10

.1016/j.brainres.2011.06.055 doi: 10.1016/j.brainres.2011.06.055 [Cited at p. 6, 7, 44]

Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011b). When persons disagree: An ERP study of unagreement in spanish. *Psychophysiology*, *48*(10), 1361–1371. Retrieved from https://doi.org/10.1111/j.1469-8986.2011.01212.x doi: 10.1111/j.1469-8986.2011.01212.x [Cited at p. 48]

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement: A reply to tanner et al. *Cortex*, *66*, 156–159. Retrieved from https://doi.org/10.1016/j.cortex.2014.06.009 doi: 10.1016/j.cortex.2014.06.009 [Cited at p. 42]

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: Erp findings and future directions. *Cortex*, *47*(8), 908-930. doi: 10.1016/j.cortex.2011.02.019 [Cited at p. 5, 39, 42]

Molinaro, N., Vespignani, F., & Joba, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research*, *1228*, 161–176. [Cited at p. 6]

Molinaro, N., Vespignani, F., Zamparelli, R., & Job, R. (2011). Why brother and sister are not just siblings: Repair processes in agreement computation. *Journal of Memory and Language*, *64*(3), 211–232. Retrieved from https://doi.org/10.1016/j.jml.2010.12.002 doi: 10.1016/j.jml.2010.12.002 [Cited at p. 42]

Münte, T. F., Matzke, M., & Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience*, *9*(3), 318–329. Retrieved from https://doi.org/10.1162/jocn.1997.9.3.318 doi: 10.1162/jocn.1997.9.3.318 [Cited at p. 39]

Münte, T. F., Szentkuti, A., Wieringa, B. M., Matzke, M., & Johannes, S. (1997). Human brain potentials to reading syntactic errors in sentences of different complexity. *Neuroscience Letters*, *235*(3), 105–108. Retrieved from https://doi.org/10.1016/s0304-3940(97)00719-2 doi: 10.1016/s0304-3940(97)00719-2 [Cited at p. 39]

Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, *1164*, 81–94. [Cited at p. 6, 41, 47, 48]

Nichols, J., & Bickel, B. (2013). Locus of marking: Whole-language typology. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/25 [Cited at p. 3]

Osterhout, L., & Mobley, L. A. (1995). Event-Related Braing Potentials Elicited by Failure to Agree. *Journal of Memory and Language*, *34*, 739–773. [Cited at p. 5]

Palolahti, M., Leino, S., Jokela, M., Kopra, K., & Paavilainen, P. (2005). Event-related potentials suggest early interaction between syntax and semantics during on-line sentence comprehension. *Neuroscience Letters*, *384*(3), 222–227. Retrieved from https://doi.org/10.1016/j.neulet.2005.04.076 doi: 10.1016/j.neulet.2005.04.076 [Cited at p. 5]

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–456. [Cited at p. 4]

Roehm, D., Bornkessel, I., Haider, H., & Schlesewsky, M. (2005). When case meets agreement: event-related potential effects for morphology-based conflict resolution in human language comprehension. *Neuroreport*, *16*(8), 875–878. [Cited at p. 6]

Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, *384*(3), 228–233. Retrieved from https://doi.org/10.1016/j.neulet.2005.04.077 doi: 10.1016/j.neulet.2005.04.077 [Cited at p. 6]

Ryding, K. C. (2005). *A reference grammar of Modern Standard Arabic*. Cambridge University Press. [Cited at p. 8]

Shlonsky, U. (1989). The hierarchical representation of subject verb agreement. *Unpublished manuscript, University of Haifa*. Retrieved from http://www.unige.ch/lettres/linguistique/shlonsky/readings/hierarchical.pdf [Cited at p. 4]

Silva-Pereyra, J. F., & Carreiras, M. (2007). An ERP study of agreement features in spanish. *Brain Research*, *1185*, 201–211. Retrieved from https://doi.org/10.1016/j.brainres.2007.09.029 doi: 10.1016/j.brainres.2007.09.029 [Cited at p. 6, 7]

Silverstein, M. (1976). Hierarchy of features and ergativity. In R. M. W. Dixon (Ed.), *Grammatical categories in Australian languages* (pp. 112–171). New Jersey, NJ: Humanities Press. [Cited at p. 4]

Simons, G. F., & Fenning, C. D. (Eds.). (2017). *Ethnologue: Languages of the World. Online version. www.ethnologue.com.* Dallas, TX: SIL International. Retrieved from https://www.ethnologue.com/language/ara [Cited at p. 8]

Steele, S. (1978). Word order variation: a typological study. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language, vol. 4: Syntax.* (pp. 585–623). Stanford University Press: Walter de Gruyter. [Cited at p. 3]

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A commentary on “grammatical agreement processing in reading: ERP findings and future directions” by molinaro et al., 2011. *Cortex*, *66*, 149–155. Retrieved from https://doi.org/10.1016/j.cortex.2014.04.007 doi: 10.1016/j.cortex.2014.04.007 [Cited at p. 40]

Tanner, D. (2018). Robust neurocognitive individual differences in grammatical agreement processing: A latent variable approach. *Cortex*. Retrieved from https://doi.org/10.1016/j.cortex.2018.10.011 doi: 10.1016/j.cortex.2018.10.011 [Cited at p. 40]

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289–301. Retrieved from https://doi.org/10.1016/j.neuropsychologia.2014.02.002 doi: 10.1016/j.neuropsychologia.2014.02.002 [Cited at p. 40]

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*, *6*. Retrieved from https://doi.org/10.3389/fpsyg.2015.00347 doi: 10.3389/fpsyg.2015.00347 [Cited at p. 4]

Van de Meerendonk, N., Chwilla, D. J., & Kolk, H. H. (2013). States of indecision in the brain: Erp reflections of syntactic agreement violations versus visual degradation. *Neuropsychologia*, *51*(8), 1383-1396. doi: 10.1016/j.neuropsychologia.2013.03.025 [Cited at p. 40]

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, *34*, 186–215. [Cited at p. 4, 50]

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*(2), 206–237. Retrieved from https://doi.org/10.1016/j.jml.2009.04.002 doi: 10.1016/j.jml.2009.04.002 [Cited at p. 4]

Wechsler, S. (2009). Agreement features. *Language and Linguistics Compass*, *3*(1), 384–405. Retrieved from https://doi.org/10.1111/

j.1749-818x.2008.00100.x doi: 10.1111/j.1749-818x.2008.00100.x [Cited at p. 3]

Zawiswewski, A., Santestegan, M., & Laka, I. (2016). Phi-features reloaded: An event-related potential study on person and number agreement processing. *Applied Psycholinguistics*, *37*(03), 601–626. Retrieved from https://doi.org/10.1017/s014271641500017x doi: 10.1017/s014271641500017x [Cited at p. 6, 41, 42, 47, 49]

Zawiszewski, A., & Friederici, A. D. (2009). Processing canonical and non-canonical sentences in basque: The case of object–verb agreement as revealed by event-related brain potentials. *Brain Research*, *1284*, 161–179. Retrieved from https://doi.org/10.1016/j.brainres.2009.05.099 doi: 10.1016/j.brainres.2009.05.099 [Cited at p. 6]