ISSN 3042-030X (Online)

STUDIES IN LANGUAGE AND MIND
6

FILOZOFSKI FAKULTET · NOVI SAD ·

Novi Sad, 2025.

UNIVERSITY OF NOVI SAD, FACULTY OF PHILOSOPHY,
21000 Novi Sad,
Dr Zorana Đinđića 2
www.ff.uns.ac.rs

On behalf of the publisher:
Prof. dr Milivoj Alanović

Cover image:
Unknown title (Laurier Canadian Excellence)

Layout
Igor Lekić

ISSN
3042-030X (Online)

URL
http://slm.ff.uns.ac.rs/

Novi Sad, 2025.

CONTENTS

Elena Gavruseva

NOT ALL NOUNS ARE ACQUIRED *EN MASSE*: COUNTABILITY AND (IN)DEFINITENESS IN SECOND LANGUAGE LEARNING OF ENGLISH ARTICLES 5

Lea Elias Isen & Kasper Boye

LANGUAGE PRODUCTION IN PTSD: EVIDENCE FOR FEATURES ASSOCIATED WITH BROCA'S APHASIA IN TWO CASE STUDIES 29

Regina Mezőlaki & Lívia Ivaskó

FACE-SAVING ACTS IN FAUX PAS STORIES: THE IMPORTANCE OF CONTEXTUAL CUES DURING INFERENTIAL PROCESSES OF AUTISTIC AND NON-AUTISTIC ADULTS 55

Marwa Mekni Toujani

PERSONALITY TRAITS MODULATE VISUAL ATTENTION IN PICTURE NAMING/CATEGORISING TASKS 75

Mauro Viganò, Giulia Gilardone, Dario Cassinelli, & Carlo Cecchetto

THE CHALLENGE OF VARIABILITY FOR SYNTACTIC ACCOUNTS OF AGRAMMATISM: A STUDY ON FEATURE DISSIMILARITY IN ITALIAN RELATIVES 93

Elena Gavruseva*
University of Iowa, United States of America
DOI: 10.19090/slm.6.12

NOT ALL NOUNS ARE ACQUIRED *EN MASSE*: COUNTABILITY AND (IN)DEFINITENESS IN SECOND LANGUAGE LEARNING OF ENGLISH ARTICLES

This study investigates how L1-Slavic/L2 English learners (N=41) acquire article use across contexts defined by noun type (count vs. mass) and (in)definiteness (definite vs. indefinite). Using an online comprehension task with four conditions (Count–A, Count–The, Mass–Null, Mass–The), we analyze accuracy and error patterns across three proficiency levels (native-like, advanced, intermediate). Results show that mass nouns pose greater challenges than count nouns. Learners frequently overgeneralize *a* to mass noun contexts and make substitution errors, revealing persistent issues with noun categorization. Statistical analyses confirm significant effects of noun type, condition, and proficiency. Our findings contribute to growing research on article acquisition and the semantics of noun classification.

Keywords: second language learning, count/mass distinction, English articles, (in)definiteness, noun semantics, semantics-pragmatics interface

1. INTRODUCTION

The acquisition of English articles remains a persistent challenge for post-puberty learners whose native languages (L1s) lack a comparable determiner system. Numerous studies have documented the variability and difficulty that L2-English learners face in mastering article use, especially in contexts where the distinction between definiteness and indefiniteness intersects with multiple discourse/pragmatic features (e.g., Abudalbuh, 2016; Derkach & Alexopoulou, 2024; Ionin, Montrul & Crivos, 2011; Ionin & Montrul, 2010; Ionin, Zubizarreta & Maldonado, 2008; Kim & Lakshmanan, 2009; Ko, Ionin & Wexler, 2010; Liu & Gleason, 2002; Montrul & Ionin, 2010; Snape, Mayo & Gureli, 2013; Tryzna, 2009; among others). An illustration of the complexity involved in article semantics is provided in example (1):

(1) A: Have you ever been to the White House?
B: No, but I've been to a white house.

Despite being nearly identical in lexical content, the two noun phrases in (1) differ fundamentally in meaning, a distinction signaled by the articles *the* and *a*. The definite noun phrase *the White House* refers to a culturally salient, unique referent, while the indefinite *a white house* introduces a non-specific, novel referent. As Hawkins (1978, 1991) notes, uses of *the* often rely on cultural or situational familiarity, requiring both linguistic and extralinguistic knowledge. For L2 learners who come

* elena-gavruseva@uiowa.edu

from articleless L1 backgrounds (e.g., Russian), this dual reliance on linguistic form and shared knowledge presents a formidable learning task. In Russian, to convey the semantic contrast shown in (1), speaker B's response would need to be expressed as 'house of white color' (*v dome belogo cveta*) or the adjective would need to appear in a diminutive form (*v belenkom dome*) to differentiate the meaning from 'the White House.'

L2 acquisition data from learners whose L1 lacks articles, such as Russian or Korean, consistently show two primary types of article errors: omission and substitution. Omission errors result in ungrammatical output (e.g., **Honda was totaled* instead of *A Honda was totaled*). In contrast, substitution errors arguably reflect incorrect semantic categorization (e.g., using *a* in a definite context, or *the* in an indefinite one). Substitution errors occur in both directions, suggesting that learners have not yet internalized the semantic/pragmatic features that differentiate *a* and *the*. For example, a Russian-speaking learner might refer to **a father of my fiancée* where the definite *the father of my fiancée* is required, or say *the beautiful silver necklace* in a context where the referent is novel and should be introduced with *a* (Ionin, Ko & Wexler, 2004).

Although much prior research has investigated article acquisition with a primary focus on the semantics of definiteness and specificity, another key factor that interacts with article choice in English is grammatical noun type, i.e. whether the noun is count or mass (Chierchia, 2010; Gardelle, 2019; Rothstein, 2010). The English article system is sensitive to this grammatical distinction: *a* combines only with singular count nouns, *the* is grammatically compatible with both count and mass nouns, and mass nouns often appear with a null article (e.g., *I like antique furniture*, not **an antique furniture*). As shown in Table 1, this creates a morphosyntactic asymmetry in the English article system:

Table 1. English article use across noun types and definiteness.

Noun type	**Indefinite**	**Definite**
Count, sg.	a book	the book
Count, pl.	Ø books	the books
Mass	Ø furniture	the furniture

Because the null article co-occurs with both plural count and mass nouns, its acquisition adds an additional layer of complexity. Prior studies suggest that article errors are not uniformly distributed across contexts but are instead sensitive to the interplay of definiteness, noun type, and referential specificity. A growing body of research has begun to examine the role of the count/mass distinction in L2 acquisition, with several studies examining its interaction with (in)definiteness and/or plural

morphology (Cerqueglini, 2022; Choi, Ionin & Zhu, 2018; Hermas, 2019, 2020; Inagaki, 2014; Sabir, 2019; Snape, 2008; Tang, Fiorentino & Gabriele, 2023; Yoon, 1993; Wu, 2025).

Contributing to this line of work, the present study investigates how learners of English with Slavic L1 backgrounds (Croatian, Czech, Polish, Russian, Slovak, and Ukrainian) acquire article use in contexts where grammatical noun type (count vs. mass) and (in)definiteness interact. Specifically, we examine learners' use of *a* vs. *the* with singular count nouns and of the *null* article vs. *the* with mass nouns. We compare our participants' performance across proficiency levels and (in)definiteness contexts and examine which article contexts pose the greatest challenge in relation to noun type.

The remainder of this article is structured as follows. Section 2 reviews the relevant background literature on article acquisition and the role of noun type; it also outlines the general characteristics of Slavic noun phrases. Section 3 describes the design of the experiment, followed by a discussion of the results. Section 4 discusses the findings in light of previous research. Section 5 concludes the article.

2. BACKGROUND

This section provides the theoretical foundation for the study. Subsection 2.1 reviews previous research on the role of noun type—the count/mass distinction—in shaping L2 acquisition of English articles. Subsection 2.2 turns to the general characteristics of noun phrases in Slavic languages, which represent the L1s of the learners in this study. Together, these subsections situate the study within broader work on article acquisition and highlight why Slavic-speaking learners face persistent challenges with English determiners.

2.1. The role of noun type in L2 article acquisition

A growing body of research in L2 acquisition has addressed the importance of noun type (count vs. mass) in shaping L2 learners' use of articles and/or plural morphology. Yoon (1993) was among the first to investigate the perception of noun countability by L1-Japanese/L2-English learners, focusing on the interaction of noun countability and article use (*null* vs. indefinite *a*). Yoon found that learners struggled not only with article semantics but also with determining how context influences the grammatical noun class (e.g., using *defiance* or *appreciation* as count nouns). Similarly, Wu (2025) investigated Chinese-speaking learners' acquisition of flexible nouns (e.g., *cake(s)*, *stone(s)*, *paper(s)*, etc.) that can shift between count and mass interpretations depending on context. While learners performed well in canonical uses, they showed difficulty with less typical meanings, such as interpreting *cake* and *stone* as mass nouns. Inagaki (2014) explored L1-Japanese learners' acquisition of the mass–count distinction in English, not through article use, but via quantity judgment tasks involving

proportional reasoning. Learners showed native-like patterns when interpreting canonical count nouns and substance-mass nouns (e.g., *mustard*), basing their judgments on number or volume as appropriate. However, their performance declined when nouns were ambiguous between mass and count uses (e.g., *string* vs. *strings* in English). These findings suggest that even when learners grasp core count–mass distinctions, noun semantic flexibility remains a learning challenge.

Building on Hua and Lee's (2005) research, Snape (2008) conducted an experimental study to investigate L2 learners' sensitivity to the count/mass distinction and their interpretation of definite descriptions. The study employed two tasks: a grammaticality judgement task to assess learners' ability to recognize acceptable mass and count noun uses, and a forced elicitation task that tested three types of definite expressions: (a) referential definites based on shared knowledge between speaker and hearer, (b) encyclopaedic definites rooted in culturally unique references (similar to *the White House* example in (1)), and (c) larger situation definites, which presuppose situational familiarity without reference to any specific entity. Crucially, noun type (count vs. mass) was manipulated across these conditions. The results revealed an asymmetry: while L1-Japanese learners generally performed well on count noun items, they frequently misjudged grammatical mass noun uses and were more likely to accept ungrammatical mass plurals (e.g., **few sunshines*). L1-Spanish learners demonstrated stronger performance with mass nouns than their Japanese counterparts, but showed similar difficulty when interpreting plural mass nouns in ungrammatical contexts. These findings suggest that even learners from typologically different L1 backgrounds may exhibit persistent challenges with the mass noun category, particularly when the interaction of plurality and definiteness is manipulated.

Hermas (2019) explored how L1-Moroccan Arabic speakers acquire mass generic noun phrases in L3 English, with special attention to the influence of non-facilitative transfer. The findings revealed a clear developmental trajectory: at the pre-intermediate level, learners inappropriately used definite articles with mass generics (indefinite contexts). Intermediate learners began to alternate between definite and bare mass noun forms, suggesting an interlanguage stage in which competing grammatical representations coexisted. Advanced learners demonstrated convergence with native speaker norms, using bare nouns appropriately to express genericity. The study thus highlights the evolving role of transfer and the gradual refinement of grammatical representations in multilingual acquisition. Sabir (2019) also focused on L1-Arabic learners of English and confirmed that mass nouns present a learning challenge for this population of L2 learners.

In aggregate, these studies converge on a key finding: article acquisition in English is sculpted not only by semantic features such as (in)definiteness and specificity, but also by the morphosyntactic

properties of the noun. Mass nouns emerge as a consistent source of difficulty, especially in indefinite contexts where the null article is required. The present study contributes to this growing literature by focusing on L1-Slavic speakers learning English. We aim to systematically compare their article use in contexts involving mass and count nouns across definite and indefinite domains.

2.2. (In)definiteness and general characteristics of Slavic noun phrases

This subsection outlines the main structural characteristics of noun phrases in the L1s of learners in this study: Croatian, Czech, Polish, Russian, Slovak, and Ukrainian. The discussion is based on Zlatić (2014), who surveys NP structure across Slavic. We draw on her examples from Polish and supplement them with our own data from Russian.

A central property of all these Slavic languages is that they are articleless: they lack a determiner system comparable to English. Noun phrases appear bare, without obligatory determiners modifying the noun. (In)definiteness interpretations are inferred from discourse/pragmatic context and word order. The default word order in Slavic is SVO, and definiteness is often associated with topicality (given information/theme), whereas indefiniteness is associated with new information (rheme). The Polish examples in (2) illustrate the typical contrasts:

(2) a. Student lubi Marie
student.Nom likes Mary.Acc
'The student likes Mary.'
b. Marie lubi student
Mary.Acc likes student.Nom
'A student likes Mary.'

In (2a), *student* is interpreted as definite when it occurs in subject/topic position, whereas in (2b) it occurs in post-verbal/rheme position and is interpreted as indefinite.

Slavic languages can also mark definiteness lexically, using demonstratives, which are optional and pragmatically licensed. Like English demonstratives, they agree with the noun in number, but unlike English, they also agree in gender and case in the singular (and only number/case in the plural). Example (3) from Russian illustrates this NP-internal agreement:

(3) Ja lublu et-u knig-u /et-i knig-i
I love this.Acc.Sg.Fem book.Acc.Sg.Fem /these.Acc.Pl books.Acc.Pl
'I love this book/these books.'

In addition, Slavic has an indefinite modifier that corresponds to the English indefinite determiner *some*. This form (e.g., *odin-Nom.Sg.Masc*/*odna-Nom.Sg.Fem* in Russian) appears in contexts requiring a specific indefinite interpretation and shows obligatory agreement in number, gender, and case:

(4) Odin čelovek pokazal nam nužnyj dom
some.Nom.Sg.Masc person.Nom.Sg.Masc showed us right house
'Some person showed us the right house.'

In this Russian example, *odin* is best translated as *some* rather than *a*, since its use signals a specific indefinite meaning. All Slavic languages spoken by the participants have equivalents of *odin/odna*, with similar discourse restrictions.[1]

The nominal system of Slavic encodes a count–mass distinction, with mass nouns marked for partitive/genitive case in object position in appropriate contexts. This is illustrated in (5) from Russian:

(5) a. Anna pokupala sol' na rynke
Anna bought salt.Acc.Sg.Fem at market
'Anna used to buy salt at the market.' (habitual action/imperfective verb)
b. Anna nasypala sol-i v solonku
Anna poured salt.Gen.Sg.Fem into saltshaker
'Anna poured some salt into the saltshaker.' (episodic/perfective verb)

In (5a), the mass feminine noun *sol'* has a generic interpretation in a habitual aspectual context and must appear in the accusative case. By contrast, in (5b) the genitive singular form *sol-i* encodes a partitive reading, denoting an unspecified amount of salt. In both uses, *sol'* is grammatically singular. When combined with demonstratives or adjectives, as in *eta sol'* ('this salt'), NP-internal agreement is expressed in number, gender, and case, just as with count nouns. It is ungrammatical to combine mass nouns with indefinite modifiers such as *odin/odna* (**odna sol'*).

Researchers differ on the precise syntactic analysis of Slavic noun phrases, with some arguing for a DP analysis and others for an NP analysis (e.g., Rutkowski, 2002; Pereltsvaig, 2007; Progovac, 1998; among others). Despite these theoretical differences, there is consensus on the core facts: (i) absence of obligatory determiners comparable to English *a/the*; (ii) agreement in number, gender, and case between demonstratives/*odin/odna* and the noun; (iii) reliance on information-structure distinctions

[1] A reviewer asks whether *odin* is pragmatically felicitous in B's response in example (1). It is not: *a white hous*e has a nonspecific indefinite interpretation in this context.

(theme/rheme) and word order for (in)definiteness interpretation; (iv) count-mass distinction in the nominal system. Thus, across L1s in our study, noun phrases share these fundamental structural properties. We can therefore treat our L2 sample as a coherent group whose article-related challenges in L2 English stem from comparable characteristics of their respective L1s.[2]

3. THE STUDY

Building on the findings from previous research, which underscore the challenges that noun type and (in)definiteness pose for L2 learners of English articles, the present study aims to investigate these dimensions more closely. While prior studies have highlighted the influence of semantic features such as specificity and atomicity, fewer studies have systematically examined how count versus mass nouns interact with article use across proficiency levels. This study aims to bridge this gap by exploring article choice in pragmatically rich dialogue contexts. Our study was guided by the following research questions:

1. How accurately do L1-Slavic/L2-English learners use articles in syntactic and semantic contexts involving different noun types (count vs. mass) and levels of definiteness (indefinite vs. definite)?
2. Are there systematic differences in article use across learner proficiency levels?
3. Which combinations of noun type and definiteness are most prone to article omission and/or substitution errors?

To address these questions, we designed a comprehension-based task, using brief conversational dialogues to elicit article judgments. The following sections describe the experimental design, report the results, and discuss their theoretical and empirical implications.

3.1 Participants

Forty one non-native speakers (NNSs) and fourteen native speakers (NSs) of English were recruited to participate in the study. Prior to the experiment, participants completed a language background questionnaire that elicited information about: (i) age at the time of testing, (ii) native language (defined as the language acquired from birth), (iii) any additional language(s) spoken in early

[2] A reviewer raised the possibility that microvariation in the nominal systems of the included Slavic languages might have subtly influenced L2 learners' accuracy rates in determiner use. In this study, we prioritized the shared core properties of Slavic nominal systems, which justify treating participants as a coherent group for the contrasts under investigation. Nonetheless, future research could examine whether nuanced microvariation across Slavic nominal systems influences how learners acquire English articles.

childhood, (iv) age of first exposure to English, (v) age of arrival (immigration) in the U.S., and (vi) total length of residence in the U.S.

All NNS had begun to learn English as an L2 in their country of birth and were residing in the U.S. at the time of testing. Their native languages included Russian (n = 27), Slovak (n = 7), Ukrainian (n = 4), Croatian (n = 1), Czech (n = 1), and Polish (n = 1). Five participants reported bilingual exposure to another Slavic language during childhood, with language pairs consisting of Ukrainian/Russian (n = 4) and Slovak/Czech (n = 1).

L2 participants also provided a self-assessment of their English proficiency by selecting one of the following categories: native-like, advanced, high-intermediate, intermediate, low-intermediate, or beginner. Based on these self-reports, participants were grouped into three proficiency levels: Native-like (n = 16), Advanced (n = 9), and Intermediate (n = 16). Table 2 summarizes demographics and language background data for each group.

Table 2. Participant demographics and English exposure history by proficiency group.

	Number (M/F)	Mean age at time of testing	Age of exposure to English in the country of birth	Age of arrival in the US	Total years of residency in the US
Native mean (SD) range	14 (3/11)	 39.9 (17.5) 18–72	–	–	–
Native-like mean (SD) range	16 (2/14)	 30.4 (8.67) 22–44	 9.1 (3.9) 5–19	 20.6 (8.56) 8–42	 8.8 (6.4) 3–23
Advanced mean (SD) range	9 (3/6)	 38.4 (8.9) 26–51	 9.2 (3.2) 5–15	 23.8 (6.3) 15–35	 13.6 (7.8) 2–23
Intermediate mean (SD) range	16 (7/9)	 47 (11.1) 22–59	 10.3 (2.1) 6–15	 32.6 (3.8) 22–38	 14.4 (9.5) 4–30

Native-like group. This group consisted of 16 participants (2 male, 14 female), with a mean age of 30.4 years. These learners were first exposed to English in their country of birth at a mean age of 9.1 years, arrived in the U.S. at a mean age of 20.6 years, and had resided in the U.S. continuously for an

average of 8.8 years. (Only two participants arrived in the U.S. before the onset of puberty, at ages 8 and 11.)

Advanced group. This group included 9 participants (3 male, 6 female), with a mean age of 38.4 years. Their mean age of first exposure to English was 9.2 years, and they arrived in the U.S. at a mean age of 23.8 years. Their average length of continuous residence in the US was 13.6 years.

Intermediate group. This group included 16 participants (7 male, 9 female), with a mean age of 47 years. These participants were first exposed to English at a mean age of 10.3 years, arrived in the U.S. at a mean age of 32.6 years, and had the longest average duration of residence in the U.S., at 14.4 years.[3]

3.2 Materials

The experimental stimuli consisted of 48 short written dialogues, evenly distributed across four conditions based on (in)definiteness and noun type: *Indefinite Count-A*, *Indefinite Mass-Null*, *Definite Count-The*, and *Definite Mass-The*. Each dialogue took the form of an informal exchange between two speakers and contained a target noun phrase embedded within a sentence (see Table 4). Each target noun appeared with a blank space where participants were asked to choose the most appropriate article from the following multiple-choice menu: *a*, *the*, or *no article needed*. Table 3 provides a list of the target nouns used in each condition.

[3] This group includes 9 learners self-rated as high-intermediate and 7 as low-intermediate. However, t-test analyses showed no significant differences in performance between these two subgroups. Therefore, they were combined into a single group, referred to as *Intermediate* throughout the study.

Table 3. Target nouns by experimental condition and noun type.

Condition	Target Nouns
Indefinite Count (a)	book, trip, package, job, shop, lamp, house, notebook, coat, bike, necklace, wallet
Indefinite Mass (null)	homework, weather, evidence, success, information, land, rice, wealth, furniture, mail, happiness, sushi
Definite Count (the)	magazine, bus, party, job, website, flash drive, car, laptop, house, key, answer, exit
Definite Mass (the)	homework, weather, evidence, success, information, land, rice, cash, water, milk, patience, butter

Each condition was matched for key structural properties: (i) all target noun phrases were first mentions (i.e., they had no explicit, prior antecedent in the dialogue), (ii) they occurred as syntactic objects of transitive predicates, (iii) they were modified by restrictive relative clauses, and (iv) each condition featured the same set of verbal predicates (*look for, hope for, plan, wait for, have, own, keep, know, get, find, lose, buy*) to ensure consistency in lexical environments. Table 4 presents sample experimental items.

Table 4. Sample experimental dialogues.

Mass-Null	Person A: Is Linda teaching this semester? She looks stressed. Person B: Yes, that's because she works too much. For tomorrow, she is planning __ homework that'll keep her students busy for two days! Person A: What's it about? Person B: Irregular verbs!
Mass-The	Person A: Our new police officer isn't good, is he? Person B: No, he isn't. Look at him. He is digging through his desk, looking for ___ evidence that we put together for him last night. Person A: Oh, he's going to get in trouble for that.

In addition to the 48 experimental items, we constructed 48 distractor dialogues followed by a three-option multiple-choice menu.

3.3 Procedure

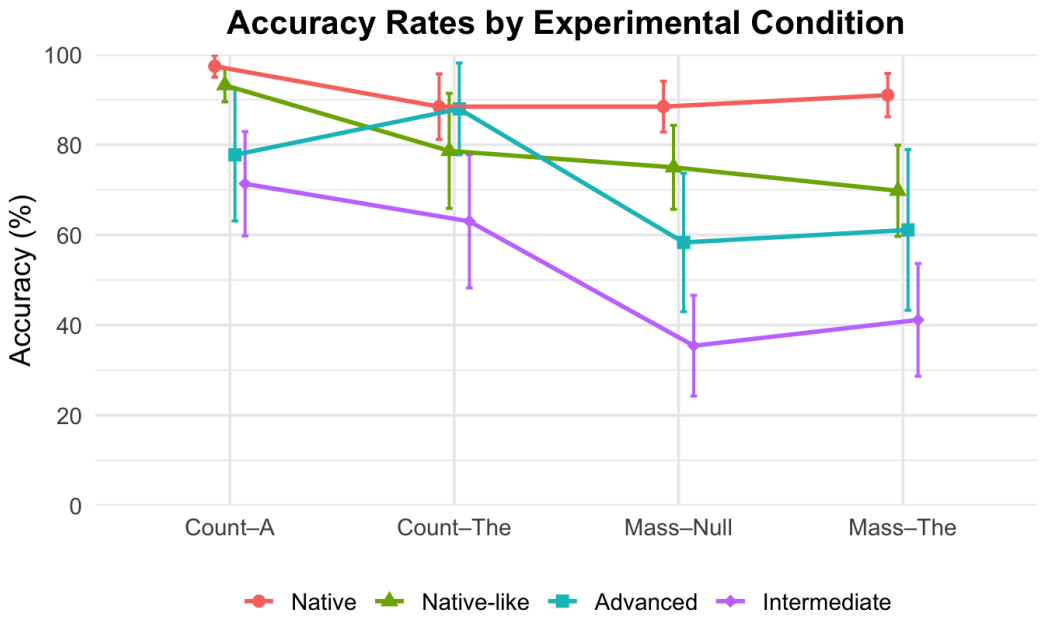
The experimental task was administered using Qualtrics, a secure web-based survey platform. All participants were recruited via the author's professional and personal network and received an anonymous link to the survey by email. They were instructed to complete the survey in a single sitting and in a quiet environment to ensure focused participation. The survey began with a language background questionnaire, which served as the first block of items. This was followed by a set of instructions introducing the task. Participants were asked to imagine the dialogues as representative of informal, everyday conversation.

Before beginning the main task, participants completed a brief practice item (unrelated to articles) to familiarize themselves with the test format. They then proceeded to the full set of experimental and distractor dialogues. The presentation order of both the dialogue items and the multiple-choice response options was randomized using a built-in randomization function in Qualtrics. This ensured that each participant received a unique item and response sequence. Non-native participants completed the task in 40 to 70 minutes, depending on their proficiency level. Upon completion of the task, participants received a $10 digital gift card as compensation for their time.

3.4 Results

Descriptive statistics. Figure 1 presents an overview of participants' accuracy rates across the four experimental conditions (Count-A, Count-The, Mass-Null, and Mass-The) by proficiency group (Native, Native-like, Advanced, Intermediate).

Figure 1. Accuracy rates by experimental condition: percent correct with 95% CI.



The graph clearly illustrates a general trend: accuracy decreases as noun type shifts from count to mass, and as proficiency declines from Native-like to Intermediate. Native speakers show near-ceiling performance in all conditions. Native-like and Advanced L2 learners perform well in the count noun conditions but exhibit greater variability in mass noun contexts. Intermediate learners show the lowest accuracy overall, with particularly sharp declines in the Mass–Null and Mass–The conditions. These group-level accuracy patterns closely align with participants' self-rated proficiency, supporting the validity of the self-assessment categories used in the analysis.

To provide a more granular view of article choice patterns, Table 5 reports the full distribution of article responses (*a*, *the*, and null) by proficiency group and condition.

Table 5. Distribution of article responses by proficiency group and condition (percent correct).

		Native	Native-like	Advanced	Intermediate
Count–A	a	96%	93%	78%	71%
	the	4%	6%	22%	26%
	null	0	1%	0	3%
Count–The	the	89%	79%	88%	63%
	a	11%	21%	12%	33%
	null	0	0	0	4%
Mass–Null	null	88%	75%	58%	35%
	the	6%	17%	28%	25%
	a	6%	8%	14%	40%
Mass–The	the	92%	70%	61%	41%
	a	0	1%	4%	25%
	null	8%	29%	35%	34%

Accuracy rates across the four article conditions (Count–A, Count–The, Mass–Null, and Mass–The) reveal effects of both proficiency level and noun type on determiner choice, with learners overusing *a* and *the* in Mass–Null contexts and overusing the *null* article in Mass–The contexts. Overuse of *a* in the Mass–The context is restricted to the Intermediate group.

ANOVA analyses. To determine whether the observed patterns in article accuracy across conditions and proficiency groups were statistically significant, we conducted a series of repeated measures analyses of variance (ANOVAs). The results are summarized in Table 6.

Table 6. Results of repeated measures ANOVA on article accuracy by proficiency level, noun type, and condition.

Effect	F-statistic and p-value
Proficiency level	$F(3, 47) = 39.281$, $p < 0.001$
Condition	$F(3, 146) = 16.052$, $p < 0.001$
Proficiency level x Condition	$F(9, 146) = 2.473$, $p < 0.01$
Noun type (count vs. mass)	$F(1, 154) = 44.210$, $p < 0.001$
Proficiency level x Noun type	$F(3, 154) = 4.585$, $p < 0.001$
(In)definiteness	$F(1, 155) = 0.242$, $p < 0.623$ (ns)
Proficiency level x (in)definiteness	$F(3, 155) = 1.434$, $p < 0.235$ (ns)

The ANOVA revealed a significant main effect of proficiency level ($p < .001$), confirming that article accuracy varied systematically across the four proficiency groups. There was also a significant main effect of condition ($p < .001$), as well as a significant level × condition interaction ($p < .01$). These

results indicate that the influence of article-noun condition differs across proficiency groups, warranting further analyses to explore the nature of this interaction.

We also tested the effect of noun type (count vs. mass) and found a significant main effect ($p < .001$), as well as a significant level × noun type interaction ($p < .001$). These findings suggest that mass nouns are more difficult overall and that this difficulty increases with lower proficiency. In contrast, the effect of (in)definiteness (definite vs. indefinite) was not significant ($p = .623$), nor was the level × (in)definiteness interaction ($p = .235$), indicating that (in)definiteness alone did not significantly impact accuracy or interact with proficiency level.

To complement the repeated-measures ANOVA analyses, we employed a binomial logistic mixed-effects regression model (GLMM) using the `glmer()` function from the `lme4` package in R. This model predicts the probability of correct article choice (coded as a binary outcome) based on fixed effects of proficiency level, condition, noun type, and (in)definiteness, along with their interactions. The model also includes random intercepts for subjects and items to account for repeated measures within participants and lexical items. The significance of each fixed effect was evaluated using Wald chi-square (χ^2) tests, which compare the fit of the full model with and without each predictor variable. These χ^2 values serve as the GLMM analogue to the F-tests in ANOVA, offering an estimation of effect significance at the trial level while accounting for both subject and item variability. The results are summarized in Table 7.

Table 7. Summary of fixed effects from binomial logistic mixed-effects regression model.

Effect	Wald χ^2 (df)	p-value
Proficiency level	$\chi^2(3) = 103.652$	p < .001***
Condition	$\chi^2(3) = 29.679$	p < .001***
Level × Condition	$\chi^2(9) = 38.007$	p < .001***
Noun type (count/mass)	$\chi^2(1) = 27.908$	p < .001***
Level × Noun type	$\chi^2(3) = 8.066$	p = .045*
(In)definiteness	$\chi^2(1) = 1.027$	p = .311
Level × (in)definiteness	$\chi^2(3) = 16.915$	p < .001***

p values: $p < .05$ = *, $p < .01$ = **, $p < .001$ = ***.

The repeated-measures ANOVA and the GLMM model yielded converging results across nearly all effects of interest. Both models identified highly significant main effects of proficiency level, condition, and noun type, as well as a robust interaction between proficiency level and condition. Additionally, both models revealed a significant interaction between proficiency level and noun type, albeit with a smaller effect size in the GLMM ($p < .05$). The primary divergence between the two models emerged in the proficiency level × (in)definiteness interaction: while the ANOVA reported this effect as non-significant, the GLMM analysis detected a statistically significant interaction ($\chi^2 = 16.915$, $p < .001$). This discrepancy likely reflects the increased sensitivity of the GLMM framework, which models trial-level data and includes subject and item random effects, thus capturing variance that is averaged out in ANOVA. Overall, the results demonstrate that both statistical approaches converge on the major patterns in the data.[4]

Across group comparisons. To further examine whether L2 learners are statistically distinguishable from native speakers, we present Table 8, which reports the results of a GLMM analysis comparing participant groups. The model first takes native speakers as the reference group and evaluates accuracy in four article–noun conditions across learner groups. This table provides insight into the developmental trajectory of article use in L2 English.

[4] We thank a reviewer for suggesting the use of a GLMM alongside repeated-measures ANOVA. Incorporating both analytic approaches allowed for a more fine-grained interpretation of the data and strengthened overall results.

Table 8. Fixed-effects estimates from binomial logistic mixed-effects regression model.

Group	Count–A	Count–The	Mass–Null	Mass–The
Native-like	–1.155 (0.627), z = –1.84, *p* = .165	–0.792 (0.363), z = –2.19, *p* = .077 †	–1.019 (0.360), z = –2.83, *p* = .013 *	–1.633 (0.381), z = –4.29, *p* < .001 ***
Advanced	–2.778 (0.623), z = –4.46, *p* < .001 ***	–0.058 (0.448), z = –0.13, *p* = .994	–1.852 (0.393), z = –4.71, *p* < .001 ***	–2.064 (0.417), z = –4.94, *p* < .001 ***
Intermediate	–3.173 (0.589), z = –5.39, *p* < .001 ***	–1.618 (0.352), z = –4.59, *p* < .001 ***	–2.905 (0.358), z = –8.12, *p* < .001 ***	–2.987 (0.381), z = –7.85, *p* < .001 ***
Constant (NS)	4.258 (0.584) ***	2.233 (0.628) ***	2.229 (0.629) ***	2.568 (0.617) ***

Note. Each column represents a model for a specific article type and noun condition. Coefficients reflect log-odds estimates relative to native speakers (reference group).
p < .05 = *, *p* < .01 = **, *p* < .001 = ***, † = marginal trend (p < .10).

The results reveal distinct accuracy patterns across L2 proficiency levels when compared to native speakers. Intermediate learners show statistically significant differences from native speakers in all four conditions (p < .001). This indicates a broad divergence from native-like article use, regardless of noun type or (in)definiteness. Advanced learners, by contrast, performed comparably to native speakers in the Count–The condition (*p* = .994), suggesting improvement with count nouns, but remained significantly different in the other three (all *p* < .001). The performance of the native-like learners does not differ from native speakers in the Count–A (*p* = .165) or Count–The (*p* = .077†) conditions, but they show statistically significant differences in the Mass–Null (*p* = .013) and Mass–The (*p* < .001) contexts. This highly restricted contrast suggests that even at very advanced stages of L2 acquisition, article use with mass nouns remains problematic.

We now turn to the results of a GLMM to draw comparisons within the L2 sample. We first focus on comparisons with native-like speakers as the reference group, and then compare advanced to intermediate learners. Table 9 summarizes the results.

Table 9. Results of GLMM comparing L2 groups across four conditions.

Condition	Groups compared	Estimate (SE)	*z*	*p* value
Count–A	Advanced vs. Native-like	–1.574 (0.444)	–3.55	p = .0008 ***
	Intermediate vs. Native-like	–1.964 (0.394)	–4.99	p < .0001 ***
Count–The	Advanced vs. Native-like	+0.732 (0.406)	+1.80	p = .131 (n.s.)
	Intermediate vs. Native-like	–0.821 (0.297)	–2.77	p = .011 *
Mass–Null	Advanced vs. Native-like	–0.843 (0.340)	–2.48	p = .025 *
	Intermediate vs. Native-like	–1.906 (0.297)	–6.42	p < .0001 ***
Mass–The	Advanced vs. Native-like	–0.433 (0.338)	–1.28	p = .339 (n.s.)
	Intermediate vs. Native-like	–1.356 (0.289)	–4.69	p < .0001 ***
Count–A	Intermediate vs. Advanced	–0.370 (0.379)	–0.98	p = .330 (n.s.)
Count–The	Intermediate vs. Advanced	–1.577 (0.414)	–3.81	p < .0001 ***
Mass–Null	Intermediate vs. Advanced	–1.085 (0.354)	–3.07	p = .002 **
Mass–The	Intermediate vs. Advanced	–0.944 (0.354)	–2.67	p = .008 **

*Note***.** The reference group for the first comparisons is the Native-like group; for the second set, Advanced learners serve as the reference. *p* values: $p < .05$ = *, $p < .01$ = **, $p < .001$ = ***.

The results reveal a clear proficiency-based trajectory in article use across noun types and (in)definiteness. When the native-like group is used as the intercept, the advanced group shows significantly lower accuracy than native-like speakers in the Count–A and Mass–Null conditions, with $p = .0008$ and $p = .025$, respectively. However, no significant differences are observed in Count–The or Mass–The contexts, indicating that advanced learners have largely converged with the Native-like group in their use of definite articles. In contrast, the intermediate group performs significantly worse than native-like learners in all four conditions, with the strongest effects observed in Mass–Null and Mass–The conditions (with $p < .0001$ in both). This suggests that intermediate learners struggle with the (in)definiteness distinctions irrespective of noun type.

When the advanced group is set as the reference level, comparisons with intermediate learners provide further insight into developmental stages. There is no significant difference in Count–A ($p = .330$), implying that intermediate learners have already acquired some competence with indefinite count nouns. However, the intermediate group performs significantly worse in the remaining three conditions: Count–The ($p = .0001$), Mass–Null ($p = .002$), and Mass–The ($p = .008$). These results suggest that definite contexts with count nouns and (in)definite contexts with mass nouns are acquired later. Taken

together, across-group comparisons point to a stage-like progression in article acquisition in L2 English: learners first master indefinite articles with count nouns, followed by definite articles with count nouns, and only later acquire target-like performance in the (in)definite contexts with mass nouns.

4. DISCUSSION

The findings of this study confirm that both noun type and proficiency level significantly influence article accuracy in the L2 English of L1-Slavic learners. As Figure 1 shows, accuracy declines in a gradient pattern: first as proficiency decreases, and second as noun type shifts from count to mass. This pattern is consistent with previous research showing that mass nouns pose unique challenges for L2 learners (Inagaki, 2014; Snape, 2008; Yoon, 1993).

In the Count–A condition, learners at all proficiency levels show relatively high accuracy. While substitution of *the* for *a* increases at lower proficiency levels (22% and 26% in Advanced and Intermediate learners, respectively), the target indefinite article *a* remains dominant. This suggests that count-singular contexts with indefinite reference are learned earlier and more robustly, possibly due to their clearer morphosyntactic marking, which may facilitate form-meaning mapping.

The Count–The condition yielded high overall accuracy among the non-native groups (63% to 88% range). Nonetheless, a substitution pattern emerged: learners across all levels occasionally used *a* instead of *the*, with the highest rates observed in Intermediate learners (33%). These results align with previously reported bidirectional substitution patterns in article choice with singular count nouns (e.g., Ionin, Ko & Wexler, 2004; Tryzna, 2009).

Mass noun conditions produced the lowest accuracy scores across the L2 groups, attesting to the difficulty of integrating (in)definiteness with noun type in article use. In the Mass–Null condition, Intermediate learners were particularly inaccurate, correctly selecting the null article only 35% of the time. Their most frequent error—substituting *a* for the null article— reflects overgeneralization of the singular indefinite marker to contexts where it is ungrammatical. This error pattern is most likely a consequence of miscategorization of some mass nouns as count singular nouns. Overuse of *the* was also observed across all proficiency levels. These findings align with earlier studies showing that L2 learners often misapply article forms in mass noun contexts, particularly when semantic individuation is unclear (Hermas, 2019; Sabir, 2019; Snape, 2008).

The Mass–Null and Mass–The conditions emerged as the most diagnostic in distinguishing native-like from native performance. While native speakers reached 88% and 92% accuracy in the two conditions, native-like learners fell to 75% and 70%, respectively, and substituted the null article 17% of

the time and definite article 29% of the time. Importantly, error patterns in the mass noun contexts were bidirectional: learners of all proficiency levels substituted the *null* determiner for *the* and the other way around. In addition, the overgeneralization of *a* to definite mass noun contexts in the Intermediate group suggests miscategorization of some mass nouns as countable, a pattern consistent with Yoon's (1993) findings.

Statistical analyses support these descriptive trends. The ANOVA and GLMM analyses revealed significant main effects of proficiency level, condition, and noun type, along with two key interaction effects: level × condition and level × noun type. These findings confirm that the difficulty of article selection is not uniform across contexts and is particularly sensitive to mass/count distinctions, especially at lower proficiency levels. By contrast, (in)definiteness as a main effect was not significant, suggesting that learners may rely more on noun type and referentiality cues than on (in)definiteness alone.

The GLMM across-group comparisons further illuminate these effects. Using native speakers as the reference group, Intermediate learners showed significant differences in all four conditions, with the most pronounced divergences in the Mass–Null and Mass–The contexts ($p < .001$). Advanced learners also struggled in both mass noun conditions ($p < .001$), especially in the Mass–The context, where substitution of the null article for *the* was frequent (35%). Even Native-like learners diverged significantly from native speakers in both mass contexts, with the strongest effect in Mass–The ($p < .001$), showing that this context remains a persistent challenge. This pattern of progressively narrowing differences across groups supports a stage-like developmental trajectory, in which increased proficiency leads to improved accuracy in count noun contexts, while difficulties with mass nouns—especially those requiring *the*—persist.

Taken together, the analyses converge in showing a significant effect of noun type, confirming that mass nouns present greater difficulty for L2 learners than count nouns. Future research will need to corroborate these findings in larger and more diverse samples of L1-Slavic speakers, with closer attention to microvariation across Slavic languages, which may contribute nuances to how learners approach English article semantics and form-meaning mappings.

5. CONCLUSION

The current study provides new evidence that L2 learners' article acquisition is shaped by complex interactions among noun type, (in)definiteness, and proficiency. Mass nouns in definite *the* contexts and indefinite contexts requiring the null article pose enduring challenges even for very advanced learners. These findings have implications not only for theories of L2 article acquisition—for

example, Ionin, Ko & Wexler's (2004) Fluctuation Hypothesis, which proposes that L2 learners fluctuate between different parameter settings when mapping articles to semantic features such as (in)definiteness and specificity. The results broadly support the idea of fluctuation, showing that it applies not only to *a/the* alternations in count noun contexts, but also to *null/the* alternations in mass noun contexts, thus extending the hypothesis beyond its original scope. This study has pedagogical implications as well: instructional approaches should more explicitly target article use with mass nouns and emphasize the semantic-pragmatic distinctions that underlie their usage.

Future research should continue to explore how semantic, syntactic, and pragmatic features interact to shape article use with mass nouns in L2 English. While the current study highlights clear proficiency-related patterns, further work is needed to disentangle what types of mass nouns (e.g., concrete vs. abstract or atomic vs. non-atomic) are more prone to misclassification and article errors. Additionally, experimental designs incorporating both comprehension and production tasks could offer a more complete view of how learners represent and use articles in mass noun contexts (Snape, 2018). Cross-linguistic comparisons involving learners from articleless languages versus those with article systems could also help refine models of transfer and universal semantic influence.

6. ACKNOWLEDGEMENTS

I would like to thank the participants in this study for supporting my research and for their patience in completing a rather lengthy test. I am also grateful to the two anonymous reviewers for their insightful comments and constructive feedback. My thanks go to the University of Iowa Social Science Research Center and to Peter Kortenkamp for their support with statistical analyses. All remaining errors are my own.

REFERENCES

Abudalbuh, M. (2016). The acquisition of English articles by Arabic L2-English learners: A semantic approach. *Arab World English Journal, 7*(2), 104–117. https://doi.org/10.24093/awej/vol7no2.7

Cerqueglini, L. (2022). Count/mass nouns and definiteness in Palestinian Arabic heritage speakers of English. *Romano-Arabica, 22*, 119–129. https://doi.org/10.62229/roar_xxii/8

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese, 174*(1), 99–149. https://doi.org/10.1007/s11229-009-9610-3

Choi, S. H., Ionin, T., & Zhu, Y. (2018). L1 Korean and L1 Mandarin L2 English learners' acquisition of the count/mass distinction in English. *Second Language Research, 34*(2), 147–177. https://doi.org/10.1177/0267658316687412

Derkach, K., & Alexopoulou, T. (2024). Definite and indefinite article accuracy in learner English: A multifactorial analysis. *Studies in Second Language Acquisition, 46*(3), 710–740. https://doi.org/10.1017/S0272263123000278

Gardelle, L. (2019). *Semantic plurality: English collective nouns and other ways of denoting plurality*. John Benjamins. https://doi.org/10.1075/slcs.210

Hawkins, J. A. (1978). *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. Croom Helm; Humanities Press.

Hawkins, J. A. (1991). On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics, 27*(2), 405–442. https://doi.org/10.1017/S0022226700012745

Hermas, A. (2019). Mass generics in L3 English: Acquisition route and transfer recovery. *International Journal of Language Studies, 13*(2), 51–72.

Hermas, A. (2020). Lexical semantics in advanced second language French: The acquisition of genericity. *Lingua, 234*, Article 102761. https://doi.org/10.1016/j.lingua.2019.102761

Hua, D., & Lee, T. H.-T. (2005). Chinese ESL learners' understanding of the English count–mass distinction. In Dekydtspotter, L., Sprouse, R. A., & Liljestrand, A. (Eds.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)* (pp. 138–149). Cascadilla Proceedings Project.

Inagaki, S. (2014). Syntax–semantics mappings as a source of difficulty in Japanese speakers' acquisition of the mass–count distinction in English. *Bilingualism: Language and Cognition, 17*(3), 464–477. https://doi.org/10.1017/S136672891300047X

Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2-acquisition: The role of specificity. *Language Acquisition, 12*(1), 3–69. https://doi.org/10.1207/s15327817la1201_2

Ionin, T., Montrul, S., & Crivos, M. (2011). A bidirectional study on the acquisition of plural marked determiners in English and Spanish. *Applied Psycholinguistics, 32*(2), 281–322. https://doi.org/10.1017/S0142716410000410

Ionin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning, 60*(4), 877–925. https://doi.org/10.1111/j.1467-9922.2010.00577.x

Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua, 118*(4), 554–576. https://doi.org/10.1016/j.lingua.2006.11.012

Kim, L. K., & Lakshmanan, U. (2009). The processing role of the Article Choice Parameter: Evidence from L2 learners of English. In García Mayo, M. P., & Hawkins, R. (Eds.), *Second language acquisition of articles* (Vol. 49, pp. 87–113). John Benjamins. https://doi.org/10.1075/lald.49.05kim

Ko, H., Ionin, T., & Wexler, K. (2010). The role of presuppositionality in the second language acquisition of English articles. *Linguistic Inquiry, 41*(2), 213–254. https://doi.org/10.1162/ling.2010.41.2.213

Liu, D., & Gleason, J. (2002). Acquisition of the article *the* by nonnative speakers of English. *Studies in Second Language Acquisition, 24*(1), 1–26. https://doi.org/10.1017/S0272263102001017

Montrul, S., & Ionin, T. (2010). Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition, 13*(4), 449–473. https://doi.org/10.1017/S1366728910000036

Pereltsvaig, A. (2007). On the universality of DP: A view from Russian. *Studia Linguistica, 61*(1), 59–94. https://doi.org/10.1111/j.1467-9582.2007.00130.x

Progovac, L. (1998). Determiner phrase in a language without determiners. *Journal of Linguistics, 34*(1), 165–179. https://doi.org/10.1017/S0022226797007062

Rothstein, S. (2010). Counting and the mass/count distinction. *Journal of Semantics, 27*(3), 343–397. https://doi.org/10.1093/jos/ffq007

Rutkowski, P. (2002). Noun/pronoun asymmetries: Evidence in support of the DP hypothesis in Polish. *Jezikoslovlje, 3*(1–2), 159–170.

Sabir, M. (2019). The L2 acquisition of mass nouns by Arab learners of English. *Advances in Language and Literary Studies, 10*(5), 152–160. https://doi.org/10.7575/aiac.alls.v.10n.5p.152

Snape, N. (2008). Resetting the nominal mapping parameter in L2 English: Definite article use and the count–mass distinction. *Bilingualism: Language and Cognition, 11*(1), 63–79. https://doi.org/10.1017/S1366728907003251

Snape, N., García Mayo, M. P., & Gureli, A. (2013). L1 transfer in article selection for generic reference by Spanish, Turkish, and Japanese L2 learners. *International Journal of English Studies, 13*(1), 1–28. https://doi.org/10.6018/ijes/2013/1/123161

Snape, N. (2018). Definite generic vs. definite unique in L2 acquisition. *Journal of the European Second Language Association, 2*(1), 83–95. https://doi.org/10.22599/jesla.46

Tang, W., Fiorentino, R., & Gabriele, A. (2023). Examining transfer in the acquisition of the count/mass distinction in L2 English. *Second Language Research, 39*(1), 231–257. https://doi.org/10.1177/0267658320947630

Tryzna, M. (2009). Questioning the validity of the Article Choice Parameter and the Fluctuation Hypothesis: Evidence from L2 English article use by L1 Polish and L1 Mandarin Chinese speakers. In García Mayo, M. P., & Hawkins, R. (Eds.), *Second language acquisition of articles* (Vol. 49, pp. 67–86). John Benjamins. https://doi.org/10.1075/lald.49.04try

Yoon, K. K. (1993). Challenging prototype descriptions: Perception of noun countability and indefinite vs. zero article use. *IRAL – International Review of Applied Linguistics in Language Teaching, 31*(4), 269–289. https://doi.org/10.1515/iral.1993.31.4.269

Wu, S. (2025). The acquisition of count-mass distinction of English flexible nouns by Chinese-speaking learners of English. *English Language Teaching and Linguistics Studies, 7*(1), 27–40.

Zlatić, L. (2014). Definiteness and structure of NP in Slavic. In L. Schürcks, A. Giannakidou, & U. Etxeberria (Eds.), *The nominal structure in Slavic and beyond* (pp. 17–38). De Gruyter Mouton. https://doi.org/10.1515/9781614512790.17

Lea Elias Isen* & Kasper Boye
DOI: 10.19090/slm.6.7

Department of Nordic Studies and Linguistics,
University of Copenhagen, Denmark

LANGUAGE PRODUCTION IN PTSD: EVIDENCE FOR FEATURES ASSOCIATED WITH BROCA'S APHASIA IN TWO CASE STUDIES

This pilot study investigates whether individuals with PTSD (IWPs), in general and during flashbacks, exhibit linguistic deficits reminiscent of Broca's aphasia. The study is motivated by two things: 1) research documenting Working Memory (WM) deficits in IWPs as well as lower activity in Broca's area (BA 44-45) during flashbacks specifically; 2) a theory suggesting that Broca's aphasia arises from general resource reductions, likely in working memory capacity (WMC), prompting more economical language production. If this theory is correct, it predicts a similar response in other groups experiencing resource reductions, including IWPs. In the present study, we analyzed the language of two IWPs and two matched controls across three comparisons: (1) IWPs in flashback vs. not in flashback, (2) IWPs not in flashback vs. control, and (3) IWPs in flashback vs. control. We found significant deficits in terms of variables associated with Broca's aphasia when IWPs were compared to controls as well as when compared to themselves while not experiencing flashback. The results support the idea of transient non-fluency during flashback as well as a degree of non-fluency in IWPs in general.

Keywords: post-traumatic stress disorder, non-fluency, Broca's aphasia, cognitive linguistics, compensatory response, cognitive resource reduction, working memory

1. INTRODUCTION

A notable minority of people who experience an event of "extremely threatening or horrific nature" (Koenen et al., 2017) will develop post-traumatic stress disorder (PTSD) (World Health Organization, 2019). Individuals with PTSD (IWPs; sg. IWP) exhibit avoidance behavior, heightened vigilance, and involuntary re-experiencing of the traumatic event (Koenen et al., 2017).

Involuntary re-experiencing, commonly referred to as *flashbacks*, is a dissociative state in which the affected individual feels as if she is reliving the traumatic event in the present (e.g., Brewin, 2003; Ehlers & Clark, 2000; Herman, 2015). Flashbacks are typically triggered by an essentially harmless stimulus that the sufferer associates with the traumatic event (Lanius, 2015), e.g., the smell of gasoline in a car crash survivor. Flashbacks can last anywhere from seconds to, in extreme cases, days (American Psychiatric Association, 2013).

Flashbacks are thought to mirror how the body responds to genuinely life-threatening situations (e.g., Payne et al., 2015; Rainnie & Ressler, 2009). This involves reallocating resources away from slow and resource-intensive higher cognitive processes to prioritize quick threat processing and alertness (e.g., LeDoux, 2012; Pavlícková et al., 2024; Rainnie & Ressler, 2009). If we understand language production as a complex cognitive ability, experiencing a flashback might affect its functionality due to competing

* Corresponding author: Lea Elias Isen; email: isen@hum.ku.dk

cognitive demands. In the case of a genuinely dangerous situation of a limited duration, this response is adaptive (McEwen, 2007). However, when overgeneralized, it becomes more harmful than helpful.

The idea that flashback-related reallocation might have linguistic consequences is supported by brain-imaging studies showing Broca's area (BA 44-45) to receive less blood flow during flashback (e.g., Pitman et al., 2001; Rauch et al., 1996; Shin et al., 1997). Although its particular contribution is controversial, Broca's area remains famous for its role in language production (Rogalsky et al., 2008), as evidenced by the fact that linguistic symptoms known as *Broca's aphasia* have been linked to the area (e.g., Carey, 2017). The link between Broca's area and Broca's aphasia has been called into question by findings showing that individuals with damage to BA 44-45 do not necessarily exhibit Broca's aphasia, and conversely, that those with Broca's aphasia do not always have damage to the area (Andrews et al., 2022; Davis et al., 2008). However, the association between Broca's aphasia and BA 44-45 persists in popular scientific literature. This is evident in how the observation that reduced activity in Broca's area is associated with both PTSD-flashbacks and Broca's aphasia has led to the hypothesis that IWPs during flashback may exhibit a form of "transient aphasia" (Cozolino, 2020, p. 168) or "effects [similar to those] of physical lesions like strokes" (van der Kolk, 2014, p. 43). In further support of this hypothesis, the presence of "unclear speech" (Chessell et al., 2019, p. 11) is used clinically as a marker of a flashback occurring. Thus, reduced blood flow to the region could be interpreted as a manifestation of resources being moved away from mechanisms involved in higher cognitive functions, e.g., language.

Beyond possible linguistic deficits during flashbacks, there is reason to suspect deficits in IWPs outside flashbacks. Some researchers argue that Broca's aphasia is the result of general resource reductions (e.g., Boye et al., 2023; Fedorenko et al., 2023; Kolk, 1995)[5] with some specifically suggesting that it is the result of a reduction in working memory capacity (WMC) (e.g., Blackwell & Bates, 1995, cf. Boye et al., 2023). WMC reductions are also common in the PTSD population (Norte et al., 2024). The idea that WMC reductions could lead to language production issues is in line with the theory that working memory (WM) contributes to our capacity for producing language (e.g., Badecker & Kuminiak, 2007; Deldar et al., 2020). Due to these shared symptoms, the features seen in Broca's aphasia could be a meaningful point of comparison for an investigation of the linguistic patterns in IWPs.

A growing body of research has already examined the patterns of language production in PTSD from a variety of perspectives. Hellawell and Brewin (2004) investigated how language use differs in the retelling of traumatic versus non-traumatic memories. Their findings supported Brewin's (2001, 2003)

[5] This theory contrasts with the interpretation of Broca's aphasia as the result of a specific deficit in syntax (e.g., Friedmann, 2006; Friedmann & Grodzinsky, 1997).

theory that traumatic memories are processed and stored differently from ordinary memories. Likewise, other studies have explored whether the linguistic structure and content of the retellings of traumatic memories reflect their distinct nature (e.g., Miragoli et al., 2019; Römisch et al., 2014; van der Kolk & Fisler, 1995). Still others have investigated whether language content or structure correlates with the presence or severity of PTSD, including potential diagnostic implications (e.g., Gayraud & Auxéméry, 2022; Kleim et al., 2018; Papini et al., 2015). While most focus on content and structure, some studies have investigated articulation specifically (Hu et al., 2024; Marmar et al., 2019), finding a tendency towards slower and more monotonous speech.

However, little attention has been paid specifically to how flashbacks and PTSD in general might impair language production. This is both a disservice to IWPs who might be suffering from unaddressed linguistic difficulties and a loss for the field of linguistic impairment research. The specific hypothesis that linguistic features resembling Broca's aphasia are present in IWPs has been tested only once previously: As part of her MA thesis, Caglar (2012) completed an investigation which yielded no significant results, possibly due to an insufficient selection of variables. Our study[6] seeks to close that gap in the research by investigating the language production of IWPs in two case studies. It seeks to test the hypothesis that IWP speech shows similarities to the speech characteristic of Broca's aphasia, and that it does so both for speech produced by IWPs in flashback and for speech produced by IWPs outside flashback. Our use of Broca's aphasia specifically as a point of comparison is theoretically motivated by the parallels in symptoms, as described above. We supplement the results of the linguistic analysis with the results of a brief qualitative survey in which we ask participants about their experience with language during flashback. The purpose of this is to evaluate whether the linguistic results reflect subjectively felt language difficulties in sufferers. The study at hand is the pilot of a larger project on language production in IWPs named *Language in PTSD* (LIP).

Section 2 outlines our understanding of Broca's aphasia. On this background, it subsequently elaborates on the link between aphasia and PTSD. Section 3 presents our specific hypotheses, and Section 4 the cases. Section 5 deals with methodological issues, Section 6 presents the results, and Section 7 discusses the results. The conclusion is found in Section 8.

2. BROCA'S APHASIA AS A RESPONSE TO RESOURCE REDUCTION

While consensus is lacking on the exact mechanisms underlying Broca's aphasia, it is widely recognized as a language disorder linked to acquired brain damage (Carey, 2017; see Section 1). The

[6] Note that this study builds on the work conducted for Isen's unpublished MA thesis (2025).

disorder is a type of non-fluent aphasia and is characterized by slower speech rate and shorter average utterance length (Menn et al., 1990; for a recent review, see Hasan et al., 2024). Additional features are subject to significant individual and cross-linguistic variation, but may include agrammatism,[7] i.e., a decrease in the conventional use of grammatical elements (including morphemes, words, constructions), and an overall reduced syntactic complexity (e.g., Kolk, 1995; Maher, 2018; Thy et al., 2024). Although traditionally associated with Wernicke's aphasia, reduced lexical diversity has also been observed in people with Broca's aphasia (Arslan et al., 2016; Sang, 2015).

At first glance, it may not seem intuitive to compare disorders as different as PTSD and Broca's aphasia. However, as mentioned in the introduction, both can be understood to involve a reduction in the availability of cognitive resources. According to a theory developed by Boye et al. (2023), some of the linguistic characteristics of Broca's aphasia must, in fact, be understood in part as a compensatory response to a cognitive resource reduction, possibly a reduced WMC.[8]

Specifically, Boye et al. proposed that the sparsity of grammatical elements sometimes seen in Broca's aphasia is an extreme example of what is known as *good-enough production* (Goldberg & Ferreira, 2022). *Good-enough production* refers to the phenomenon of speakers producing utterances that are incomplete, less complex, or generally suboptimal in order to preserve cognitive resources. Despite imperfections, *good-enough utterances* are typically sufficient for successful communication due to listeners being cooperative (Grice, 2002 [1975]) and able to infer meaning from context (Tanenhaus et al., 1995). The view that lower resource availability might lead to aphasia-like language production is further supported by the finding that neurotypical individuals can be induced to produce agrammatic language if put under stress (e.g., Blackwell & Bates, 1995; Ishkhanyan et al., 2019). In this view, Broca's aphasia is not unique in kind but in degree: it represents an extreme utilization of compensatory strategies that all speakers might employ when experiencing excessive cognitive load.

The theory mentioned above provides a theoretical link between possible linguistic features characteristic of PTSD and features characteristic of Broca's aphasia, as it allows us to see them as resulting from the same basic phenomenon, viz. resource reduction. In the case of Broca's aphasia, the resource reduction is due to brain damage. In the case of PTSD flashbacks, we hypothesize that the reduction is due to a reallocation of cognitive resources to better accommodate perceived acute survival

[7] While this feature of aphasic speech has received ample attention in the literature, not all individuals with Broca's aphasia display agrammatism (see e.g., Berndt, 1987). Additionally, omission of grammatical elements might not be a feature of Broca's aphasia at all in speakers of certain languages, e.g., Bengali (Bose, 2025) and Kalaallisut (Nedegaard et al., 2020).
[8] Multiple studies have found reduced WMC in individuals with aphasia (e.g., Caspari et al., 1998; Mayer & Murray, 2012). Therefore, WMC is a plausible candidate for the particular resource in deficit.

needs. The WMC reduction found in IWPs more broadly might likewise result in linguistic difficulties even outside of flashbacks, although possibly to a lesser extent. In short, if Broca's aphasia is due to resource reductions, we would expect to see features in the language of IWPs which are parallel to those found in Broca's aphasia.

2.1 Specific features of Broca's aphasia as compensatory responses

If a speaker has fewer resources available than what typical language production requires, she must adjust her language production to accommodate it. The basic compensatory response gives priority to those elements that convey the communicatively essential points. As a result of this prioritization, less essential elements may be left out entirely, reduced, or substituted with others. The utterance in (1) is an example of language from an individual with aphasia (IWA).[9] Assuming that the utterance represents an alternative to the utterance *the stepmother is really ugly*, we might interpret the lack of both copula and article as the result of a prioritization in response to resource limitations:

(1) "Stepmother … really ugly" (Thompson & Bastiaanse, 2012)

Now, what governs which elements are retained, and which ones are 'left out'? According to Boye et al. (2023), attentionally foregrounded parts of messages tend to be retained, while backgrounded parts tend to be omitted, reduced, or substituted. In light of the theory developed in Boye and Harder (2012), this explains why grammatical elements tend to be omitted, reduced or substituted. In this theory, grammatical elements – e.g., *the* and *is* – are defined as being attentionally backgrounded (or, discourse secondary) and thus as dependent on combination with other words, while lexical words – e.g., *stepmother* and *ugly* – are defined as having the potential to be foregrounded (discourse primary). In a speaker with a reduction in available cognitive resources – e.g., IWA – it therefore makes sense that we specifically see a reduction of grammatical words as opposed to lexical ones, since the information loss resulting from 'omitting' grammatical words is relatively benign and thus a more effective use of resources. Note now that if the speaker who produced (1) 'left out' *stepmother* and *ugly* instead of the grammatical words – leaving only *the is really* – the main point of the utterance is lost. A similar point was made by Fedorenko et al. (2023), who went as far as calling agrammatism a rational strategy in the face of resource deficits.

[9] While this abbreviation can be used for individuals with other types of aphasia, we will exclusively use it for references to individuals with Broca's aphasia in this paper.

Leaving out elements inherently leads to shorter utterances, and shorter average utterance length is an indicator of lower complexity (Szmrecsányi, 2008). Shorter average utterance length is also a defining feature of Broca's aphasia (Menn et al., 1990). In addition, grammatical elements are often obligatory, for which reason their omission can be described as a type of error, i.e., a deviation from conventions. Thus, lowering complexity may be associated with higher error rates.

This brings us to another possible means of adjusting resource expenditure: speech rate. Coupé et al. (2019) measured the information rate (bits/s) in language production and found that while speech rate varies, bits/s remains stable across languages (~39 bits/s). Based on this, they argued that the typical speech rate of a given language is determined by the typical information processing speed of the human brain. If this assertion is correct, a slower speech rate could be a way of lowering the bits/s, thus distributing production costs over a larger span of time. This might manifest as individual words being pronounced more slowly or as a higher production rate of disfluencies. Both of these responses create more time between information-carrying units, as exemplified in (2).

(2) "(uhm) (well) (uhm) worked (uhm) autodesk (uhm) [/] sales <sales> (uhm) worldwide" (tactustherapy, 2017)

Finally, lowering lexical diversity might represent a way of preserving resources by relying on more entrenched vocabulary, thus possibly reducing retrieval costs (e.g., Steinkrauss & Schmid, 2016; Hatchard, 2015).

3. HYPOTHESES

The previous section introduced the features we are looking for in our two cases. We will be investigating whether these features can be found in our two cases (see Section 4). This brings us to the specific hypotheses we are testing. These are seen in Table 1.

Table 1. Hypotheses.

H1	The language of an IWP experiencing flashback shows features associated with Broca's aphasia when compared to the language of an IWP not experiencing a flashback.
H2	The language of an IWP not experiencing a flashback shows features associated with Broca's aphasia when compared to the language of a neurotypical control.
H3	The language of an IWP experiencing flashback shows features associated with Broca's aphasia when compared to the language of a neurotypical control.

As summarized in Table 2, the hypotheses concern relative differences between individual IWPs and either themselves in a non-flashback state (H1) or neurotypical matched controls (H2, H3). Testing

H1 will shed light on the linguistic consequences of the presence of a flashback, flashback being the independent variable, while H2 concerns the effect of PTSD without flashback, PTSD being the independent variable. Finally, H3 compares the flashback condition directly with the neurotypical control condition. It needs to be emphasized that all hypotheses concern *relative similarity to Broca's aphasia*: flashback relative to non-flashback, non-flashback PTSD relative to neurotypicality, and flashback relative to neurotypicality.

Table 2. Conditions.

	Focal condition	Control
H1	IWPs during a flashback (+flashback/+PTSD)	IWPs in non-flashback state (–flashback/+PTSD)
H2	IWPs in non-flashback state (–flashback/+PTSD)	Control without PTSD (–flashback/–PTSD)
H3	IWPs in flashback (+flashback/+PTSD)	Control without PTSD (–flashback/–PTSD)

As an example of how these hypotheses would be supported, consider H1. Support for H1 is found if, for instance, the median utterance length or median speech rate is lower in IWPs' speech during a flashback than in IWPs outside a flashback.

4. DATA

In the following section, the data collection procedure as well as the collected data is described. The procedure received IRB approval from *The Research Ethics Committee of the Faculty of Humanities and the Faculty of Law at University of Copenhagen* and all participants provided informed consent.*4.1 Data collection, IWPs*

IWPs were recruited both via Danish NGOs supporting the population specifically[10] and those serving populations with a high incidence rate of PTSD.[11] Inclusion criteria for IWPs were (a) being a native speaker of Danish, and (b) having received a formal PTSD diagnosis.

The data collection method was developed specifically for this study. In the design, we prioritized participants' mental safety over data quality when the two were at odds. IWPs were provided with a definition of flashbacks and then had to record themselves during flashbacks up to three times over the span of one month (used as the focus condition for H1 and H3). Additionally, the IWPs had to record themselves up to three times while in a neutral or positive mood (used as the control for H1 and the focus for H2).

[10] *Landsforeningen for PTSD-ramte og pårørende* (IWPs and their loved ones) and *Angst & StressForeningen* (individuals suffering from anxiety and/or stress).
[11] *Landsforeningen Spor* (adult victims of childhood sexual abuse), *Søstre mod vold og kontrol* (victims of gendered violence), and *Voldtægtsofres Vilkår* (victims of rape).

We decided against on-site recording by the researchers, as this would have necessitated the ethically dubious task of having to trigger flashbacks in the participants (cf. Pitman et al., 2001; Rauch et al., 1996; Shin et al., 1997). Having the participants record themselves instead allowed us to capture flashbacks they would have experienced regardless of their participation in the study.

Each recording had to be between 2 and 5 minutes as we estimated that this would provide us with a sample of at least 300 words per participant per condition (see Prins & Bastiaanse, 2004) without overwhelming the participants with the scale of the task, thus possibly increasing the risk of participant dropout.

In the recordings, participants describe the current day in detail without discussing the traumatic memory itself. This topic was chosen because it closely resembles grounding exercises used in PTSD treatment (e.g., Rothschild, 2021), thus being judged safer for participants. In addition, due to the transient nature of flashbacks, asking the participants to perform a more structured task, e.g., picture description, was judged impractical: by the time a participant had located the elicitation material, their flashback state might have passed.

After the collection of language samples, IWPs had to fill out a questionnaire on demographic information, details about PTSD diagnosis and treatment history, and their subjective experience of linguistic difficulty.

4.2 Data collection, controls

The controls (non-IWPs) were recruited subsequent to IWP-recruitment for the sake of matching gender and age. Inclusion was additionally based on (a) being a native speaker of Danish, and (b) not being diagnosed with PTSD. Non-IWPs were asked to complete the same task as the IWPs, but only once for 3-5 minutes while in a neutral or positive mood (used as controls for H2 and H3). This was decided to approximate the lengths of recordings provided by IWPs. Non-IWPs were additionally asked to report their gender, age, and PTSD status.

4.3 Data description

Two IWPs (M35-39, F60-64) completed data collection. M35-39 will henceforth be referred to as Peter, F60-64 as Pia. Pia's control will be referred to as Carla, Peter's control as Chris. The details of the data collected can be seen in Table 3.

Table 3. Data.

	+flashback/+PTSD		−flashback/+PTSD		−flashback/−PTSD	
	Peter[12]	Pia	Peter	Pia	Chris	Carla
Duration	00:02:17	00:03:57	00:01:59	00:03:16	00:03:43	00:03:55
Length	150	354	193	480	298	554

5. METHOD OF ANALYSIS

Our method of analysis is pairwise comparison. For non-categorical variables, the Mann-Whitney U-test (Mann & Whitney, 1947) is applied; for categorical variables, Fisher's Exact Test (Fisher, 1922). The specific tests were chosen for their reliability in the case of small sample sizes (Morgan, 2017), e.g., in case studies. We selected our dependent variables based on literature describing Broca's aphasia (see Section 1.2). The variables are introduced in the following subsection.

5.1 Dependent variables and their relation to Broca's aphasia

Utterance length is measured by median length of utterance (MdLU).[13] In our definition, an utterance can either consist of a main clause (including 0 or more subordinate clauses), or a stand-alone phrase, e.g., NP or VP. In order to relatively resemble Broca's aphasia, utterances must be shorter on average.

We measure speech rate in both words per minute (WPM) and the rate of disfluencies. In order to relatively resemble Broca's aphasia, production must contain a higher rate of disfluencies and lower WPM.

Grammatical error rate is measured firstly by frequency of errors (insertion, deletion, and substitution) (Dulay et al., 1982) and secondly by frequency of topic drop (e.g., Nygard, 2018). Topic drop is a type of deviation from convention in Danish. However, we distinguish between this type of deviation and genuine errors as topic drop is relatively common in informal registers (Jensen, 2015). Proportion of definite articles (DEF) (rel. to nouns) is included as a third measurement of error rate.

Complexity is measured by comparing proportions of parts of speech where one category is more discourse prominent in Danish. These include proportion of subordinate clauses (SUBR) (rel. to main clauses, MAIN), transitive verbs (TR) (rel. to intransitive verbs, INTR), and auxiliary verbs (AUX) (rel. to main verbs, MV). Complexity is additionally measured by contrasting the frequency of nouns (N)

[12] Note that Peter only managed to make one recording per condition, meaning that in his number of words per condition did not reach 300.

[13] Mean length of utterance (MLU) is more commonly used for this purpose. However, we decided to use MdLU because medians are less sensitive to extreme outliers – a common issue with smaller samples.

with verbs (V), articles (ART), and anaphoric pronouns (ANA), given that nouns are typically discourse primary.

Finally, lexical diversity is measured by Type–Token Ratio (TTR). Both lexical and grammatical words are included in this count, while disfluencies are excluded. An overview of all variables can be found in Table 4.

Table 4. Linguistic features associated with Broca's aphasia and the variables representing them.

Linguistic feature	Variable
Shorter average utterance length	• Median length of utterance (MLU)
Slower speech rate	• Words per minute (WPM) • Ratio of disfluencies to words
Higher grammatical error rate	• Ratio of grammatical errors to words • Ratio of topic drops to words • Ratio of DEF to N
Lower complexity	• Prop. of SUBR • Prop. of TR • Prop. of AUX • Ratio of V to N • Ratio of ART to N • Ratio of ANA to N
Lower lexical variety	• Ratio of types to tokens

6. RESULTS

When results are presented in terms of p-values (Table 5 and Table 7), red denotes a non-significant result, darker green a result significant at the Bonferroni-corrected significance level ($\alpha_{corrected} = \alpha/13 = 0.003846154$), and lighter green a result significant at the nominal significance level ($\alpha = 0.05$).[14] For tables displaying results in terms of medians and proportions (Table 6 and Table 8), light blue denotes that the +flashback condition is significantly different from the −PTSD condition but not from the −flashback/+PTSD condition, dark blue that it is both significantly different from −PTSD and from +flashback/−PTSD. Purple denotes that the −flashback/+PTSD condition differs significantly from −PTSD.

[14] See e.g., VanderWeele and Mathur (2019) for a discussion of when Bonferroni-correction might be too conservative as well as benefits of including both Bonferroni-corrected and nominal significance level.

6.1 Peter

With vs. without PTSD

WPM (U = 0, z = -2.51, p = 0.0060), rate of topic drop (p = 0.0174), and TTR (p = 0.0424)[15] are significantly different from the neurotypical control at the nominal level. Disfluency rate is significantly different at the corrected level (p = 0.0007). This suggests Peter speaks more slowly than Chris, both in terms of disfluency rate (26.94 per 100 words) and WPM (Mdn = 93.6). He additionally produces more deviations of the topic drop type (4.15 per 100 words).

In vs. outside of flashback

When the speech of Peter during flashback is compared to his non-flashback speech, only WPM is significantly different (U = 4, z = -1.67, p = 0.0475) and only at the nominal level. This indicates slower production (Mdn = 62.1) during flashback.

In flashback vs. without PTSD

When Peter in flashback is compared to Chris, WPM (U = 0, z = -2.51, p = 0.0060), disfluency rate (p = 0.0039), and TTR (p = 0.0493) are significantly different at the nominal level. MdLU (U = 405, z = 2.84, p = 0.0023) is significantly different at the corrected level. This means that Peter in flashback produces shorter utterances (Mdn = 6.00). This indicates that his utterances are also less complex. In addition, his speech production appears to be slower, both in terms of disfluencies and WPM.

Table 5. Results, p-values, Peter and Chris.

	With vs. without PTSD	In vs. outside of flashback	In flashback vs. without PTSD
MdLU	0.0571	0.1314	0.0023
WPM	0.0060	0.0475	0.0060
Disfluency	0.0007	0.8124	0.0039
Errors	0.7590	0.3816	0.1437
Topic drop	0.0174	0.1976	0.6054
DEF	0.7153	0.6416	0.7144
SUBR	0.7767	1.0000	0.5412
TR	0.8515	0.0814	1.0000
VB	0.3843	0.8535	0.7459
ART	0.4906	1.0000	0.4239
ANA	0.2530	1.0000	0.3031
AUX	0.6090	1.0000	0.7801
TTR	0.0424	0.9257	0.0493

[15] At first glance, Peter's language appears to be less lexically diverse (51.81 types per 100 tokens). However, this is most likely due to a mismatch in sample size lengths, which TTR is highly sensitive too (Wright et al., 2003). In future studies, this could be mitigated either by matching sample sizes or by using Moving-Average Type–Token Ratio (MATTR) (Covington & McFall, 2010) instead of a classic TTR.

Table 6. Results, proportions and medians, Peter and Chris.

	+PTSD/+flashback	+PTSD/−flashback	−PTSD
Median length of utterance	6.0	7.0	8.0
Median WPM	62.1	93.6	147.4
Disfluencies per 100 words	25.3	26.9	12.1
Errors per 100 words	4.7	2.6	2.0
Topic drop per 100 words	1.3	4.1	0.7
DEF per 100 N	14.3	6.7	11.3
SUBR per 100 MAIN	21.1	27.3	34.5
TR per 100 INTR	28.0	73.9	67.2
V per 100 N	152.4	133.3	172.6
ART per 100 N	19.0	23.3	35.5
ANA per 100 N	19.0	20.0	37.1
AUX per 100 MV	18.5	22.9	16.3
Types per 100 tokens	50.7	51.8	70.8

6.2 Pia

With vs. without PTSD

When Pia is compared to Carla, both MdLU (U = 1374.5, z = -1.83, p = 0.0336) and proportion of subordinate clauses (p = 0.0245) are significantly reduced at the nominal level. The proportion of topic drops (p = 0.0005) is significantly different at the corrected level. This suggests that Pia produces shorter and thus less complex utterances (Mdn = 6.5). Additionally, her utterances appear less complex in terms of subordination (17.86 per 100 MAIN) and contain a higher rate of a benign type of deviation from convention (2.71 topic drops per 100 words).

In vs. outside of flashback

Pia's speech rate (U = 2, z = -2.48, p = 0.0066) and error rate (p = 0.0108) both significantly differ from her non-flashback condition at the nominal significance level. MdLU (p = 0.0009) and disfluency rate (p = 0.0004) differ significantly at the corrected level. This suggests Pia speaks more slowly during flashbacks, both in terms of WPM (Mdn = 91.7) and rate of disfluency production (18.08 per 100 words). Additionally, her utterances appear to contain more errors (3.39 per 100 words) and to be shorter (Mdn = 5.0) and thus less complex.

In flashback vs. without PTSD

When Pia in flashback is compared to Carla, she has a significantly higher error rate (p = 0.0112), disfluency rate (p = 0.0113) and proportion of subordinate clauses (p = 0.0216) at the nominal

significance level. MdLU (U = 907, z = -4.57, p = 1.00E-5), WPM (U = 0, z = p = 0.0026) and topic drop rate (p = 3.793E-6) are significantly different at the corrected level. This indicates that Pia in flashback produces utterances that are less complex in terms of both subordination and length while containing more deviations from convention, both topic drops and errors of the general type. Additionally, her speech is slower.

Table 7. Results, p-values, Pia and Carla.

	With vs. without PTSD	In vs. outside of flashback	In flashback vs. without PTSD
MdLU	0.0336	0.0009	1.00E-05
WPM	0.3446	0.0066	0.0026
Disfluency	0.2525	0.0004	0.0113
Errors	1.0000	0.0108	0.0112
Topic drop	0.0005	0.1858	3.79E-06
DEF	1.0000	0.7687	0.5930
SUBR	0.0245	1.0000	0.0216
TR	0.1128	0.1564	1.0000
VB	0.4453	0.3626	0.7779
ART	0.3019	0.7078	0.7450
ANA	0.5036	0.4231	0.8841
AUX	0.7909	0.5145	0.3107
TTR	0.9077	0.8969	0.8008

Table 8. Results, proportions and medians, Pia and Carla.

	+PTSD/+flashback	+PTSD/−flashback	−PTSD
Median length of utterance	5.0	6.5	9.0
Median WPM	91.7	154.9	145.0
Disfluencies per 100 words	18.1	8.5	11.0
Errors per 100 words	3.4	0.8	0.9
Topic drops per 100 words	4.5	2.7	0.2
DEF per 100 N	6.3	9.3	9.4
SUBR per 100 MAIN	17.9	19.0	47.4
TR per 100 INTR	27.2	42.2	28.3
V per 100 N	182.8	149.5	171.0
ART per 100 N	25.0	20.6	29.0
ANA per 100 N	37.5	28.9	35.5
AUX per 100 MV	18.2	22.9	25.5
Types per 100 tokens	41.8	40.8	40.3

7. DISCUSSION

This study is only a pilot study and has several major methodological limitations. Before we discuss the possible interpretations of the study, the limitations should be noted. Firstly, the study consists of only two case studies, meaning that the results cannot be generalized. Secondly, the match between IWPs and non-IWP controls is less than ideal. Thirdly, the fact that sample sizes vary might skew the results, especially in the case of variables known to be sensitive to sample size, e.g., TTR (Wright et al., 2003). These factors should be taken into account when interpreting the results and in the design of future studies by LIP and others.

7.1 Interpreting the results

We aimed to test three hypotheses, specifically whether the language produced by our IWPs (Peter and Pia) during flashback and outside flashback showed features associated with Broca's aphasia when compared to respectively the language produced by the IWPs outside flashback (H1) and the language produced by their neurotypical controls (H2; H3).

Our results provide support for all three hypotheses. The support is found in both case studies. It is however found only for some of the variables investigated, and for distinct sets of variables depending on speakers and hypotheses. Specifically, we find a slower speech rate and shorter utterances, i.e., the defining features of Broca's aphasia, but we do not find evidence of agrammatism. In summary, the speech in IWPs can be said to be relatively non-fluent but not agrammatic, and the parallel to Broca's aphasia is thus only partial.

Consider again H1: *The language of an IWP experiencing flashback shows features associated with Broca's aphasia when compared to the language of an IWP not experiencing a flashback*. The support for H1 might be explained by our theory that flashbacks lead to resource reallocation, which in turn leads to fewer resources available for language production (see Section 1).

Now, consider H2: *The language of an IWP not experiencing a flashback shows features associated with Broca's aphasia when compared to the language of a neurotypical control*. The support for this hypothesis could be interpreted as suggesting that the WMC reductions often seen in PTSD (see Section 1) have negative consequences for language production.

Lastly, consider H3: *The language of an IWP experiencing flashback shows features associated with Broca's aphasia when compared to the language of a neurotypical control*. This hypothesis was naturally supported for more variables than H1 and H2. In accordance with the interpretations suggested

for H1 and H2, the support for H3 may be interpreted as being due to an accumulation of the effects of temporary resource allocation and of a general deficit in WMC.

7.2 Subjective experience of IWPs

To learn whether the quantitative analysis might reflect subjectively felt impairments in the IWPs, we asked the IWPs about their self-perceived linguistic difficulties (see Section 3.1). Specifically, we asked: 'Are you yourself aware of having experienced language difficulties during flashbacks?', and: 'If yes, which ones?'.

Peter describes how his 'vocabulary is very small', which possibly reflects a deficit in lexical retrieval, which was not pronounced enough to show up in the quantitative analysis. Additionally, Peter describes how he 'has a hard time putting sentences together'. This could reflect reductions in his resource availability, resulting in subthreshold effects on his grammatical production. Pia also describes experiencing linguistic difficulties during flashbacks. She specifically describes a spectrum where she feels that her language either 'becomes faltering' or 'disappears'. This might reflect the quantitative finding that she produces slower, more disfluent speech in flashback.

Overall, the answers support an interpretation of the quantitative results as possibly reflecting an actual, subjectively felt impairment during flashbacks (World Health Organization, 2001) with negative linguistic consequences, not simply a neutral difference in language production.

7.3 Interpreting the results based on qualitative similarity to Broca's aphasia

Our results suggest that the presence of flashbacks and PTSD respectively appears to influence language production in a way which is in part parallel to the presence of Broca's aphasia. We did not include a control group of IWAs, and quantitative similarity can thus not be determined. However, we can say that the speech of IWPs and IWAs share some qualitative similarities.

Keeping this similarity in mind, we may begin to understand our results in terms of interpretations of Broca's aphasia. As stated in Sections 1–2, researchers have argued that features of Broca's aphasia may be understood as the result of a compensatory response to resource reduction (Boye et al., 2023) or as a rational strategy in the face of resource deficits (Fedorenko et al., 2023). This interpretation entails that features characteristic of Broca's aphasia may also be present in other types of language conditioned by resource reductions, e.g., the language of IWPs. Additionally, some of the features are also seen in speakers with disorders other than Broca's aphasia, including disorders without structural damage to Broca's area, e.g., slower speech and shorter utterances in individuals with frontotemporal dementia

(Coppieters et al., 2024) or those with a developmental language disorder (DLD) (Wilder & Redmond, 2022).

As predicted by our theory (see Section 1), several features of Broca's aphasia were found in the language production of our IWPs, including the defining features, i.e., slower speech rate and shorter utterances. This and other evidence (see Section 2) considered, we suggest interpreting the results and the qualitative similarity to Broca's aphasia as the consequence of cognitive resource reductions which are not specific to language. WMC could be the specific resource which is reduced in Broca's aphasia as well as PTSD, given that WMC reductions are often seen in IWPs (Norte et al., 2024) as well as in IWAs (e.g., Caspari et al., 1998; Mayer & Murray, 2012). This is, however, purely speculative as we did not dissociate WMC deficits from other possible cognitive consequences of PTSD in our design.

Other features of Broca's aphasia were not present, chiefly those representing reduced grammatical complexity. For instance, while Pia's language production contains a reduced proportion of subordinate clauses compared to Carla's – one variable representing complexity (see Section 5.1) – it does not display a reduced proportion of auxiliary verbs – a second variable representing complexity. In other words, we did not find convincing evidence of agrammatism. At first glance, this might be taken to speak against our interpretation of Broca's aphasia and PTSD as sharing the underlying cause, e.g., a general resource reduction. An interpretation based on the syntactic deficit theory of Broca's aphasia would be that grammatical knowledge is spared in IWPs while it is damaged in IWAs with agrammatism, implying differences in the underlying processes of the disorders. However, an alternative interpretation based on the compensatory response theory would be that agrammatism is only present in the case of more severe resource deficits, i.e., those seen in IWAs with agrammatism but not in IWPs. This theory might be investigated in controlled experiments in which neurotypical speakers are exposed to varying degrees of cognitive load.

This study was first and foremost motivated by the finding that flashbacks are associated with a decrease in blood flow to Broca's area (e.g., Pitman et al., 2001; Rauch et al., 1996; Shin et al., 1997) and claims that this might cause deficits in linguistic production parallel to Broca's aphasia (e.g., Cozolino, 2020; van der Kolk, 2014). In addition, we argued that a parallel might even be found between the speech of IWAs and IWPs not in flashback, given the shared symptom of reduced WMC.

Our results support, if only partly, the idea that flashbacks might lead to deficits in linguistic production that are qualitatively similar to Broca's aphasia. We hypothesize that this is due to the reduction in blood flow to Broca's area during flashback, which might in turn be a manifestation of temporary resource reallocation. Our results also support, though again only in part, the idea that the

presence of PTSD might lead to deficits in linguistic production that are qualitatively similar to Broca's aphasia. We hypothesize that this is due to the WMC deficits associated with PTSD.

However, our study does not allow us to determine the exact chain of causality between the independent variables (PTSD and flashbacks) and the linguistic features of Broca's aphasia, i.e. the dependent variables. In particular, our suggestion that the relative increase in features associated with Broca's aphasia during flashback is due to reduced blood flow to Broca's area is put into question by the incomplete link between damage to the area and Broca's aphasia (see Section 1). Thus, while we argue that the qualitative similarity between the speech of IWAs and IWPs likely has to do with reduced resource availability in both disorders, the role of Broca's area specifically remains unclear.

Moreover, because we only find a partial qualitative parallelism between the linguistic production in IWAs and our IWPs, we cannot entirely reject the possibility that the linguistic deficits seen in the disorders might be caused by fully distinct processes.

Our study cannot establish causality, but if it is in fact the case that flashbacks – and to a lesser extent the long-term resource reductions found in PTSD – cause deficits in language production similar to those seen in Broca's aphasia, then it aligns with our view of Broca's aphasia as a syndrome that must be understood as a compensatory response to resource reduction, and which does not differ qualitatively, but may differ in terms of degree, from other cases of language produced under resource reduction. Thus, our study adds to the list of groups without structural damage to Broca's area (e.g., those with DLD, see above) who display the defining features of Broca's aphasia – slower speech rate and shorter utterances – thus increasing the burden of proof for those who argue that Broca's aphasia represents a distinct, specifically linguistic deficit and that this deficit points towards Broca's area serving specific syntactical functions (e.g., Friedmann, 2006; Friedmann & Grodzinsky, 1997).

8. CONCLUSION

Our hypotheses were based on an understanding of Broca's aphasia as the extreme of a commonplace phenomenon, i.e., linguistic reductions in response to resource reductions (Boye et al., 2023). This understanding entails that aphasia-like language production may also be found in populations experiencing resource reduction without having brain damage, e.g., IWPs. We found partial support for all three hypotheses. Firstly, language produced during flashback shows features associated with Broca's aphasia when compared to the language of an IWP not experiencing a flashback (H1). Secondly, language produced by an IWP not experiencing a flashback shows features associated with Broca's aphasia when compared to the language of a neurotypical control (H2). Thirdly, language produced during flashback

shows features associated with Broca's aphasia when compared to the language of a neurotypical control (H3). However, with one exception, features associated specifically with agrammatism were not found. What we did find was mainly those features associated with non-fluency more broadly. On this basis, the idea of IWPs suffering from "transient aphasia" during flashbacks (Cozolino, 2020, p. 168; cf. van der Kolk, 2014) is not unfounded, but its focus on aphasia is misleading. Instead, we might say that IWPs suffer from *transient relative non-fluency* during flashbacks.

In addition, it should be noted that while our study documents relative similarities between the linguistic response in IWPs and IWAs, it does not and cannot establish causality. However, given that both PTSD and Broca's aphasia entail resource reductions, we consider the results to most likely reflect lower resource availability in our focus conditions relative to control conditions. More research is needed to confirm this explanation. However, if this interpretation is correct, what we find in IWPs is not a linguistic impairment per se, but a general cognitive impairment with possible linguistic consequences (see Section 7.3). The findings overall suggest language production as a meaningful lens for understanding cognitive impairments in PTSD.

In addition to evidence of relative non-fluency in IWPs, we also found indications of a subjective experience of linguistic impairment in our cases (see Section 7.2). Given the fundamental role of language in everyday interactions, even slight language difficulties may negatively influence an individual's daily functioning and ability to participate in public life on equal footing with the non-impaired population. This knowledge might inform the design of inclusion measures. Additionally, documentation of negative linguistic consequences of PTSD might inform psychoeducation, supporting the self-understanding of patients. However, further research with a larger, more diverse sample and a more comprehensive survey of IWPs' ability to meet their communicative needs is required to draw firm conclusions about both the nature of language in PTSD and the subjective experience of sufferers.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Fifth edition.). American Psychiatric Association.

Andrews, J. P., Cahn, N., Speidel, B. A., Chung, J. E., Levy, D. F., Wilson, S. M., Berger, M. S., & Chang, E. F. (2022). Dissociation of Broca's area from Broca's aphasia in patients undergoing neurosurgical resections. *Journal of Neurosurgery*, *138*(3), 847–857. https://doi.org/10.3171/2022.6.JNS2297

Arslan, S., Bamyacı, E., & Bastiaanse, R. (2016). A characterization of verb use in Turkish agrammatic narrative speech. *Clinical Linguistics & Phonetics*, *30*(6), 449–469. https://doi.org/10.3109/02699206.2016.1144224

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language*, *56*(1), 65–85. https://doi.org/10.1016/j.jml.2006.08.004

Blackwell, A., & Bates, E. (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience*, *7*(2), 228–257. https://doi.org/10.1162/jocn.1995.7.2.228

Berndt, R. S. (1987). Symptom co-occurrence and dissociation in the interpretation of agrammatism. In M. Coltheart, R. Job, & G. Sartori (Eds.), *The Cognitive Neuropsychology of Language (Psychology Revivals)* (pp. 221–233). Psychology Press.

Bose, A. (2025, September). *Cross-linguistic variation in agrammatic manifestation in post-stroke aphasia: Findings from South Asian languages*. Science of Aphasia, University of Copenhagen.

Boye, K., & Harder, P. (2012). A usage-based theory of grammatical status and grammaticalization. *Language*, *88*(1), 1–44. https://doi.org/10.1353/lan.2012.0020

Boye, K., Bastiaanse, R., Harder, P., & Martínez-Ferreiro, S. (2023). Agrammatism in a usage-based theory of grammatical status: Impaired combinatorics, compensatory prioritization, or both? *Journal of Neurolinguistics*, *65*, 101108. https://doi.org/10.1016/j.jneuroling.2022.101108

Brewin, C. R. (2001). Memory processes in post-traumatic stress disorder. *International Review of Psychiatry*, *13*(3), 159–163. https://doi.org/10.1080/09540260120074019

Brewin, C. R. (2003). *Posttraumatic stress disorder: Malady or myth?* Yale University Press.

Caglar, L. R. (2012). *Traumatized Broca's Area A linguistic analysis of speech in posttraumatic stress disorder* [Unpublished MA thesis, Utrecht University]. https://studenttheses.uu.nl/handle/20.500.12932/12173

Carey, D. P. (2017). Broca's and Wernicke's areas. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science*. Springer International Publishing. https://doi.org/10.1007/978-3-319-16999-6_3339-1

Caspari, I., Parkinson, S. R., LaPointe, L. L., & Katz, R. C. (1998). Working memory and aphasia. *Brain and Cognition*, *37*(2), 205–223. https://doi.org/10.1006/brcg.1997.0970

Chessell, Z. J., Brady, F., Akbar, S., Stevens, A., & Young, K. (2019). A protocol for managing dissociative symptoms in refugee populations. *The Cognitive Behaviour Therapist*, *12*, e27. https://doi.org/10.1017/S1754470X19000114

Coppieters, R., Bouzigues, A., Jiskoot, L., Montembeault, M., Tee, B. L., Rohrer, J. D., Bruffaerts, R., Russell, L. L., Foster, P. H., Ferry-Bolder, E., Convery, R., Bocchetta, M., Cash, D., Goldsmith, S., Samra, K., Thomas, D. L., van Swieten, J. C., Seelaar, H., Poos, J., … Korhonen, V. O. (2024). A systematic review of the quantitative markers of speech and language of the frontotemporal degeneration spectrum and their potential for cross-linguistic implementation. *Neuroscience & Biobehavioral Reviews*, *167*, 105909. https://doi.org/10.1016/j.neubiorev.2024.105909

Coupé, C., Oh, Y. M., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, *5*(9), eaaw2594. https://doi.org/10.1126/sciadv.aaw2594

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). *Journal of Quantitative Linguistics*, *17*(2), 94–100. https://doi.org/10.1080/09296171003643098

Cozolino, L. (2020). *The pocket guide to neuroscience for clinicians*. W. W. Norton & Company.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language Two*. Oxford University Press.

Davis, C., Kleinman, J. T., Newhart, M., Gingis, L., Pawlak, M., & Hillis, A. E. (2008). Speech and language functions that require a functioning Broca's area. *Brain and Language*, *105*(1), 50–58. https://doi.org/10.1016/j.bandl.2008.01.012

Deldar, Z., Gevers-Montoro, C., Khatibi, A., & Ghazi-Saidi, L. (2020). The interaction between language and working memory: A systematic review of fMRI studies in the past two decades. *AIMS Neuroscience*, *8*(1), 1. https://doi.org/10.3934/Neuroscience.2021001

Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, *38*(4), 319–345. https://doi.org/10.1016/s0005-7967(99)00123-0

Fedorenko, E., Ryskin, R., & Gibson, E. (2023). Agrammatic output in non-fluent, including Broca's, aphasia as a rational behavior. *Aphasiology*, *37*(12), 1981–2000. https://doi.org/10.1080/02687038.2022.2143233

Friedmann, N. (2006). Speech production in Broca's agrammatic aphasia: Syntactic tree pruning. In Y. Grodzinsky & K. Amunts (Eds.), *Broca's Region* (pp. 63–82). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195177640.003.0005

Friedmann, N., & Grodzinsky, Y. (1997). Tense and agreement in agrammatic production: Pruning the syntactic tree. *Brain and Language*, *56*(3), 397–425. https://doi.org/10.1006/brln.1997.1795

Fisher, R. A. (1922). On the interpretation of χ 2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society*, *85*(1), 87–94. https://doi.org/10.2307/2340521

Gayraud, F., & Auxéméry, Y. (2022). Identification of the marks of psychic trauma in spoken language: Definition of the "SPLIT-10" diagnostic scale. *Annales Médico-Psychologiques, Revue Psychiatrique*, *180*(3), 195–212. https://doi.org/10.1016/j.amp.2021.09.016

Goldberg, A. E., & Ferreira, F. (2022). Good-enough language production. *Trends in Cognitive Sciences*, *26*(4), 300–311. https://doi.org/10.1016/j.tics.2022.01.005

Grice, H. P. (2002). Logic and conversation. In D. J. Levitin (Ed.), *Foundations of cognitive psychology: Core readings* (pp. 719–732). MIT Press.

Hasan, F., Kalra, N., Siddique, N. A., Kumar, P., & Kumar, V. (2024). Broca's aphasia: A detailed review of its features, origins, diagnosis, and management. *International Journal of Chemical Studies*, *12*(5), 08–14.

Hatchard, R. (2015). *A construction-based approach to spoken language in aphasia* [Doctoral dissertation, University of Sheffield]. https://etheses.whiterose.ac.uk/10385/

Hellawell, S. J., & Brewin, C. R. (2004). A comparison of flashbacks and ordinary autobiographical memories of trauma: Content and language. *Behaviour Research and Therapy*, *42*(1), 1–12. https://doi.org/10.1016/S0005-7967(03)00088-3

Herman, J. (2015). *Trauma and recovery: The aftermath of violence*. Basic Books.

Hu, J., Zhao, C., Shi, C., Zhao, Z., & Ren, Z. (2024). Speech-based recognition and estimating severity of PTSD using machine learning. *Journal of Affective Disorders*, *362*, 859–868. https://doi.org/10.1016/j.jad.2024.07.015

Ishkhanyan, B., Boye, K., & Mogensen, J. (2019). The meeting point: Where language production and working memory share resources. *Journal of Psycholinguistic Research*, *48*(1), 61–79. https://doi.org/10.1007/s10936-018-9589-0

Jensen, E. S. (2015). Subjektellipse eller informationsstrukturel organisering i sms. In I. S. Schoonderbeek & T. T. Hougaard (Eds.), *15. Møde om Udforskningen af Dansk Sprog* (pp. 171–188).

Kleim, B., Horn, A. B., Kraehenmann, R., Mehl, M. R., & Ehlers, A. (2018). Early linguistic markers of trauma-specific processing predict post-trauma adjustment. *Frontiers in Psychiatry*, *9*, 645. https://doi.org/10.3389/fpsyt.2018.00645

Koenen, K. C., Ratanatharathorn, A., Ng, L., McLaughlin, K. A., Bromet, E. J., Stein, D. J., Karam, E. G., Meron Ruscio, A., Benjet, C., Scott, K., Atwoli, L., Petukhova, M., Lim, C. C., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Bunting, B., Ciutan, M., de Girolamo, G., … Kessler, R. C. (2017). Posttraumatic stress disorder in the World Mental Health Surveys. *Psychological Medicine*, *47*(13), 2260–2274. https://doi.org/10.1017/S0033291717000708

Kolk, H. (1995). A time-based approach to agrammatic production. *Brain and Language*, *50*(3), 282–303. https://doi.org/10.1006/brln.1995.1049

Lanius, R. A. (2015). Trauma-related dissociation and altered states of consciousness: A call for clinical, treatment, and neuroscience research. *European Journal of Psychotraumatology*, *6*, 27905. https://doi.org/10.3402/ejpt.v6.27905

LeDoux, J. (2012). Rethinking the emotional brain. *Neuron*, *73*(4), 653–676. https://doi.org/10.1016/j.neuron.2012.02.004

Maher, L. M. (2018). Broca's aphasia and grammatical processing. In A. M. Raymer & L. J. Gonzalez Rothi (Eds.), *The Oxford Handbook of Aphasia and Language Disorders* (pp. 145–160). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199772391.013.9

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Marmar, C. R., Brown, A. D., Qian, M., Laska, E., Siegel, C., Li, M., Abu-Amara, D., Tsiartas, A., Richey, C., Smith, J., Knoth, B., & Vergyri, D. (2019). Speech-based markers for posttraumatic stress disorder in US veterans. *Depression and Anxiety*, *36*(7), 607–616. https://doi.org/10.1002/da.22890

Mayer, J. F., & Murray, L. L. (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders*, *45*(5), 325–339. https://doi.org/10.1016/j.jcomdis.2012.06.002

McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: Central role of the brain. *Physiological Reviews*, *87*(3), 873–904. https://doi.org/10.1152/physrev.00041.2006

Menn, L., Obler Loraine, K., & Miceli, G. (1990). *Agrammatic aphasia: A cross-language narrative sourcebook*. John Benjamins.

Miragoli, S., Camisasca, E., & Di Blasio, P. (2019). Investigating linguistic coherence relations in child sexual abuse: A comparison of PTSD and non-PTSD children. *Heliyon*, *5*(2), e01163. https://doi.org/10.1016/j.heliyon.2019.e01163

Morgan, C. J. (2017). Use of proper statistical techniques for research studies with small samples. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, *313*(5), L873–L877. https://doi.org/10.1152/ajplung.00238.2017

Nedergaard, J. S. K., Martínez-Ferreiro, S., Fortescue, M. D., & Boye, K. (2020). Non-fluent aphasia in a polysynthetic language: Five case studies. *Aphasiology*, *34*(6), 675–694. https://doi.org/10.1080/02687038.2019.1643000

Norte, C. E., Vargas, A. L. V., & de Carvalho Silveira, A. (2024). Post-traumatic stress disorder and working memory: A systematic review. *Trends in Psychology*, *32*(2), 612–623. https://doi.org/10.1007/s43076-022-00206-2

Nygard, M. (2018). *Norwegian discourse ellipsis: Clausal architecture and licensing conditions*. John Benjamins Publishing Company.

Papini, S., Yoon, P., Rubin, M., Lopez-Castro, T., & Hien, D. A. (2015). Linguistic characteristics in a non-trauma-related narrative task are associated with PTSD diagnosis and symptom severity. *Psychological Trauma*, *7*(3), 295–302. https://doi.org/10.1037/tra0000019

Pavlícková, K., Gärtner, J., Voulgaropoulou, S. D., Fraemke, D., Adams, E., Quaedflieg, C. W. E. M., Viechtbauer, W., & Hernaus, D. (2024). Acute stress promotes effort mobilization for safety-related goals. *Communications Psychology*, *2*(1). https://doi.org/10.1038/s44271-024-00103-7

Payne, P., Levine, P. A., & Crane-Godreau, M. A. (2015). Somatic experiencing: Using interoception and proprioception as core elements of trauma therapy. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00093

Pitman, R. K., Shin, L. M., & Rauch, S. L. (2001). Investigating the pathogenesis of posttraumatic stress disorder with neuroimaging. *The Journal of Clinical Psychiatry*, *62*(Suppl 17), 47–54.

Prins, R., & Bastiaanse, R. (2004). Analysing the Spontaneous Speech of Aphasic Speakers. *Aphasiology*, *18*(12), 1075–1091. https://doi.org/10.1080/02687030444000534

Rothschild, B. (2021). *Revolutionizing trauma treatment: Stabilization, safety, & nervous system balance*. W. W. Norton & Company.

Rainnie, D. G., & Ressler, K. J. (2009). Physiology of the amygdala: Implications for PTSD. In P. J. Shiromani, T. M. Keane, & J. E. LeDoux (Eds.), *Post-Traumatic Stress Disorder: Basic science and clinical practice* (pp. 39–78). Humana Press. https://doi.org/10.1007/978-1-60327-329-9_3

Rauch, S. L., van der Kolk, B. A., Fisler, R. E., Alpert, N. M., Orr, S. P., Savage, C. R., Fischman, A. J., Jenike, M. A., & Pitman, R. K. (1996). A symptom provocation study of posttraumatic stress

disorder using positron emission tomography and script-driven imagery. *Archives of General Psychiatry*, *53*(5), 380–387. https://doi.org/10.1001/archpsyc.1996.01830050014003

Rogalsky, C., Matchin, W., & Hickok, G. (2008). Broca's area, sentence comprehension, and working memory: An fMRI study. *Frontiers in Human Neuroscience*, *2*. https://doi.org/10.3389/neuro.09.014.2008

Römisch, S., Leban, E., Habermas, T., & Döll-Hentschker, S. (2014). Evaluation, immersion, and fragmentation in emotion narratives from traumatized and nontraumatized women. *Psychological Trauma*, *6*(5), 465–472. https://doi.org/10.1037/a0035169

Shin, L. M., Kosslyn, S. M., McNally, R. J., Alpert, N. M., Thompson, W. L., Rauch, S. L., Macklin, M. L., & Pitman, R. K. (1997). Visual imagery and perception in posttraumatic stress disorder: A positron emission tomographic investigation. *Archives of General Psychiatry*, *54*(3), 233–241. https://doi.org/10.1001/archpsyc.1997.01830150057010

Steinkrauss, R., & Schmid, M. S. (2016). 16. Entrenchment and language attrition. In H.-J. Schmid (Ed.), *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge* (pp. 367–384). De Gruyter Mouton. https://www.degruyterbrill.com/document/doi/10.1515/9783110341423-017/html

Szmrecsányi, B. (2008). On operationalizing syntactic complexity. In *Le Poids des Mots. Proceedings of the 7th international conference on textual data statistical analysis* (Vol. 2, pp. 1032–1039). Leuven University Press.

tactustherapy (2017, June 14). *Broca's Aphasia (Non-Fluent Aphasia)* [Video]. YouTube. https://www.youtube.com/watch?v=JWC-cVQmEmY

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Thompson, C. K., & Bastiaanse, R. (2012). Introduction to agrammatism. In R. Bastiaanse & C. K. Thompson (Eds.), *Perspectives on Agrammatism* (pp. 1–16). Psychology Press.

Thy, A. D. M., Gerona, J., Martínez-Ferreiro, S., Popov, S., & Boye, K. (2024). Deictic vs. anaphoric pronouns: A comparison of fluent and non-fluent aphasia in English and Tagalog. *Language, Cognition and Neuroscience*, *39*(7), 909–923. https://doi.org/10.1080/23273798.2024.2368114

van der Kolk, B. A. (2014). *The body keeps the score: Brain, mind, and body in the healing of trauma*. Viking.

van der Kolk, B. A., & Fisler, R. (1995). Dissociation and the fragmentary nature of traumatic memories: Overview and exploratory study. *Journal of Traumatic Stress*, *8*(4), 505–525. https://doi.org/10.1007/BF02102887

VanderWeele, T. J., & Mathur, M. B. (2019). Some desirable properties of the Bonferroni correction: Is the Bonferroni correction really so bad? *American Journal of Epidemiology*, *188*(3), 617–618. https://doi.org/10.1093/aje/kwy250

Wilder, A., & Redmond, S. M. (2022). The reliability of short conversational language sample measures in children with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *65*(5), 1939–1955. https://doi.org/10.1044/2022_JSLHR-21-00628World Health Organization. (2001). *International classification of functioning, disability and health: ICF*. World Health Organization. https://iris.who.int/handle/10665/42407

World Health Organization. (2019). 6B40 Post traumatic stress disorder. In *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/browse/2024-01/mms/en#2070699808

Wright, H. H., Silverman, S., & Newhoff, M. (2003). Measures of lexical diversity in aphasia. *Aphasiology*, *17*(5), 443–452. https://doi.org/10.1080/02687030344000166

Regina Mezőlaki* **& Lívia Ivaskó**
University of Szeged, Hungary

DOI: 10.19090/slm.6.13

FACE-SAVING ACTS IN FAUX PAS STORIES: THE IMPORTANCE OF CONTEXTUAL CUES DURING INFERENTIAL PROCESSES OF AUTISTIC AND NON-AUTISTIC ADULTS

A faux pas occurs when someone unintentionally says or does something socially inappropriate. Recognition of a faux pas requires cognitive and affective theory of mind ability. The Faux Pas Recognition test consists of short stories with and without faux pas, followed by target questions assessing mental state and emotion attribution abilities.
The aim of this study was to gain deeper insight into the social cognitive abilities of autistic (N=12) and non-autistic (N=11) adults. We examined (a) whether affective and cognitive mental state attribution in autistic and non-autistic participants can be assessed more effectively through a modified version of the Faux Pas Recognition Test, and (b) whether the representation of an avoidance process as a contextual cue within the target stories of the Faux Pas Recognition Test influences the inferential and affective mental state attribution processes in autistic and non-autistic individuals. Our results indicate that the modified Faux Pas Recognition Test provides a more comprehensive assessment of participants' social cognitive abilities, and that the presence of avoidance processes as contextual cues supports emotion attribution and consideration of the implicit meaning of an utterance.

Keywords: faux pas, facework, autism, contextual cue, avoidance process

1. INTRODUCTION

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), autism spectrum disorder[16] is characterized by persistent deficits in social communication and social interaction, such as difficulties in social-emotional reciprocity, nonverbal communicative behaviours, or developing, understanding, and maintaining relationships (APA, 2013). For example, adjusting one's behaviour to various social contexts can be a challenge for autistic individuals.

Another diagnostic criterion for autism is the presence of restricted, repetitive patterns of behaviour, interests, or activities, and these characteristics can also highly affect one's social behaviour in an interaction. Social cognitive abilities are linked to pragmatic ones, and theory of mind abilities are necessary for perspective-taking to engage in an interaction successfully. Social cognitive differences are related to pragmatic deficits and the understanding of the self in autism spectrum disorder (condition) (Bosco et al. 2018; Lombardo & Baron-Cohen, 2010; APA, 2013). Multiple studies have shown that autistic people might favor honesty and might be less concerned about others' and their own face and self than non-autistic individuals (Howlin, 1996; Belek, 2018; Koskinen et al., 2021; Williams & Happé,

* Corresponding author: Regina Mezőlaki; email: mezolaki.regina@gmail.com

[16] The authors prefer the term *autism spectrum condition* over *autism spectrum disorder*, as the latter reflects a more medicalized perspective. Furthermore, throughout the study, the terms *autism* and *autistic person* are primarily used, in line with the neurodiversity paradigm and consistent with the terminology preferred by a large part of the autistic community (Bottema-Beutel et al. 2021).

2009), while other scholars claim that autistic people can successfully take part in interactions in line with the so-called social norms, especially when conventionalized forms and contextual cues facilitate successful interaction (Sirota, 2004; Kissine et al., 2015). In this study, we explore how faceworking acts (avoidance processes, such as changing the topic when an awkward situation occurs) as contextual cues affect the understanding of faux pas stories in the case of autistic and non-autistic adults' performance.

2. THEORETICAL BACKGROUND

This paper focuses on the phenomenon of faux pas, which, by definition, involves a false belief and a feeling of awkwardness. According to Baron-Cohen and his colleagues (1999:408) part of the definition of a faux pas should include [an] "uh-oh" reaction […], best summed up by the phrase "I wish I hadn't said that!"'. The following story from the Faux Pas Recognition test (Stone et al., 1998) exemplifies a faux pas in social interaction, in which Sarah probably wishes she hadn't revealed the surprise party.

> 'Story 2. Helen's husband was throwing a surprise party for her birthday. He invited Sarah, a friend of Helen's, and said, "Don't tell anyone, especially Helen." The day before the party, Helen was over at Sarah's and Sarah spilled some coffee on a new dress that was hanging over her chair.
>
> "Oh!" said Sarah, "I was going to wear this to your party!"
>
> "What party?" said Helen.
>
> "Come on," said Sarah, "Let's go see if we can get the stain out."' (Stone et al., 1998)

For the successful recognition and interpretation of a faux pas, it is essential to have a comprehensive understanding of human cognitive processes, social relations and accepted social norms.

2.1. A cognitive pragmatic perspective on communication

Ostensive-inferential communication means that the communicator indicates, by verbal or non-verbal ostension, their intention to communicate a given set of information, which is interpreted by their interlocutor through decoding and inference processes. Both verbal and nonverbal communication can only be achieved through both the expression and recognition of intentions. Therefore, successful communication requires that the communicator's informative and communicative intentions are fulfilled and mutually manifested. These intentions depend on the actual context of the interaction, which builds up from the physical, cognitive, and social contexts (Sperber & Wilson, 1995; Ivaskó & Németh T., 2002).

According to Frith and Frith (2007), social cognition is the sum of processes that allow individuals of the same species to interact with each other, and social cognitive differences are related to

the differences and deficits in pragmatic abilities (Bosco et al., 2018). The intentions with which the communicator and their interlocutor engage in interactions and the intentions they attribute or can attribute to each other depend on their cognitive abilities, for instance, on their theory of mind ability, the ability to attribute mental states, emotions, and intentions to each other, and to predict the behaviour of the other person based on these attributions (Premack & Woodruff, 1978). Mental state attribution is essential for perspective-taking and to be able to interpret the implied meaning of the context, therefore, essential for successful communication (Baron & Cohen, 1995; Sperber & Wilson, 1995; Győri, 2006).

The term 'faux pas' frequently occurs in studies in cognitive psychology, since the difficulty in recognizing and interpreting a situation involving a faux pas may indicate an underdeveloped (Gál, 2015) or impaired (Baron & Cohen et al., 1999; Varga et al., 2008; Gál et al., 2011) social cognitive ability.

2.2. A sociopragmatic perspective on communication

Brown and Levinson (1987:61) assume that 'all competent adult members of a society have (and know each other to have) […] face, the public self-image that every member wants to claim for himself […], [and] certain rational capacities, in particular consistent modes of reasoning from ends to the means that will achieve those ends.' Rationally thinking members have two wants: the freedom of action (negative face) and the positive self-image claimed by interactants (positive face) (Goffman, 1967; Brown & Levinson, 1987).

When someone has or maintains face, they present an image of themselves that is internally consistent and supported by judgements of other participants (Goffman 1967). With facework, one can counteract incidents with symbolic implications that threaten one's face, as face-threatening acts run contrary to the face wants of the addressee and/or of the speaker. With redressive action such as modifications or additions, the speaker can save the face of the addressee of the potential face-threatening action and indicate that the threat was neither intended nor desired (Brown & Levinson, 1987).

As Goffman (1967: 14-15) claims, there are three levels of responsibility that one may have for a face-threatening action that he has created. First, the speaker is innocent, his offence is unintended, and he would have tried to avoid it if he had seen the negative consequences of it. Goffman names this kind of threats faux pas, gaffes, boners, or bricks. Secondly, it may happen that the speaker commits the threat intentionally to cause an open insult. Lastly, incidental offences can happen when the speaker acts despite the negative consequences, but not because of them. In these cases, the threats are unintended but possibly predictable by-products of the act. These threats can address both the speaker's and their interlocutor's faces.

There are two basic kinds of facework: the avoidance process and the corrective process. The former means that the best way to prevent threats is to avoid contacts that are likely to create an incident. When a person chances an encounter, they can keep off sensitive topics and avoid activities that would lead to the expression of a threat. Also, they can change the topic of the conversation if they recognize the potential for a threat. In situations when the person who commits the faux pas fails to prevent the incident, he can attempt to maintain the fiction that the threat has not happened. The corrective process refers to an act when the interactants acknowledge the incident as a threat, deem it worthy of direct attention and attempt to mitigate its threats (Goffman, 1967: 5-21).

3. THE FAUX PAS RECOGNITION TEST (ADULT VERSION)

According to Baron-Cohen and his colleagues (1999: 408), a faux pas occurs when someone says something that the other person might not want to hear or know, and which typically has negative consequences that the speaker never intended. The speaker usually feels a mix of regret and embarrassment, feeling bad for the hearer after the faux pas has been committed. In this section, we present the Faux Pas Recognition Test designed for adults (Stone et al. 1998; Baron-Cohen et al. 1999; Gregory et al. 2002).

3.1. The design of the original Faux Pas Recognition Test

Detecting a faux pas requires the understanding that the speaker's and the hearer's knowledge states are different, and that the utterance might have an emotional impact on the hearer. This means that the recognition and understanding of a faux pas requires first-order theory of mind ability to correctly attribute thoughts and beliefs to the characters of a story in which a faux pas was committed. Additionally, it also requires second-order theory of mind ability to understand that the characters have different beliefs and thoughts about each other. Also, with the affective theory of mind ability, one should be able to recognize the emotional impact caused by the faux pas, while with the cognitive theory of mind ability, it is important to detect that the speaker committed the faux pas unintentionally, without any malicious or offensive intention, because of his false belief (Stone et al. 1998; Baron-Cohen et al. 1999; Gregory et al. 2002).

The Faux Pas Recognition test assesses the participant's cognitive and affective mental state attribution abilities by requiring responses to target questions concerning short stories that either contain a conversational faux pas (10 target stories) or do not (10 control stories, containing minor inconveniences that are resolved). The first target question ('*Did anyone say something they shouldn't have said or*

something awkward?') always follows the stories, while the rest of the five target questions are only asked if the participant detects the faux pas in the target stories, or falsely detects one in the control stories. The two control questions are raised following all the stories, both target and control stories. The target questions assess: (1) detection of the faux pas or the lack of faux pas, (2) identification of the person who committed the faux pas ('*Who said something they shouldn't have said or something awkward?'*), (3) understanding inappropriateness: explanation of why the faux pas was inappropriate *('Why shouldn't they have said it or why was it awkward?'*), (4) intentions: attribution of intention to the character who committed faux pas ('*Why do you think they said it?*'), (5) belief: recognition the false belief of the character who committed faux pas (*'Did X know/realize that Y?'*), and (6) empathy: understanding of the emotional state of the addressee of the faux pas (*'How do you think X felt?'*). The control questions are designed to examine whether the participant has comprehended the story. To successfully complete the 20 short tasks (stories), the participant must recognize that the speaker made an inappropriate remark, identify the speaker's false belief, acknowledge the absence of offensive intent, and infer the negative emotional consequences (Stone et al., 2003).

The full set of stories, questions and instruction for scoring is available on the website of the Autism Research Centre.[17]

3.2. Previous studies using the original test

Versions of the Faux Pas Recognition Test have been used to assess both children (Baron-Cohen et al. 1999) and adults (Stone et al., 1998; Gregory et al., 2002). Baron-Cohen and colleagues (1999) found that despite doing well on the first- and second-order theory of mind tasks, autistic children and children with Asperger's Syndrome[18] performed significantly lower on the Faux Pas Recognition Test than 11-year-old non-autistic boys and girls. Also, their results showed that both groups could identify that no one had said anything awkward in the control stories, however in the target stories autistic children identified faux pas significantly less often than the control group (Baron-Cohen et al. 1999).

Zalla et al. (2009) investigated mindreading abilities in a group of adults diagnosed with Asperger Syndrome by using the faux pas task. They found that their research participants generally successfully detected faux pas (detection of faux pas, person identification), but generally were unable to provide correct justifications regarding reasons and intentions. They failed to recognize the mistaken belief or forgetfulness that caused the faux pas, and they did not understand that the protagonists in the

[17] https://docs.autismresearchcentre.com/tests/FauxPas_Adult.pdf
[18] Since the 5th ed. of the APA Diagnostic Manual (2013), Asperger's Syndrome falls into the category of the autism spectrum disorders.

stories did not commit the faux pas intentionally and thus did not offend the addressee deliberately. Also, they did not understand the emotional impact resulting from the involuntary inappropriate act.

Assessing typically developing Hungarian adults (sample ranging from 11 to 75 years of age) on the adult version of the Faux Pas Recognition test, Gál (2015: 102) found an increase in Faux Pas Recognition Test total scores between ages 11 and 12, followed by a slight decline in performance beginning in the 13–17 age group, but there was no statistical difference between 18–29- and 30–59-year-olds. Theory of mind performance was lowest in the older adult group (60–75 years), compared to all younger age groups. The results show that both the cognitive and affective components of faux pas understanding decline with advancing age. The results of her study showed no significant differences across the different age groups on the control stories, suggesting that the age-related differences reflect genuine variations in theory of mind ability. These findings are consistent with previous research demonstrating age-related decreases in theory of mind performance (Maylor et al., 2002; Duval et al., 2011), and with studies showing reduced performance among older adults on the Faux Pas Recognition Test (Wang & Su, 2006).

3.3. Cognitive and sociopragmatic approach to the test material

In this section, we present the categorization of the faux pas stories from the Faux Pas Recognition Test based on whether they contain facework/ redressive action/ avoidance process or not. In our previous research (Mezőlaki & Ivaskó, 2024a;b;c), a comparison of the stories containing faux pas was carried out according to the following criteria: 1) the original knowledge about the state of the world, 2) the truth-content of the utterance involving a faux pas, 3) whether the interactants employed any redressive action after the faux pas occurred, and 4) whether the speaker regretted committing a faux pas. Here, we focus on the third and fourth aspects of the analysis.

Out of the 10 faux pas stories, only 5 (number 2 – quoted in Section 2., 11, 12, 15, 18) contain an utterance with which one of the participants aims to maintain the fiction that the threat has not happened and, therefore, these contain facework. The face-working participants in these 5 stories always try to save the offended person's face by changing the topic, and none of the 10 faux pas stories contains any corrective process. On the other hand, some of the control stories do contain a corrective process (e.g. number 20), in which the participants recognize and address the misunderstanding, and try to resolve it. This raises the question of whether the character committing the faux pas regrets the threat, and whether the research participant can attribute this mental state to the speaker or not.

None of the target stories mention explicitly whether the speaker who committed the faux pas regrets causing the embarrassment. There are two target stories (2 and 12) the content of which implies that the speaker might regret their utterance. In story 2, this is more obvious, since the speaker herself performs the redressive action. In story number 12, the addressee of the faux pas was the overhearer, and it was another person, the hearer, who tried to apply the avoidance process and save everybody's face. If there is no facework in the stories, or it is not the speaker who commits it, then it is not clear whether the person committing the faux pas regrets their face-threatening act or not. The only way to resolve this is to assume and imagine that after the story the speaker realizes that they committed a faux pas and to assume that they have not intentionally offended their partner. The following story is an example of a faux pas story from the Faux Pas Recognition test (Stone et al. 1998) that does not contain facework. In the story, Lisa does not know that Jill is the one who bought the curtains and commits a faux pas by criticizing them. Unlike story 2, there is no topic shifting in the text.

> 'Story 4. Jill had just moved into a new flat. Jill went shopping and bought some new curtains for her bedroom. When she had just finished decorating the flat, her best friend, Lisa, came over. Jill gave her a tour of the flat and asked, "How do you like my bedroom?" "Those curtains are horrible," Lisa said. "I hope you are going to get some new ones!"'

In our previous research (Mezőlaki & Ivaskó, 2024a;b;c), we added two more questions to the Faux Pas Recognition Test to get a more accurate picture of the research participants' cognitive and affective mental state attribution abilities, metapragmatic abilities, and their inferential processes. In these questions, 'X' is the person who commits the faux pas and 'Y' is the offended person. The following two questions were added to the test:

1. Did X want Y to hear what X was saying? Why?
2. Does X regret saying what X was thinking? Why?

By addressing the first question, we hoped to gain a deeper understanding of the research participant's beliefs about the character's informative and communicative intentions, and it might also be revealed who the research participant believes to be the addressee of the utterance. The second question aims to recognize the emotions of the character who committed the faux pas, therefore, we can learn more about the research participant's affective mental state attribution ability. Answering this question may also reveal whether participants take into consideration the redressive action (topic shifting) or the lack of it in the stories. If the participants answer 'If the person who committed the faux pas realizes their mistake, they will regret it!', we can also get information about the extent to which the participants rely on the context of the particular story and the extent to which they can build on the implications of the text. Based on the results of the modified version of the test, it was found that the explicit formulation of the avoidance

process (such as topic shifting) might affect whether the research participants attribute immediate regret to the person committing the faux pas (Mezőlaki & Ivaskó, 2024a;b;c).

4. ANALYSIS OF THE APPLICATION OF THE MODIFIED FAUX PAS RECOGNITION TEST

In the following, we present how autistic and non-autistic adults perform on the modified version of the Faux Pas Recognition Test.

4.1 Materials and method

The Hungarian adaptation of the adult version of the Faux Pas Recognition Test (Gál, 2015) was used, complemented with our two previously introduced target questions. These questions were only asked if the participants detected a faux pas, and were asked before the control questions.

All procedures were in accordance with the ethical standards of the Medical Research Council, University of Szeged, and an ethical permission[19] was provided by the Medical Research Council, Hungary in 2021. The recruitment process and data collection took place between March 2022 and April 2023. The recruitment was carried out on social media sites, and institutions and organizations dealing with autism helped us spread the call for research. Participation in the research was voluntary, and the data was processed anonymously. Each participant filled in an informed consent form. Research participants could withdraw from the research at any point and request the deletion of their data.

The data collection took place in person, in a quiet, low-stimulus environment. Only the research participant and the examiner were present in the room. A tablet and a voice recorder were used to record data. Participants were provided with individual printed copies of the target and control stories of Faux Pas Recognition Test, which they could follow while the researcher read the stories aloud. The questions were presented by the researcher exclusively in oral form. Basic demographic data were collected via a Jotform questionnaire administered on a tablet device. Jamovi (version 2.2) software was used for statistical analysis. Descriptive statistics were conducted on all major variables, Mann-Whitney U tests were used for pairwise between-group comparisons. Friedman tests and Wilcoxon tests were used for in-group comparisons.

Scoring the Faux Pas Recognition Test was in accordance with Stone and her colleagues' work (1998). For each faux pas story (target story), the research participant gets 1 point for each target question answered correctly. Overall, there was a total score of 60 points that research participants can get on the

[19] Case number: IV/3899- 1 /2021/EKU, Title in Hungarian: Udvariassági stratégiák autizmus spektrum zavarban: homlokzatfenyegető aktusok tipikus és atipikus nyelvhasználók körében.

6 target questions on the 10 faux pas stories. For control stories, participants get 1 point for detecting that there is nothing inappropriate in the stories, and get 1 point for each unanswered target question (total score of 60 points). The maximum reachable score in ratio is 1.0 on each question. For further information about the scoring, read the instructions: https://docs.autismresearchcentre.com/tests/FauxPas_Adult.pdf (Stone et al. 1998: 23–27).

Responses for the additional questions (Did X want Y to hear what X was saying? Why?, Does X regret saying what X was thinking? Why?) were scored as follows. We categorized the participants' answers into groups, and summed up the total answers per categories. The additional questions were only asked if the participants detected a faux pas in the story. As there were 10 faux pas stories in the test, one participant could answer a question in a certain way 10 times. Later, we divided the faux pas stories into two groups based on whether they involved an avoidance process or not: 5 stories that involve avoidance process (topic shifting) and 5 stories that do not. Scores were measured at group level.

4.2. Participants

The group of research participants included a total of 23 persons, 12 adult autistic persons with level 1 severity for autism spectrum condition, and 11 adult non-autistic persons. All participants were of Hungarian nationality and spoke Hungarian as their primary language. The average age of autistic people was 36.7 years (SD=7.94), while the average age of non-autistic people was 29 years (SD=6.02). All participants were between 19 and 52 years of age at the time of data collection. Based on the findings of Gál (2015), the age differences among our participants should not have a significant statistical effect on the theory of mind ability of the participants.

5. RESULTS

5.1. Results on the original elements of the Faux Pas Recognition Test

Table 1. shows the performances of the two research groups on the Faux Pas Recognition Test. Scores are in ratio. The assumption of normality was tested using the Shapiro–Wilk test. Results indicated that normality was not met within groups ($p < 0.05$), therefore, Mann-Whitney U tests were used for pair-wise between-group comparisons. We found no significant differences between autistic and non-autistic individuals' performances, neither in the aggregated total scores nor in the item-level comparisons of the target (faux pas) questions (Q1-Q6).

Table 1. Group comparisons of total, subscale, and item-level scores on the original Faux Pas Recognition Test (scores are in ratio).

Test elements	**U**	**p**	**Autistic participants** median	**Non-autistic participants** median	**Autistic participants** mean (SD)	**Non-autistic participants** mean (SD)
Faux pas Recognition total scores including control stories (no control questions)	55	0.518	0.879	0.925	0.846 (0.132)	0.864 (0.135)
Faux pas stories scores (no control stories and control questions)	63	0.88	0.816	0.850	0.753 (0.271)	0.758 (0.240)
Control stories scores (no control questions)	42.5	0.118	0.925	1	0.937 (0.06)	0.970 (0.06)
Control questions (of both faux pas and control stories)	60.5	0.384	1	1	0.998 (0.007)	1 (0.00)
Target questions of target (faux pas) stories						
Q1. FP Detection	61	0.772	0.900	0.900	0.833 (0.277)	0.809 (0.221)
Q2. FP Detection (person identification -Who said it?)	60.5	0.750	0.900	0.900	0.808 (0.294)	0.791 (0.234)
Q3. Understanding Inappropriateness	64	0.924	0.900	0.900	0.775 (0.302)	0.791 (0.234)
Q4. Intentions	66	1.00	0.7500	0.800	0.669 (0.302)	0.673 (0.290)
Q5. Belief	58	0.640	0.700	0.800	0.635 (0.303)	0.700 (0.249)
Q6. Empathy	64.5	0.959	0.900	0.900	0.800 (0.273)	0.791 (0.230)

Performances on the target questions were also compared within groups. Because the assumption of normality was not met for all variables, Friedman tests were conducted. The Friedman test was significant in case of the autistic group $\chi^2(5) = 28.2$, $p < 0.001$, as well as of the non-autistic group $\chi^2(5) = 32.9$, $p < 0.001$. These findings indicate that there are fundamental differences among the variables within both groups. In case of both the autistic and the non-autistic research groups, Q1 (FP detection), Q2 (Person identification), Q3 (Understanding inappropriateness) and Q6 (Empathy) scores are significantly higher than Q4 (Intentions) and Q5 (Belief) scores, measured by Durbin-Conover tests.

5.2. Results of the additional questions of the Faux Pas Recognition Test

Five categories were created based on the answers the participants provided to the first additional question 'Did [the character who committed faux pas] want [the addressee] to hear [the character who committed faux pas]?'. These are the following: 1) 'I do not know': when the participants replied that 'I don't know the answer', or 'One cannot know the answer', or 'I cannot decide'; 2) 'Yes': we applied this category to the answers that indicated that the participant attributed offensive intent to the speaker; 3) 'No': this category encompassed responses interpreted as reflecting the participant's perception of the speaker's intent as not offensive; 4) 'Spoken to them': when the participant took the literal meaning of the question, and answered that the speaker indeed spoke to the other, therefore, wanted to be heard; 5)

‘Without offensive intent’: when the participant explicitly answered that the speaker wanted to be heard, but the offence was not intentional. Target questions, and additional questions were not raised if participants did not recognize the faux pas in the target stories, these instances are categorized as ‘Omitted questions’.

We created 4 categories based on the answers the participants provided to the second additional question ‘Does [the character who committed faux pas] regret saying what [the character who committed faux pas] was thinking?’. These are the following: 1) ‘I do not know’: when the participants replied that they did not know the answer or that one cannot know the answer or that they were unable to decide; 2) ‘No’: if the participants replied that the speaker does not regret committing the faux pas; 3) ‘Yes’: when the participants said that the speaker does regret committing the faux pas; and 4) ‘Would regret’: when the participants answered that the speaker would regret committing the faux pas after realizing the ‘mistake’, negative emotional consequences. Again, ‘Omitted question’ marks the instances when the participants did not detect the faux pas in the target stories.

Table 2. shows how frequently each response was given by the two groups, as well as the between-group comparisons of these responses. Scores are calculated as the sum of item responses (out of 10 / question). The assumption of normality was tested using the Shapiro–Wilk test. Results indicated that in most cases normality was not met within groups ($p < 0.05$), therefore, Mann-Whitney U tests were used for pair-wise between-group comparisons. We found only one significant difference between the groups: autistic research participants responded ‘I do not know’ more frequently to the question ‘*Does [the character who committed faux pas] regret saying what [the character who committed faux pas] was thinking*?’ than members of the non-autistic group ($p = 0.013$).

Table 2. Response frequencies and group comparisons on the additional questions of the Faux Pas Recognition Test (scores are summed).

Response categories	**U**	**p**	**Autistic participants** median	**Non-autistic participants** median	**Autistic participants** mean (SD)	**Non-autistic participants** mean (SD)
Question 'Did [the character who committed faux pas] want [the addressee] to hear [the character who committed faux pas]?'						
'I do not know'	49.5	0.092	0.00	0.00	0.250 (0.452)	0.000 (0.000)
'Yes'	59.0	0.672	1.00	1.00	1.083 (1.676)	1.091 (1.136)
'No'	58.0	0.604	1.00	1.00	1.167 (1.403)	0.727 (0.467)
'Spoken to them'	52.0	0.391	3.00	0.00	2.917 (2.503)	2.273 (3.197)
'Without offensive intent'	50.0	0.332	2.00	3.00	2.667 (2.015)	4.000 (2.864)
Omitted question	64.5	0.949	1.00	1.00	1.917 (2.712)	1.909 (2.212)
Question 'Does [the character who committed faux pas] regret saying what [the character who committed faux pas] was thinking?'						
'I do not know'	30.0	**0.013***	1.00	0.00	1.42 (1.38)	0.273 (0.905)
'Yes'	46.5	0.229	1.50	2.00	1.75 (1.71)	2.273 (1.191)
'No'	62.0	0.819	1.00	2.00	2.08 (2.19)	1.727 (1.104)
'Would regret'	53.0	0.438	3.50	3.00	3.00 (2.04)	3.818 (2.359)
Omitted question	64.5	0.949	1.00	1.00	1.75 (2.73)	1.909 (2.212)

* Significant at < 0,05 between autistic and non-autistic groups with Mann-Whitney U test.

Frequencies of responses were also compared within groups. Because the assumption of normality was not met for all variables, Friedman tests were conducted. In case of the first added question (Did [the character who committed faux pas] want [the addressee] to hear [the character who committed faux pas]?), the Friedman test was significant in case of the autistic group $X^2(5) = 18.3$, $p = 0.003$, as well as of the non-autistic group $X^2(5) = 19.2$, $p = 0.002$. These findings indicate that there are fundamental differences among the variables within both groups. In case of the autistic group, 'Spoken to them' was responded significantly more often than 'I do not know' ($p < 0.001$), 'Yes' ($p = 0.01$), and 'No' ($p = 0.034$). Also, 'Without offensive intent' was also responded significantly more often by autistic participants than 'I do not know' ($p < 0.001$), 'Yes' ($p = 0.012$) and 'No' ($p = 0.039$). Non-autistic group of participants' 'Without offensive intent' responses were significantly more frequent than 'I do not know' ($p < 0.001$), 'Yes' ($p = 0.013$), 'No' ($p = 0.015$) and 'Spoken to them' ($p = 0.013$). Also 'Spoken to them' was responded significantly more often than 'I do not know' ($p = 0.015$). Non-autistic participants responded 'No' significantly more often than 'I do not know' ($p = 0.013$), and 'Yes' than 'I do not know ($p = 0.015$).

In case of the second added question (Does [the character who committed faux pas] regret saying what [the character who committed faux pas] was thinking?), the Friedman test was not significant in case of the autistic group $X^2(4) = 8.55$, $p = 0.073$, meaning that there is no difference among the variables fundamentally, therefore, results of the Durbin-Conover test should be interpreted with caution. In the case of the non-autistic group, the Friedman test was significant $X^2(4) = 16.1$, $p = 0.003$, resulting in a

fundamental difference among the variables. The autistic group of participants supplied the answer 'Would regret' significantly more often than 'I do not know' ($p = 0.029$) and 'Yes' ($p = 0.041$). Non-autistic participants responded significantly more often with 'Would regret' than 'I do not know' ($p < 0.001$). Also, they replied significantly more often with 'Yes' ($p < 0.001$) and 'No' ($p = 0.011$) than 'I do not know'.

5.3. Faux Pas Recognition Test results by presence of avoidance process (topic shifting) in the target questions

By categorizing the faux pas stories based on whether they explicitly include facework (avoidance process, topic shifting specifically) or not, we aimed to gain a deeper understanding of the mental state attribution abilities of the research participants. Pairwise in-group comparisons using Wilcoxon tests revealed no significant effect of the presence of the avoidance process on the original target questions of the Faux Pas Recognition Test (Q1–Q6) in either group.

For the first additional question ('Did [the character who committed the faux pas] want [the addressee] to hear [the character who committed the faux pas]?'), we examined whether the presence of topic shifting affected participants' responses. The assumption of normality was tested using the Shapiro–Wilk test. Results indicated that normality was not met within groups ($p<0.05$), therefore, Wilcoxon tests were used for pair-wise within-group comparisons. In the case of the autistic participants' group, we found that they replied 'Spoken to them' significantly more often when the faux pas stories did not include avoidance process than when they did ($T = 0$, $p = 0.02$). Also, autistic participants responded 'No' significantly more often if the faux pas stories included avoidance process ($T = 36$, $p = 0.01$). Analyzing the non-autistic research group's responses to the first additional question, we found one significant difference related to the presence of the topic shifting: they replied 'No' significantly more often when the stories included avoidance process ($T = 36$, $p = 0.006$).

As for the second additional question (Does [the character who committed faux pas] regret saying what [the character who committed faux pas] was thinking?), we also examined the differences in responses in relation to the presence of the avoidance process. Again, the assumption of normality was tested using the Shapiro–Wilk test. The results indicated that normality was not met within groups ($p < 0.05$) and because of this, Wilcoxon tests were used for pair-wise within-group comparisons. The autistic research group responded 'Yes' significantly more frequently when the faux pas story included facework (avoidance process) ($T = 28$, $p = 0.018$), and so did the non-autistic research group ($T = 28$, $p = 0.02$).

6. DISCUSSION

6.1. Discussion of the original variables of the Faux Pas Recognition Test

In the present study, we first examined how autistic and non-autistic individuals performed on the original items of the Faux Pas Recognition Test. We found that the two groups achieved similar total scores as well as similar scores on the target questions (Q1–Q6). This suggests that autistic and non-autistic participants in our research may have similar theory of mind abilities, or alternatively, that they may relied on cognitive processes that differ from those of non-autistic participants, and employed different problem-solving, compensatory strategies that resulted in statistically indistinguishable performances. The results are also similar to the average performance reported by Gál (2015) in a typically developing research group (18–59-year-olds).

When analyzing the target questions within groups, we found that, irrespective of the neurotype, participants had significantly more difficulty identifying the intentions of the character committing the faux pas (Q4 – Intentions) and recognizing that the speaker's behavior was based on a false belief (Q5 – Belief) compared to detecting the faux pas (Q1), identifying the character who committed it (Q2), understanding its inappropriateness (Q3), or recognizing its emotional consequences (Q6). These results converge with the findings of Zalla et al. (2009), who reported that autistic individuals generally succeeded in detecting faux pas situations but performed less well when providing correct justifications regarding reasons and intentions or in recognizing the emotional consequences. In contrast to Zalla et al.'s findings, however, the same pattern was also observed in the non-autistic group in our study. Furthermore, even the lowest median score (in the autistic group, on the Belief question: Med = 0.7) still reflects a relatively good level of performance. In our sample, differences in the ability to attribute affective and cognitive mental states cannot be demonstrated between groups with different neurotypes.

6.2. Discussion of the results of the additional questions of the Faux Pas Recognition Test

To the Faux Pas Recognition Test, we added two additional questions designed to provide a deeper insight into participants' inferential processes and mental state attributions. Autistic and non-autistic groups responded very similarly to these questions, with only one response category showing a significant difference between groups: autistic individuals were significantly more likely than non-autistic research participants to respond that they did not know, or that it could not be known whether the character who committed the faux pas regrets their utterance. Thus, while both groups performed similarly on the original target questions aimed at 'Belief' and 'Empathy', autistic participants attributed affective mental state to the character committing a faux pas with significantly greater uncertainty than non-autistic

research participants. Autistic participants often justified their 'I do not know' answers by saying that the stories typically end before the speaker could learn that their false belief has led them to say something unintentionally hurtful.

However, within-group analysis of responses revealed that autistic participants answered significantly more often that the character who committed the faux pas would feel regret if they discovered that they had unintentionally offended the other person due to a false belief than that they did not know the answer or that it could not be determined. Moreover, autistic research participants appeared to rely on their inferential processes, and answered more frequently that the speaker would regret the faux pas than that the speaker is feeling regret in the present. This distinction is critical, as in the target stories the speaker is never explicitly confronted by the other participants with the fact that they said something awkward, nor realizes in half of the target stories that they said something awkward, making it hard for their current affective mental state to involve regret.

Non-autistic participants provided definite answers to the question of whether the character committing faux pas regrets saying what they said, as the response 'I don't know / it cannot be determined' was chosen significantly less frequently compared to all other response categories. However, no significant differences emerged among their other responses to this question at group level, which means it cannot be concluded whether they relied more on the explicitly stated content of the story (no mention of regret) or on the implicitly suggested interpretation (that the speaker would likely feel regret afterwards). In case of such ambiguous results, we considered it particularly important to examine the responses in light of whether the target stories contained avoidance process, or not.

By asking if the character committing the faux pas wants the other person to hear what was said or not, we aimed to examine whether participants would respond based on the explicit, literal meaning of the question (the addressee of the utterance is the listener), or whether their answers would instead reflect consideration of the speaker's false belief and that the offense was unintentional. Since the autistic group more frequently gave the answers 'Spoken to them' and 'Without offensive intent' compared to the other answers, but no significant difference was found between these two responses, we may infer that autistic participants identified the listener as the addressee of the utterance. However, depending on the specific content of the given faux pas story, they appeared to consider either the explicit or the implicit meaning of the question.

In contrast, non-autistic participants showed a significant tendency to consider the implicit meaning of the question rather than the explicit one. Their responses ('Without offensive intent')

emphasized that, although the listener was indeed the addressee, the faux pas utterance was not intended to be offensive.

6.3. Discussion of the Faux Pas Recognition Test results by presence of avoidance process (topic shifting) in the target questions

The faux pas stories were grouped according to whether they included an avoidance process or not. The results showed that autistic participants were more likely to consider the explicit meaning of the first added question (identifying the listener as the intended addressee of the utterance) when the story did not include facework, topic shifting. Also, they were more likely to respond with 'No' (that the speaker did not want to be heard), when the story did include avoidance process. These findings suggest that the presence of the avoidance process as a contextual cue influences whether autistic individuals attend to the explicit or the implicit content of an utterance (in this case, the content of the added question).

Non-autistic participants also appeared to consider the presence of topic shiftings, although only for one response category. Specifically, when the faux pas stories contained topic shifting, they were significantly more likely to respond that the character committing the faux pas did not want their partner to hear the utterance, compared to when no topic shifting was present.

As for the second added question, both the autistic and the non-autistic groups were significantly more likely to attribute affective mental state (regret) in the present to the character committing the faux pas when the target stories included an avoidance process than when they did not include one. This suggests that although in the target stories containing avoidance process it is not the speaker who committed the faux pas who performs the topic change in 4 out of 5 cases, and it is not explicitly revealed to the speaker that their false belief led them to make an offensive comment, research participants in both groups nevertheless attributed regret to the speaker (who committed the faux pas) when a character shifted the topic. This may indicate that topic shifting, as a contextual cue, facilitates affective mental state attribution.

Koskinen et al. (2021) found that autistic individuals manage facework differently in social interactions compared to non-autistic individuals. In contrast to their findings, our results indicate that autistic participants were sensitive to the presence of topic shifting as a form of facework, which requires complex inferential processing and socio-cognitive operations. Topic shifting is fundamentally an implicit interactional act, and despite the commonly reported difficulty in interpreting nonliteral meaning within the autistic population, participants in our study were able to understand the function of it.

Consistent with the findings of Sirota (2004) and Kissine et al. (2015), our results suggest that contextual cues, such as topic shifting, support autistic individuals in navigating social interactions.

7. CONCLUSION

The aim of this study was to gain deeper insight into the social cognitive abilities of autistic (N=12) and non-autistic (N=11) adults. We examined (a) whether affective and cognitive mental state attribution in autistic and non-autistic participants can be assessed more effectively through a modified version of the Faux Pas Recognition Test, and (b) whether the representation of an avoidance process as a contextual cue within the target stories of the Faux Pas Recognition Test influences the inferential and affective mental state attribution processes in autistic and non-autistic individuals.

Our results indicate that the modified Faux Pas Recognition Test provides a more comprehensive assessment of participants' social cognitive abilities, and that the presence of avoidance processes as contextual cues supports emotion attribution and consideration of the implicit meaning of an utterance.

To the best of our knowledge, this study represents an original contribution to the literature, as no prior research has systematically investigated how the presence of an avoidance process (topic shifting) within faux pas stories influences their interpretation. Also, several limitations should be acknowledged. The sample size was relatively small, which limits the universality of the findings. Second, all participants were Hungarian, raising the possibility that cultural or linguistic factors may have influenced the results; future research should therefore examine more diverse populations.

8. ACKNOWLEDGEMENTS

The authors would like to thank Márk Grastyán, Bachelor's graduate and Master's student in Psychology, for performing the statistical tests, the reviewers for their valuable comments and suggestions that helped improve the paper, and the research participants for their time and contribution.

REFERENCES

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. *Diagnostic and Statistical Manual of Mental Disorders*. https://doi.org/10.1176/APPI.BOOKS.9780890425596

Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. In *Mindblindness: An essay on autism and theory of mind.* The MIT Press.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger Syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*(5), 407–418. https://doi.org/10.1023/a:1023035012436

Belek, B. (2018). Autism and the Proficiency of Social Ineptitude: Probing the Rules of "Appropriate" Behavior. *Ethos*, *46*(2), 161–179. https://doi.org/https://doi.org/10.1111/etho.12202

Bosco, F. M., Tirassa, M., & Gabbatore, I. (2018). Why Pragmatics and Theory of Mind Do Not (Completely) Overlap. *Frontiers in Psychology*, *9*.

Bottema-Beutel, K., Kapp, S. K., Lester, J. N., Sasson, N. J., & Hand, B. N. (2021). Avoiding Ableist Language: Suggestions for Autism Researchers. *Autism in adulthood: challenges and management*, *3*(1), 18–29. https://doi.org/10.1089/aut.2020.0014

Brown, P., & Levinson, S. C. (1987). Politeness: Some Universals in Language Usage. In *Studies in Interactional Sociolinguistics*. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511813085

Duval, C., Piolino, P., Bejanin, A., Eustache, F., Desgranges, B. (2011). Age effects on different components of theory of mind. Consciousness and Cognition, 20. 3. 627-642.

Frith, C. D., & Frith, U. (2007). Social Cognition in Humans. *Current Biology*, *17*(16), R724–R732. https://doi.org/10.1016/J.CUB.2007.05.068

Gál, Z., Egyed, K., Pászthy, B., & Németh, D. (2011). Tudatelméleti deficit anorexia nervosában. [Impaired theory of mind in anorexia nervosa]. *Psychiatria Hungarica*, 26. 1. 12–25.

Gál, Z. (2015). *Age-related differences of the theory of mind, and it's relationship with working memory and sociometric position*. University of Szeged.

Goffman, E. (1967). *Interaction ritual: essays on face to face behavior*.

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., & Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain: a journal of neurology*, *125*(Pt 4), 752–764. https://doi.org/10.1093/brain/awf079

Győri, M. (2006). *Autism and Cognitive Architecture: Domain Specificity and Psychological Theorising on Autism (Neurocognitive Development and Impairments)*. Akadémiai Kiadó.

Howlin, P. (1996). *Autism: Preparing for Adulthood* (1st ed.). Routledge.

Ivaskó, L., & Németh T., E. (2002). Types and reasons of communicative failures. A relevance theoretical approach. *Modern Filológiai Közlemények 4*: 31–43.

Ivaskó, L. (2004). *A pragmatikai kompetencia szerepe a sikertelen diskurzusokban.* Doktori (PhD) disszertáció. Szegedi Tudományegyetem Nyelvtudományi Doktori Iskola, Szeged.

Ivaskó, L. (2000). Az elszólás szerepe a diskurzusban. In: Szécsényi Tibor szerk., LingDok 1. Nyelvész-doktoranduszok dolgozatai. SZTE Elméleti nyelvészet doktori program, Szeged. 171–177.

Kissine, M., Cano-Chervel, J., Carlier, S., De Brabanter, P., Ducenne, L., Pairon, M.-C., Deconinck, N., Delvenne, V., & Leybaert, J. (2015). Children with Autism Understand Indirect Speech Acts: Evidence from a Semi-Structured Act-Out Task. *PLOS ONE*, *10*(11), e0142191-. https://doi.org/10.1371/journal.pone.0142191

Koskinen, E., Stevanovic, M., Peräkylä, A. (2021). The Recognition and Interactional Management of Face Threats: Comparing Neurotypical Participants and Participants with Asperger's Syndrome. *Social Psychology Quarterly*, *84*(2), 132–154. https://doi.org/10.1177/01902725211003023

Lombardo, M. V., & Baron-Cohen, S. (2010). Unraveling the paradox of the autistic self. *Wiley interdisciplinary reviews. Cognitive science*, *1*(3), 393–403. https://doi.org/10.1002/wcs.45

Maylor, E. A., Moulson, J. M., Muncer, A., Taylor, L. A. (2002). Does performance on theory of mind tasks decline in old age? *British Journal of Psychology*, 93. 465-485.

Mezőlaki, R., & Ivaskó, L. (2024a). A faux pas (társalgási baklövés) jelenség pragmatikai értelmezési és felismerési lehetőségei az explicit nyelvi megformálás függvényében: 1. rész. [The possibilities of pragmatic interpretation and recognition of the faux pas phenomenon depending on explicit linguistic formulation: Part 1.] Magyar Nyelv, 120 (2): 150–158. https://doi.org/10.18349/MagyarNyelv.2024.2.150

Mezőlaki, R., & Ivaskó, L. (2024b). A faux pas (társalgási baklövés) jelenség pragmatikai értelmezési és felismerési lehetőségei az explicit nyelvi megformálás függvényében: 2. rész. [The possibilities of pragmatic interpretation and recognition of the faux pas phenomenon depending on explicit linguistic formulation: Part 2.] Magyar Nyelv, 120 (3): 277–290. https://doi.org/10.18349/MagyarNyelv.2024.3.277

Mezőlaki, R., & Ivaskó, L. (2024c). A faux pas (társalgási baklövés) jelenség pragmatikai értelmezési és felismerési lehetőségei az explicit nyelvi megformálás függvényében: 3. rész. [The possibilities of pragmatic interpretation and recognition of the faux pas phenomenon depending on explicit linguistic formulation: Part 3.] Magyar Nyelv, 120 (4): 409–417. https://doi.org/10.18349/MagyarNyelv.2024.4.409

Premack, D. & Woodruff, G. (1978). Does a chimpanzee have a theory of mind. *Behavioral and Brain Sciences*, *1*(04), 515–526.

Sirota, K. G. (2004). Positive politeness as discourse process: politeness practices of high-functioning children with autism and Asperger Syndrome. *Discourse Studies*, *6*(2), 229–251. https://doi.org/10.1177/1461445604041769

Sperber, D., & Wilson, D. (1995). *Relevance: communication and cognition*. Blackwell.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, *10*(5), 640–656. https://doi.org/10.1162/089892998562942

Stone, V. E., Baron-Cohen, S., Calder, A., Keane, J., & Young, A. (2003). Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia*, *41*(2), 209–220. https://doi.org/10.1016/s0028-3932(02)00151-3

Varga, E., Tényi, T., Fekete, S., & Herold, R. (2008). Mentalizációs deficit vizsgálata Faux Pas Teszttel skizofréniában. [The evaluation of mentalization deficit by the faux pas test in schizophrenia].*Neuropsychopharmacologia Hungarica*, *10*(2), 67–72.

Wang, Y., & Su, Y. (2006). Theory of mind in old adults: the performance on Happé's stories and faux pas stories. *Psychologia*, 49. 4. sz. 228-237.

Williams, D. M., & Happé, F. (2009). What Did I Say? Versus What Did I Think? Attributing False Beliefs to Self Amongst Children With and Without Autism. *Journal of Autism and Developmental Disorders*, *39*, 865–873. https://api.semanticscholar.org/CorpusID:12683956

Zalla, T., Sav, AM., Stopin, A., Ahande S, Leboyer M. (2009). Faux Pas Detection and Intentional Action in Asperger Syndrome. A Replication on a French Sample. *J Autism Dev Disord* 39, 373–382. https://doi.org/10.1007/s10803-008-0634-y

Marwa Mekni Toujani*

[1]Higher Institute of Applied Languages and
Computer Sciences of Beja (ISLAIB),
Jendouba University, Tunisia
[2]"Langue et Formes Culturelles" Research Lab,
Higher Institute of Languages,
University of Carthage (LR19ES19), Tunisia

DOI: 10.19090/slm.6.11

PERSONALITY TRAITS MODULATE VISUAL ATTENTION IN PICTURE NAMING/CATEGORISING TASKS

The main question motivating this study was to which extent overt visual selective attention depends on personality traits and on whether the participant is performing a linguistic or non-linguistic task. To answer this question, this study investigated the effects of personality traits and task demands (i.e., categorising and naming pictures) on visual selective attention. The visual selective attention of 64 Dutch participants was assessed using eye-tracking. Their personality traits were measured using the Big-Five model (BFI-2-S). The results indicated that higher degrees of extraversion, open-mindedness and agreeableness affect visual selective attention during language production tasks. That is, more extraverted, more open-minded and more agreeable people could be more distracted by their partners' pictures more than conscientious as well as emotionally negative people could be. Those who have more degrees of the latter personality traits selectively attend to their own pictures and overlook their partners'. Thus, this study contributed to the considerable number of studies investigating the effect of individual differences on cognitive processing demonstrating that some personality traits could affect selective visual attention in picture naming/categorising tasks.

Keywords: personality traits, selective visual attention, social attention, eye tracking, linguistic/non-linguistic tasks.

1. INTRODUCTION

This study was motivated by Matthews' (2008) claim that there is a correlation between personality traits and information processing. Attention can be considered the first stage in information processing as it makes the information read or heard available to the subsequent processing stages. More precisely, selective attention plays a pivotal role in cognitive processing in general and in language processing in particular. Due to the limited capacity of the processing system, selective attention seems an essential step in the successful processing of information read or heard. It filters unnecessary information from important information, which enhances the ability of the cognitive system. Selective attention could be either visual or auditory. This study focused on visual selective attention. Particularly, in this study, visual selective attention means that the information from naming the target picture (i.e., the participants' pictures) should be processed, yet information from partners' pictures, displayed on the same screen, should be ignored (see Method, for details).

Language production research demonstrates a close coupling between visual attention and speech planning. For example, Griffin and Bock (2000) show that speakers typically fixate referents

* email: marwaISLT@gmail.com

shortly before naming them. In spoken-language comprehension, Tanenhaus et al. (1995) similarly found that listeners direct their gaze toward objects in a display as the corresponding linguistic expressions unfold. Both sets of studies, however, focus on participants' attention to their own visual material rather than to those of an interaction partner. Visual attention to a partner's image has been examined more directly in Hoedemaker and Meyer (2019), who reported that while engaging in a turn-taking language task, speakers need not fixate a partner's picture to the same extent that they fixate their own. Likewise, Brehm et al. (2019) observed that fixations to a partner's picture occurred primarily after speech onset, suggesting that they may reflect residual processing time at the end of a trial rather than a functional requirement for production. Collectively, these studies indicate that although visual attention to a picture is often correlated with hearing or producing corresponding linguistic material, individuals vary considerably in the degree to which they attend to others' information in a partner's scene. This variation may be attributable to additional contextual or cognitive factors.

McIntyre and Graziano (2016) previously claimed that "these differences have been recognized in various forms as orientations, interests, and preferences, but empirical work examining these differences at a cognitive level is scarce" (p. 1258). This effect was probed in different joint action tasks (e.g., Dalmaso et al., 2021), which argues for "pervasive effects of real social stimuli on attentional mechanisms" (p. 469). The same conclusion was also drawn by Tufft and Gobel (2021), who asserted that both the interpretation of the social context and the gender of the social interactants could influence social visual attention. In fact, many factors are believed to alter selective attention. Those factors can be either bottom-up (low-level), such as image properties, or top-down (high-level) factors such as personality traits (Kaspar & Konig, 2012). Existing research provides ample evidence that attention can be modulated by individual differences. Working memory (Couperus et al., 2021), motivational relevance of the information (Cummins et al., 2016), emotion (Kaspar & Konig, 2012), personality traits (Wilson et al., 2016), etc., could be related to selective attention. In view of all this, the present study focuses on exploring whether personality traits are related to visual attention in non-communicative language tasks, i.e., naming (language production task) and categorising (language perception task).

This focus on personality traits is informed by Matsumoto et al.'s (2010) argument that personality might influence eye-movement patterns during visual information processing. Because personality traits reflect relatively stable dispositions, they may systematically shape how individuals attend to and engage with their environment. In fact, Wu et al. (2014) expanded on this topic as they explored the relationship between personality traits and social attention, and they found that "who a person is affects how they move their eyes to social stimuli" (p. 25). In other words, their results

confirmed that there is a relationship between personality and eye movements towards social stimuli. More precisely, their results revealed that while both extraversion and agreeableness relate positively "to the amount of attention committed to the eyes of others, openness to experience [negatively relates] to the attention committed to eyes" (Wu et al., 2014, p. 28).

As pointed out earlier, the present study aims to explore the effects of personality traits and task demands on selective visual attention in non-communicative interactions. It thereby aims to contribute to the literature on individual differences and selective visual attention in general and to the limited literature on personality traits and selective visual attention in particular. Therefore, the main questions raised in the present study were the following:

(1) Do personality traits direct visual selective attention? and

(2) Do those traits impact visual selective attention differently when someone is speaking (linguistic task) versus making a forced-choice button press categorising decision (non-linguistic task)?

Eye-tracking was used to explore whether selective attention in a non-communicative language production task (i.e., naming pictures) and perceptive tasks (i.e., categorising) is modulated by individual differences. In other words, patterns of fixations were used to probe whether, depending on their personality traits, participants pay selective attention towards their stimuli while naming or categorising them, or whether they are distracted by their partners' stimuli. Therefore, the Big Five Inventory–2 Short Form (BFI-2-S) (Soto & John, 2017) was used to measure participants' personality traits.

To address the aims of this study, the eye fixations of the speakers were analysed using multiple logistic regression models. In fact, the eye fixations were analysed to explore whether the current participants' personality traits modulate their attention to their partners' pictures while naming or categorising their own pictures. The dependent variable was the fixation to partners' stimuli. Within these models, it was assessed whether there is a main effect of each personality trait, for both production and perception tasks, and whether there is an interaction with the performed tasks. On the one hand, it was predicted that the likelihood of fixating the partner object would be higher in the production task than in the perception task. On the other hand, it was speculated that more extraverted, more open-minded, and more agreeable people would look at their partners' pictures more than conscientious and emotionally negative people would do in the speaking tasks. These predictions were formulated on the basis of the social motivations associated with each personality trait (Costandi, 2013).

2. METHOD

2.1 Participants

Sixty-four participants, who are native speakers of Dutch and whose ages range from 18 to 30, were selected from the database of the Max Planck Institute for Psycholinguistics. The number of participants was chosen following the G*Power tool (Faul et al., 2007), which suggested that 48 participants in each of the between-participant groups would allow observation of an effect f of 0.1 or larger (a small effect size in an ANOVA-style design) (Brehm et al., 2019, p. 3). All participants read and signed an informed consent form for behavioural experiments. All participants had normal or corrected to normal vision as well as normal hearing. They were invited as pairs to the lab, introduced to each other, and were randomly assigned to one of the experiment's trials and one of the computers. Participants were remunerated for taking part in the experiment.

2.2 Design

Participants were visually presented with 180 experimental trials, which were classified into six blocks. There were 30 trials per block. There were two people per session. Each participant sat in front of a computer screen on which two objects were presented simultaneously. In other words, each participant saw his/her own picture and his/ her partner's picture. Their partners' pictures could be similar to theirs (i.e., congruent) or different (i.e., incongruent). Trials involved naming pictures aloud or categorising the pictures via button pressing. Table 1 presents how the experimental materials were divided across trials, participants, and picture congruency. The order of trial display was counterbalanced, ensuring the appearance of each block type. Participants' eye fixations were monitored via a dual eye-tracking system (Eye-Link 1000 Desktop Mount).

Table 1. Experimental Materials Across Trials, Participants, and Picture Congruency.

Block	**Speaker's task**	**Co-speaker's task**	**Picture Congruency**
1	Name	Name	Similar
2	Name	Categorise	Similar
3	Categorise	Name	Similar
4	Name	Name	Different
5	Name	Categorise	Different
6	Categorise	Name	Different

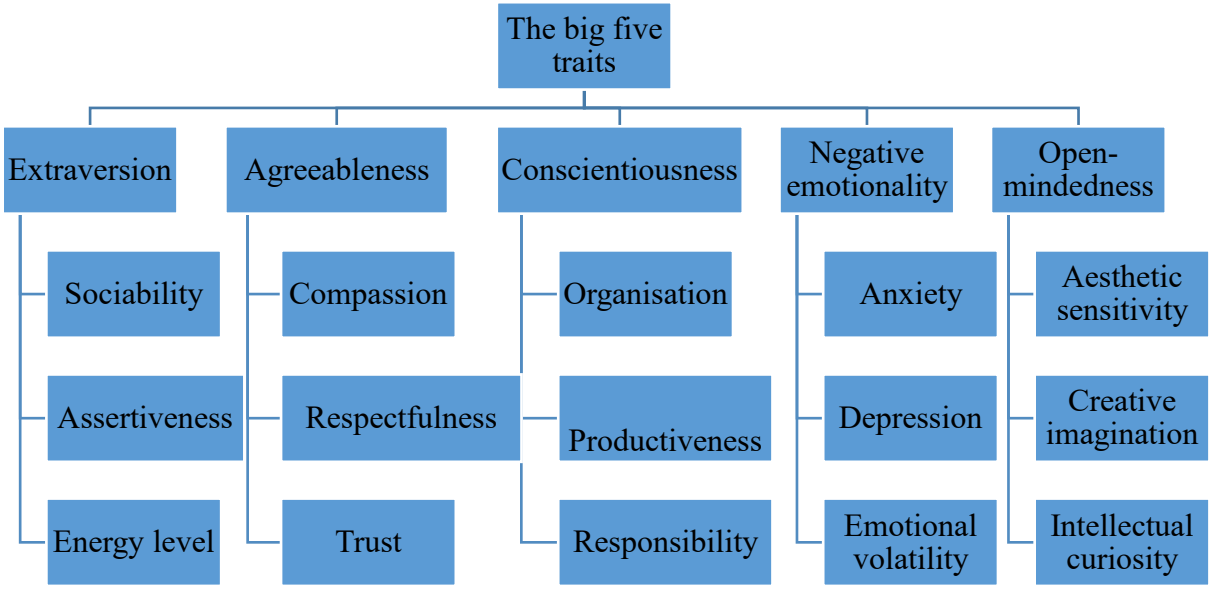
2.3 Materials

Thirty black-and-white line drawings, adopted from Severens et al.'s (2005) study on timed picture naming norms in Dutch, were the stimuli used in the present study (for a full list of stimuli, see Appendix A). From a semantic standpoint, the stimuli depicted either animate or inanimate entities. The

animate category comprised line drawings of animals, fruits, and vegetables, whereas the inanimate category included line drawings of clothing, furniture, and vehicles. All images were presented against a uniform white background.

Besides measuring their attention levels, information about participants' personality traits was also gathered. The Big Five Inventory–2 Short Form (BFI-2-S) was used (Soto & John, 2017; unpublished Dutch version translated, validated, and shared by Jaap Denissen, personal communication). It consists of 30 items aimed at testing the following personality categories: (1) Extraversion, (2) Agreeableness, (3) Conscientiousness, (4) Negative Emotionality, and (5) Open-Mindedness. Each category is divided into three sub-categories with two items per sub-category. In their third study, Soto and John (2017) examined the internal consistency of the BFI-2-S (Cronbach alpha). The replication of the reliability analyses is reported in Appendix B. Figure 1 presents the sub-categories within the five personality traits. Table 2 in Results section (Section 3) reports the range scores for each personality trait in this study.

Figure 1. Sub-categories Within the Five Personality Traits.



2.4 Procedure

As this experiment is the same as Experiment 2 reported in Brehm et al. (2019), it followed the same procedure. That is, before the beginning of the experiment, participants were introduced to each other. They received general instructions about the setup of the experiment, and they read a detailed

version of the instructions in Dutch. Participants were presented with 30 experimental items along with their corresponding names and categories. To collect responses while preventing them from hearing their partner's responses, participants wore microphone headsets in combination with noise-cancelling headphones that played pink noise. Their eye movements were calibrated using the EyeLink 1000, and participants were instructed to rest their forehead against a headrest to minimize unwanted movements. Finally, to familiarize them with the experimental stimuli and to verify the effectiveness of the sound-cancelling equipment, participants completed six practice trials: three for naming and three for categorisation. Each trial started with two fixation crosses: one per side for each participant. The display of those crosses lasted 1000 ms and was followed by a 500 ms blank screen. The stimuli were displayed for 2000 ms and then followed by a blank interval for 1500 ms. Participants were only instructed to name loudly their own pictures or categorise them by pressing buttons, and they did not receive any instructions to attend to their partners' pictures. Then, the following trial automatically started. Subsequent trials proceeded as the preceding one. At the end of the experiment, participants were given the full version of the Big Five Inventory–2 Short Form (BFI-2-S) translated into Dutch. They were instructed to rate on a 5-point Likert scale to what extent each statement applied to them. Participants indicated their level of agreement with each statement on a five-point Likert scale: (1) strongly disagree, (2) disagree a little, (3) neutral, (4) agree a little, or (5) strongly agree. The confidentiality of their responses was emphasized.

2.5 Data analysis

Mixed-effects models, namely the *lme4* package, version 1.1–18 (Bates et al., 2015) in R version 3.5.1 (R Core Team, 2018) were used to analyse the data. Primary dependent measures were log transformed proportion of fixating one or more times on the partner object during naming and categorising trials. Predictors were own task (name/categorise) and personality traits. All contrasts were coded as – 0.5 and 0.5. The analyses of the naming latencies are beyond the scope of this paper. Personality traits were manually coded, and their means were calculated using Excel.

All models were initially specified with maximal random effects, including random slopes for participants, trials (i.e., picture pairs), speaker task, speaker picture, partner picture (i.e., picture congruency), and partner task at various stages. However, the models did not converge fully. Consequently, only the random intercepts for speaker picture and partner picture were retained. Random effects that were estimated as zero (i.e., partner picture) were subsequently removed from all models. Thus, high correlations were observed between the interaction terms and the personality traits showing

effects. To address this, each personality trait was scaled to reduce multicollinearity. For the linear models, p-values were derived through model comparison.

3. RESULTS

Participants were significantly more likely to fixate their partner objects during naming trials than during categorising trials. While proportions of fixation were 0.0446 in categorising, they were 0.0794 in naming. This is confirmed with a logistic mixed-effect model (see subsequent analyses).

The likelihood of fixating the partner's objects at any point in the trials modulated by personality traits (see Table 2 for the range scores for each personality trait) was analysed using logistic mixed-effect models. The results revealed that more extraverted, more agreeable or more open-minded participants (see Tables 3, 4, and 5, respectively) were significantly more likely to fixate the partner's object. Nonetheless, conscientiousness and negative emotionality do not account for the likelihood of fixating partner's objects (see Tables 6 and 7, respectively).

Table 2. Descriptive Statistics for the BFI-2-S Domains in the Current Study.

Personality Traits	**Min**	**Mean**	**Max**
Extraversion Score	2.000	3.504	4.667
Agreeableness Score	2.500	3.977	5.000
Conscientiousness Score	2.167	3.362	4.833
Negative Emotionality Score	1.167	2.760	4.167
Open mindedness Score	2.333	3.697	5.000

Table 3. Logistic Mixed Effect Model for Odds of Fixating the Partner Object for Extraversion Trait in Both Tasks (Naming and Categorising).

Fixed effects	**Estimate**	**SE**	**z value**	**P($x^{2)}$**
Intercept	-5.01104	0.09932	-50.45	< 0.001
Speaker Task	-0.76063	0.01105	-68.84	< 0.001
Scaled Extraversion	2.26023	0.07013	32.23	< 0.001
Speaker Task X Scaled Extraversion	0.29520	0.01198	24.64	< 0.001

Random effects	**Term**	**Variance**
Speaker Task	Intercept	0.051
Scaled Extraversion	Intercept	0.098
Speaker Task X Scaled Extraversion	Intercept	-0.070

Table 4. Logistic Mixed Effect Model for Odds of Fixating the Partner Object for Agreeableness Trait in Both Tasks (Naming and Categorising).

Fixed effects	**Estimate**	**SE**	**z value**	**P($x^{2)}$**
Intercept	-4.99800	0.08131	-61.47	< 0.001
Speaker Task	-0.70844	0.01082	-65.49	< 0.001
Scaled Agreeableness	1.27548	0.06215	20.52	0.0393
Speaker Task X Scaled Agreeableness	0.15996	0.01466	10.91	< 0.001

Random effects	**Term**	**Variance**
Speaker Task	Intercept	0.006
Scaled Agreeableness	Intercept	-0.053
Speaker Task X Scaled Agreeableness	Intercept	-0.002

Table 5. Logistic Mixed Effect Model for Odds of Fixating the Partner Object for Openness Trait in Both Tasks (Naming and Categorising).

Fixed effects	**Estimate**	**SE**	**z value**	**P($x^{2)}$**
Intercept	-4.838052	0.098268	-49.233	< 0.001
Speaker Task	-0.670148	0.010418	-64.325	< 0.001
Scaled Openness	0.769552	0.087994	8.746	< 0.001
Speaker Task X Scaled Openness	-0.023849	0.009315	-2.560	0.01

Random effects	**Term**	**Variance**
Speaker Task	Intercept	0.022
Scaled Openness	Intercept	0.046
Speaker Task X Scaled Openness	Intercept	-0.004

Table 6. Logistic Mixed Effect Model for Odds of Fixating the Partner Object for Conscientiousness Trait in Both Tasks (Naming and Categorising).

Fixed effects	**Estimate**	**SE**	**z value**	**P($x^{2)}$**
Intercept	-4.84008	0.10982	-44.073	< 0.001
Speaker Task	-0.68931	0.01044	-66.047	<0.001
Scaled Conscientiousness	-0.02621	0.09599	-0.273	0.785
Speaker Task X Scaled Conscientiousness	-0.16657	0.01201	-13.873	< 0.001

Random effects	**Term**	**Variance**
Speaker Task	Intercept	0.031
Scaled Conscientiousness	Intercept	0.203
Speaker Task X Scaled Conscientiousness	Intercept	0.060

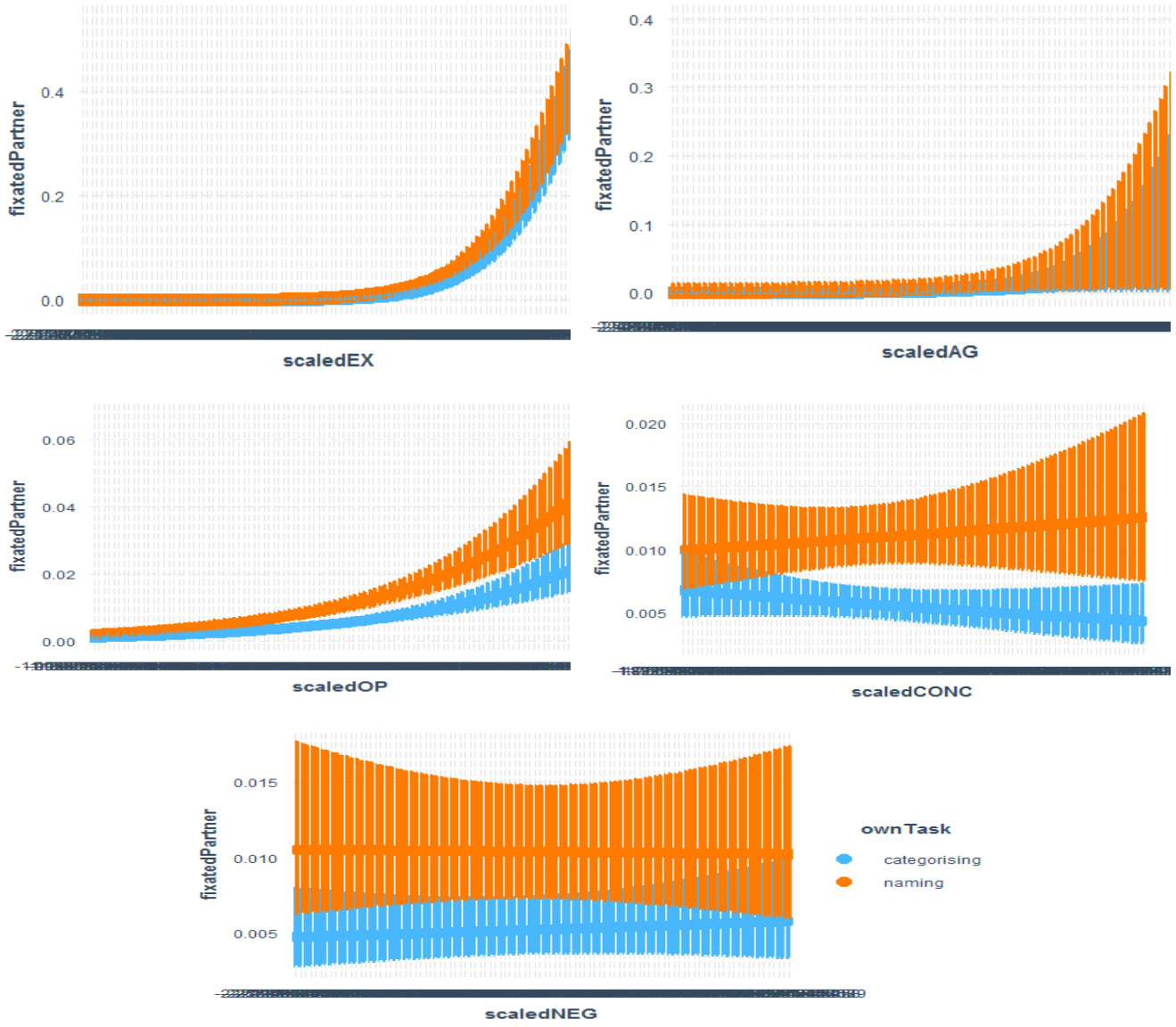
Table 7. Logistic Mixed Effect Model for Odds of Fixating the Partner Object for Negative Emotionality Trait in Both Tasks (Naming and Categorising).

Fixed effects	**Estimate**	**SE**	**z value**	**P($x^{2)}$**
Intercept	-4.89792	0.18165	-26.963	< 0.001
Speaker Task	-0.68087	0.01047	-65.054	< 0.001
Scaled Negative Emotionality	0.01992	0.09404	0.212	0.832
Speaker Task X Scaled Negative Emotionality	0.05412	0.01288	4.203	< 0.001

Random effects	**Term**	**Variance**
Speaker Task	Intercept	-0.024
Scaled Negative Emotionality	Intercept	0.100
Speaker Task X Scaled Negative Emotionality	Intercept	0.053

A closer look at the significant interactions reveals that the effect of personality traits differs according to the task performed by the speaker. In other words, the effect of personality traits was greater during the naming task than during the categorising task. For example, extraverts showed a higher tendency to fixate on their partners' pictures than introverts, with this effect being stronger during naming than during categorisation. The same conclusion applies to the remaining four personality traits. Figure 2 shows that the likelihood of fixating on the partner's pictures was influenced by the participants' own task, as indicated by interactions between task type and the five personality traits.

Figure 2. Likelihood of Fixations to Partner Pictures: Interactions Between Naming/Categorising Tasks and the Five Personality Traits.



4. DISCUSSION

The results of the present study indicated that participants with higher levels of extraversion, agreeableness, and openness were more likely to attend to their partners' irrelevant pictures than those with lower levels of these traits. There was a higher likelihood for these fixations during the naming task than during the categorisation task. These results suggest that extraversion, agreeableness, and open-mindedness are associated with a high tendency to be distracted by irrelevant stimuli. The effects of emotional negativity and conscientiousness on the frequency of fixations were not significant.

This study revealed that people fixate their partners' pictures during the naming task more than during the categorising task, which confirms the first prediction. On the one hand, this difference could be explained by the fact that the participants visualized themselves as partners in a dialogue and assumed that they should interact. This interaction was materialised through fixating the partner's pictures. Such a conclusion is supported by Costandi's (2013) claim that "people have a tendency to follow the gaze of others [in] joint action [activities]" (p. 86). In this case, the speaking activity forms a dyadic speech situation which requires joint action rather than a non-speaking situation. As explained by Brehm et al. (2019), the present conditions simulate "naming in chorus" situations (p. 9). On the other hand, the non-speaking activity requires the categorisation of pictures into living or non-living items. This activity is more of a conceptual/perceptual activity that does not normally necessitate joint action. Perceptual activities are usually individualistic activities that happen inside one's mind. The task also requires some sensorimotor action (i.e., pressing a button), so the participants' full attention was devoted to their own actions. Thus, while the naming task could evoke a "naming in chorus" situation, the categorising task (i.e., perceptual activity that necessitates sensorimotor action) could consume the participants' attention.

This "naming in chorus" setting may have promoted greater social attention among participants. In fact, Salley and Colombo (2016) claim that "the term social attention has been as a synonym for nonverbal social communication, or joint attention behaviors (p.3)". Nummenmaa and Calder (2009) define joint attention as "sharing a common focus of attention, such as an object or a spatial location with another individual" (p. 135). Therefore, participants' attention to their partners' pictures may have mirrored their attention to their partners' behaviour during the task. According to Salley and Colombo (2016), there are three categories of social attention. The last category, which "can be considered as attention (orienting, focusing and disengagement of visual systems) in the context of social streams of information" (p. 3) could fit within the present results. When they were engaged in the naming task, the present participants may have directed some attention to the social stimulus, which is, in this case, the partner's picture.

The previously formulated predictions regarding the influence of personality traits on visual attention were also confirmed. More extraverted, more open-minded, and more agreeable people would look at their partners' pictures more than conscientious and emotionally negative people would. These results could be explained by the underlying criteria of the personality traits. Consistent with the sociability criterion of extraverts, the compassion criterion of agreeable people, and the intellectual curiosity criterion of open-minded people, more extraverted, more agreeable, and more open participants could look at their partners' pictures more than their more introverted, less agreeable, and less open-minded peers would. Consistent with the organization criterion of conscientious people and the self-consciousness criterion of emotionally negative people, neither conscientiousness nor negative emotionality are expected to affect visual selective attention. This explanation is consistent with the previously outlined social attention framework. Individuals with more socially oriented traits (i.e., more extraverted, more open-minded, and more agreeable) are inclined to allocate greater attention to social stimuli compared to individuals with less socially oriented traits (i.e., conscientious and emotionally negative people).

The more socially oriented participants (i.e., more extraverted, more open, and more agreeable) tended to attend to objects that might otherwise have gone unnoticed, even when their partner's pictures were outside the focus of their covert attention (see Brehm et al., 2019, for further details). They deliberately chose to look at their partners' pictures. As discussed earlier, this choice was further influenced by both the performed task and personality traits. Partner pictures might have been seen as distractions for conscientious people, who are described as efficient, neat, and deliberate (Costandi, 2013). Emotionally negative people might have turned intentionally blind to their partners' pictures as they tended to focus their attention on their own tasks (Costandi, 2013). Conversely, more extraverted people who "[take] pleasure in social events" (Costandi, 2013, p. 69) may have perceived the situation as a genuine dialogue that warranted interactive engagement. For more agreeable people who are "cooperative and sympathetic" (Costandi, 2013, p. 69), fixating on the partners' pictures might be a form of cooperation in joint action activities "[including] paying attention to the physical environment (e.g., noticing the bend in the stairs)" (Hoedemaker & Meyer, 2019, p. 748). Finally, more open-minded people who are driven by their curiosity might have looked at their partners' picture to satisfy this need. This explanation was adopted by Brehm et al. (2019), who suggest that "fixations to partner objects might have occurred simply because time remained at the end of the trial and there were only two possible pictures to fixate" (p. 10).

The previously outlined explanations for the observed likelihood of fixating partners' pictures for more extraverted, more agreeable, and more open-minded people could lead to the speculation that these people simulated the dialogue situation. Simulation of events is the baseline of embodied cognition, where observed behaviour could have resulted "from dynamic interactions between brain, body and environment" (Costandi, 2013, p. 54). Thus, more extraverted, more agreeable, and more open people were more able to embody the dyadic situations. However, further studies are required to provide evidence of this speculation.

Lastly, this study explored the effects of personality traits on visual selective attention in a joint picture naming/categorisation task. Indeed, it has been proven that the social motives behind each personality trait could play a role in our attentional resources. Attention is the first gate for language processing, and the information that is attended will be processed, and the unattended information will be filtered out. Therefore, overloading the cognitive system with unnecessary information can result in processing deficiency (for details, cf. studies on Moses illusion and contradiction paradigms). Within this framework, individuals who are more extraverted, agreeable, and open-minded might initially appear at a disadvantage compared with those who are more conscientious or emotionally negative, as they tend to focus on their own pictures rather than attending to their partner's. However, the embodied scenarios observed in the task may mitigate or offset this potential disadvantage. Still, it is a speculation that needs to be further investigated in more psycholinguistic experiments where visual attention, accuracy, and instances of embodied language comprehension are explored regarding personality traits. This represents a potential limitation of the current research, as a single study of this kind cannot provide conclusive evidence to support such speculation. Hence, this study was an attempt to explore how social motives and contexts could influence cognitive processing in general and language processing in particular, which is in line with previous research on cognitive processes (e.g., Dalmaso et al., 2021; Tufft & Gobel, 2021). Ultimately, this study could be considered promising in terms of taking into consideration that personality traits could affect the result of language processing studies and could account for the variability of the surface-level contradictory results.

5. CONCLUSION

Intrigued by the difference in allocating attention to partners' pictures, this study looked up answers for the following two questions: (1) Do task demands affect selective visual attention in non-communicative language tasks? (2) Do personality traits affect selective visual attention? It found that there is a higher tendency for attending to partners' pictures during the naming task compared to the

categorising task. This attention is further influenced by some personal traits. Namely, the present research proves that while conscientious and emotionally negative people selectively attend to their own pictures, more extraverted, more agreeable, and more open-minded individuals are more likely to be distracted by their partners' pictures. This suggests that the attention processes are not generic, and personality traits can account for the observed variability.

6. ACKNOWLEDGEMENTS

I would like to thank the anonymous reviewers for their constructive suggestions which significantly improved the quality of the manuscript.

This experiment was conducted in Max Planck Institute for Psycholinguistics (MPI) as a part of the current author's three-month internship in the Department of Psychology of Language.

The naming and eye-movement data were published in Brehm, L., Taschenberger, L., & Meyer, A. (2019).

REFERENCES

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Brehm, L., Taschenberger, L., & Meyer, A. (2019). Mental representations of partner task cause interference in picture naming. *Acta psychologica*, *199*, 102888.

Costandi, M. (2013). 50 Human Brain Ideas You Really Need to Know. UK: Quercus Publishing.

- Couperus, J.W., Lydic, K. O., Hollis, J. E., Roy, J. L., Lowe, A.R., Bukach, C. M., & Reed, C. L. (2021). Individual Differences in Working Memory and the N2pc. *Front. Hum. Neurosci.* 15:620413. 10.3389/fnhum.2021.620413

Cummins, R. G., Gong, Z., & Kim, H.-S. (2016). Individual differences in selective attention to information graphics in televised sports. *Communication & Sport*, *4*(1), 102–120. https://doi.org/10.1177/2167479513517491

Dalmaso, M., Castelli, L., Scatturin, P., & Galfano, G. (2021). Can attitude similarity shape social inhibition of return? *Visual Cognition*, *29*(7), 463–474. https://doi.org/10.1080/13506285.2021.1922566

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, *39*(2), 175–191.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological science*, *11*(4), 274–279. https://doi.org/10.1111/1467-9280.00255

Hoedemaker, R. S., & Meyer, A. S. (2019). Planning and coordination of utterances in a joint naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(4), 732–752. https://doi.org/10.1037/xlm0000603

Kaspar, K., & König, P. (2012). Emotions and personality traits as high-level factors in visual attention: A review. *Frontiers in Human Neuroscience, 6,* Article 321. https://doi.org/10.3389/fnhum.2012.00321

Matsumoto, K., Shibata, S., Seiji, S., Mori, C., & Shioe, K. (2010). Factors influencing the processing of visual information from non-verbal communications. Psychiatry and Clinical Neurosciences, *64*, 299–308.

Matthews, G. (2008). *Personality and information processing: A cognitive-adaptive theory.* In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment, Vol. 1. Personality theories and models* (p. 56–79). Sage Publications, Inc. https://doi.org/10.4135/9781849200462.n3

McIntyre, M. M., & Graziano, W. G. (2016). Seeing people, seeing things: Individual differences in selective attention. *Personality and Social Psychology Bulletin, 42(9),* 1258-1271. https://doi.org/10.1177/0146167216653937

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. Trends in Cognitive Sciences, *13*(3), 135-143. https://doi.org/10.1016/j.tics.2008.12.006

R Core Team (2018). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Salley, B., & Colombo, J. (2016). Conceptualizing Social Attention in Developmental Research. *Social development (Oxford, England)*, *25*(4), 687–703. https://doi.org/10.1111/sode.12174

Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. Acta Psychologica, *119* (2), 159–187.

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113,* 117-143.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science (New York, N.Y.)*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Tufft, M. R. A., & Gobel, M. S. (2021). Gender and perceived cooperation modulate visual attention in a joint spatial cueing task. *Visual Cognition*, *30*(1–2), 6–27. https://doi.org/10.1080/13506285.2021.1976892

Wilson, K.E., Lowe, M.X., Ruppel, J., Pratt, J. & Ferber, S. (2016). The scope of no return: Openness predicts the spatial distribution of inhibition of return. Attention, Perception, & Psychophysics *78*, 209–217 https://doi.org/10.3758/s13414-015-0991-5

Wu, D. W.-L., Bischof, W. F., Anderson, N. C., Jakobsen, T., & Kingstone, A. (2014). The influence of personality on social attention. *Personality and Individual Differences*, *60*, 25–29. https://doi.org/10.1016/j.paid.2013.11.017

Appendix A: Items

Category	Dutch Name	English Name	Visual Complexity	Log Frequency	Number Syllables	Name Agreement	Normed RT
Non-living	broek	pants	16.6	1	1	91	804
Non-living	hoed	hat	9.87	1.61	1	100	862
Non-living	trui	sweater	12.4	1.30	1	77	977
Non-living	jas	coat	14.5	1.69	1	82	1067
Non-living	rok	skirt	10	1.49	1	85	1065
Non-living	tafel	table	12.9	2.39	2	100	734
Non-living	bed	bBed	14.5	2.48	1	97	830
Non-living	stoel	armchair	12.4	1.18	2	73	1017
Non-living	kast	dresser	21.3	1.68	1	81	1082
Non-living	bureau	desk	18.2	1.94	2	87	1118
Non-living	auto	car	10.4	2.32	2	100	694
Non-living	fiets	bike	24.4	1.68	1	100	759
Non-living	vliegtuig	plane	17.3	1.72	2	100	831
Non-living	bus	bus	23.3	1.76	1	100	955
Non-living	boot	boat	17.7	1.83	1	87	1061
Living	slang	snake	23.9	1.43	1	100	772
Living	olifant	elephant	24.5	1	3	100	788
Living	konijn	rabbit	12.2	1.36	2	100	813
Living	muis	mouse	14	1.32	1	86	997
Living	bij	bee	13	3.58	1	54	1252
Living	appel	pple	9.37	1.38	2	100	856
Living	citroen	lemon	9.68	1.04	2	97	901
Living	peer	pear	19.3	1	1	100	914
Living	aardbei	strawberry	17.3	0.70	2	100	917
Living	kers	cherry	5.8	0.70	1	89	1019
Living	paddestoel	mushroom	9.51	0.95	3	95	872
Living	wortel	carrot	14.1	1.57	2	94	885
Living	tomaat	tomato	9.57	0.95	2	97	912
Living	sla	lettuce	17.6	0.48	1	66	986
Living	mais	corn	16.7	0.90	1	92	1005

Appendix B: Reliability Analyses

The BFI-2-S reliability analyses were carried out. They revealed that the BFI-2-S facets and domains scales showed moderate to strong reliability.

Reliability Coefficients of the BFI-2-S Facets and Domains in this Study

Facets/**Domains**	Alpha
Sociability	0.71
Assertiveness	0.72
Energy level	0.71
Extraversion	0.70
Compassion	0.68
Respectfulness	0.71
Trust	0.70
Agreeableness	0.69
Organization	0.71
Productiveness	0.72
Responsibility	0.70
Conscientiousness	0.70
Anxiety	0.71
Depression	0.71
Emotional volatility	0.72
Negative Emotionality	0.70
Aesthetic sensitivity	0.71
Intellectual curiosity	0.72
Creative imagination	0.72
Open-Mindedness	0.70
Mean	0.72

Mauro Viganò,[1*] Giulia Gilardone,[2] Dario Cassinelli,[2] & Carlo Cecchetto[3,4]

DOI: 10.19090/slm.6.9

[1]UMR7320 Bases, Corpus, Langage,
CNRS & Université Côte d'Azur, Nice, France
[2]Department of Neurorehabilitation Sciences,
Casa di Cura IGEA, Milan, Italy
[3]UMR7023 Structures Formelles du Langage,
CNRS & Université Paris 8, Paris, France
[4]Department of Psychology, University of Milan-Bicocca, Milan, Italy

THE CHALLENGE OF VARIABILITY FOR SYNTACTIC ACCOUNTS OF AGRAMMATISM: A STUDY ON FEATURE DISSIMILARITY IN ITALIAN RELATIVES

This study explores syntactic processing in ten Italian-speaking individuals with agrammatism, testing predictions from two competing linguistic accounts: the Trace Deletion Hypothesis and Generalized Minimality. We examined subject, object, and passive relatives in comprehension (sentence-to-picture matching task) and production (sentence priming task), manipulating number mismatch as a possible facilitation factor. According to Generalized Minimality, mismatch should reduce intervention effects in object relatives, while the Trace Deletion Hypothesis predicts uniform impairment of derived structures. Comprehension results showed an overall subject advantage, with mismatch-related facilitation observed in a subset of participants. However, high interindividual variability challenged strong generalisations. No clear performance pattern emerged in production, possibly due to concurring deficits and task-related limitations. Passive relatives, expected to be easier in line with studies on language acquisition, were as impaired as object relatives, supporting a specific deficit with passives in agrammatism. Linguistic theory played a central role in accounting for the variability observed, with Generalized Minimality emerging as the most explanatory, yet not unchallenged account.

Keywords: aphasia; agrammatism; relative clauses; morphosyntax; relativized minimality; language assessment; syntactic movement; trace deletion hypothesis; sentence comprehension; sentence production.

1. INTRODUCTION

Agrammatism, a possible symptom of aphasia, is characterized by nonfluent speech with short utterances and frequent omissions of morphological elements. Since Caramazza and Zurif's (1976) seminal paper, it has been shown that individuals with agrammatism also display specific comprehension difficulties, especially with noncanonical reversible sentences, supporting the idea that agrammatism constitutes a robust clinical construct (e.g., Grodzinsky, 2000).

Theoretical linguistics has greatly advanced the study of agrammatism, aiming to identify the source of the deficit. In this context, sentences with relative clauses have been a critical structure, examined by several studies across languages. The present study aims at comparing two generative

* Corresponding author: Mauro Viganò; email: dott.mauro.vigano@gmail.com

** Author contributions: MV conceived the research protocol, conducted the analyses and interpretation of the results, and wrote the manuscript; GG and DC carried out data collection; CC supervised the project and contributed to data interpretation; all authors reviewed and approved the final version of the manuscript. We thank Arianna Ravidà for creating the images used in the study.

linguistic accounts of agrammatism by investigating the comprehension and production of relatives in Italian-speaking individuals with post-stroke agrammatism.

1.1 The Trace Deletion Hypothesis

Grodzinsky (2000) proposed the Trace Deletion Hypothesis, which claims that individuals with agrammatism cannot represent traces of syntactic movement. In object relatives, such as in (1), the deletion of the trace disrupts the ability to represent the dependency between the moved constituent and its original position, where the thematic role is assigned.

(1) John watches the cat_i that the dog chases t_i

As a consequence, people with agrammatism rely on an extra-syntactic agent-first strategy for interpretation. In (1), since the trace is not represented, the Agent thematic role is assigned both to 'the cat', following the agent-first strategy, and to 'the dog', since it is in the standard subject position.[20] As a result, people with agrammatism are expected to perform at chance in comprehension, since both DPs equally compete for the Agent role. Conversely, in subject relatives, such as in (2), the trace is also deleted, but no difficulty arises in thematic role assignment: the moved constituent and its trace are contiguous, and the agent-first strategy converges with the standard syntactic processing.

(2) John watches the dog_i that t_i chases the cat

The Trace Deletion Hypothesis can thus be defined as a structural account and, assuming a representational deficit, yields deterministic predictions about agrammatic performance.

1.2 The Generalized Minimality account

Different studies have criticised this structural account, as it cannot explain the wide variability observed in agrammatic performance and its correlation with other cognitive resources, particularly short-term memory (e.g., Caplan et al., 2007).

Grillo (2008) proposed an alternative hypothesis, called Generalized Minimality, to account for both the variability and the regularities found in agrammatism. According to Grillo, the underlying cause of morphosyntactic deficits in aphasia is a reduction in cognitive resources, which triggers a Feature Underspecification. This is defined as an impoverished representation of the full array of morphosyntactic features associated with syntactic categories, and, in combination with the locality

[20] According to the Trace Deletion Hypothesis, the movement of the subject from Spec,VP to Spec,IP (Koopman & Sportiche, 1991) is not considered problematic for people with agrammatism.

restrictions of syntactic representation, it accounts for the asymmetries observed in agrammatic production and comprehension.

Relativized Minimality (e.g., Rizzi, 2018) is the key locality constraint for this account, as it captures the idea that syntactic relations are disrupted if there is a suitable intervener. Here follows a formal definition of the principle:

> In [. . . X . . . Z . . . Y . . .] a dependency between X and Y is disrupted when
> (i) Z intervenes in the dependency between X and Y (i.e., X c-commands Z and Z c-commands Y), and
> (ii) Z matches X in terms of relevant syntactic features.[21]
> (iii) The degree of disruption is a function of the featural distinctness of X with respect to Z.

In subject relatives, such as in (2), no syntactic category intervenes in the dependency between the head of the relative and its trace. However, in object relatives, such as in (1), the subject DP of the relative clause structurally intervenes in the dependency. In adult unimpaired processing, this intervention does not trigger a minimality effect because the head of the relative and the intervening DP differ in their featural configuration: the head bears a wh-feature that triggers movement. In agrammatism, however, Feature Underspecification may lead to the loss of the wh-feature, resulting in identical featural configurations between the head and the intervener. This triggers a minimality effect and causes impaired performance.

As noted above, Feature Underspecification is defined as a deficit linked to limited processing resources. Since resource availability varies not only across but also within individuals (e.g., due to fatigue, attentional fluctuations, or cognitive load), the occurrence and degree of underspecification may fluctuate accordingly. This account thus better explains performance variability than strictly structural hypotheses such as the Trace Deletion Hypothesis. At the same time, Generalized Minimality captures the structural selectivity of the linguistic impairment, since underspecification leads to deficits only in specific syntactic contexts constrained by locality conditions, unlike purely processing-based accounts (e.g., Caplan et al., 2007).

A key property of this account is that it predicts that dissimilarities in morphosyntactic features between the two DPs in object relatives, such as grammatical number in (3), can provide a facilitation.

(3) John watches the cat_i that the dogs chase t_i

[21] Syntactic features are defined as relevant if they define syntactic positions, by licensing applications of External and Internal Merge (e.g., features involved in argument-verb agreement, triggers of movement to functional positions, etc.).

Indeed, even in cases of wh-feature underspecification, the DPs 'the cat' and 'the dogs' present distinct featural configurations, one being singular and the other being plural.

1.3 Experimental evidence

The comprehension asymmetry between subject and object relatives has been found in people with agrammatic aphasia across languages: English (see Grodzinsky, 2000 for a review), German (Adelt et al., 2017, 2020), Greek (Varlokosta et al., 2014), Hebrew (Friedmann, 2008), Italian (Grillo, 2008; Gilardone et al., 2023), Russian (Friedmann et al., 2010), and Serbo-Croatian (Lukatela et al., 1995).

Moreover, studies on language acquisition have reported robust evidence for a facilitation effect induced by feature dissimilarities between the head of the relative and the intervener (Belletti et al., 2012; Adani et al., 2010). As for agrammatism, results are more mixed. Friedmann and colleagues (2010) found no facilitation for gender and case features in Russian. Terzi and Nanousi (2018) observed a facilitation for gender mismatch in Greek, though notably also in subject relatives, which do not involve intervention. Adelt and colleagues (2017, 2020) found that, in German, case marking provides a better facilitation in comparison with number marking, as it provides earlier disambiguation within the sentence. Hence, they concluded that people with aphasia require a high degree of feature dissimilarity to successfully process structures characterised by intervention.

To address the controversy surrounding previous findings, this study is the first to investigate the comprehension and production of subject and object relatives in Italian agrammatic speakers, focusing on the unexplored facilitation effect of number dissimilarity. In particular, the Trace Deletion Hypothesis predicts a selective impairment in object relative clauses, with no effect of number-feature mismatch between the two DPs in the sentence. In contrast, Generalized Minimality also expects increased difficulty with object relatives, but crucially predicts a modulation of performance, whereby grammatical number mismatch yields a facilitation effect.

2. METHODS

The study was conducted at the Department of Neurorehabilitation Sciences, Casa di Cura IGEA in Milan (Italy). The protocol was approved by the local ethics committee (Milan Area 2; Resolution: 743_2022bis). All participants gave their informed consent in a manner appropriate to their abilities.

2.1 Participants

Ten participants were selected from the medical records of the clinic according to these criteria. Inclusion criteria: ischemic or haemorrhagic focal brain injury in the left hemisphere, right-handedness,

chronic phase (time post-onset > 12 months), nonfluent aphasia with agrammatism, and Italian as native language. Exclusion criteria: neurodegenerative diseases, non-linguistic cognitive impairments, and affective or psychiatric disorders.

The characteristics of the participants are reported in Table 1. Neuroimaging data were not available, and information regarding lesion sites was retrieved from medical records.

The experimental protocol was also administered to a control group of 20 native Italian speakers with no history of neurological or psychiatric conditions, nor of language or learning disorders. Control participants were recruited among the relatives of individuals with aphasia and the clinical staff, and were matched to the aphasic group for gender, age, and educational level.

Table 1. Characteristics of the participants.

ID	**Sex**	**Age** *years*	**Ed.** *years*	**Hand.**	**Aetiology**	**Lesion site**	**TPO** *month*
I01	F	68	9	R	IS - MCA occlusion	LH - cortical (fronto-temporal)	132
I02	F	55	13	R	IS - MCA occlusion	LH - cortical and subcortical (fronto-insular, nucleo-capsular)	164
I03	M	48	17	R	IS - MCA occlusion	LH - cortical (fronto-temporal)	146
I04	F	52	17	R	IS - MCA occlusion	LH - cortical (fronto-temporal)	123
I05	F	75	16	R	IS - cardioembolic	LH - cortical and subcortical (fronto-temporo-insular)	43
I06	F	71	8	R	IS - cardioembolic	LH - cortical and subcortical (fronto-temporal)	79
I07	M	51	18	R	IS - MCA occlusion	LH - cortical and subcortical (frontal and nucleo-capsular)	32
I08	F	54	17	R	IS - ICA dissection	LH - cortical (fronto-temporo-parietal)	76
I09	F	60	13	R	IS - MCA occlusion	LH - cortical (fronto-temporo-parietal)	155
I10	F	65	18	R	IS - ICA dissection	LH - cortical (fronto-temporo-parietal)	73

Notes. M=male; F=female; R=right-handed; IS=Ischemic stroke; MCA=Middle Cerebral Artery; ICA=Internal Carotid Artery; LH=Left Hemisphere; TPO=Time Post-Onset.

2.2 Standardised aphasia assessment

Each participant underwent standardised assessment of language and verbal short-term memory, using the following tests: the Italian version of the Aachen Aphasia Test (AAT; Luzzatti et al., 2023), the auditory grammaticality judgment task, the memory task of the Batteria per l'Analisi dei Deficit Afasici (BADA; Miceli et al., 1994), and both forward and backward Digit Span (Monaco et al., 2013).

2.3 Experimental materials

Examples of the stimuli used to test the comprehension and production of sentences with relative clauses are provided in (4–8). The protocol included subject (4), object (5), and passive (6) relatives with only singular DPs, as well as subject (7) and object (8) relatives featuring number mismatch between the two DPs in the sentence. Examples of the corresponding images used in the tasks are shown in Figure 1.

To minimize lexical processing demands, the stimuli included only eight nouns for human male-gendered entities: *Bambino* – 'child', *cuoco* – 'cook', *dottore* – 'doctor', *ladro* – 'burglar', *mago* – 'magician', *prete* – 'priest', *soldato* – 'soldier', and *vecchietto* – 'oldie'. For the same reason, we selected only six transitive action verbs: *Applaudire* – 'to applaud', *bagnare* – 'to wet', *colorare* – 'to paint', *colpire* – 'to hit', *inseguire* – 'to chase', and *spingere* – 'to push'. All sentences were reversible and items were pseudorandomised in the protocol.

(4) *Indica il mago$_i$ che* t$_i$ *spinge il cuoco*

'Point to the magician$_i$ that t$_i$ pushes the cook'

(5) *Indica il mago$_i$ che il cuoco spinge* t$_i$

'Point to the magician$_i$ that the cook pushes t$_i$'

(6) *Indica il mago$_i$ che* t$_i$ *viene spinto* t$_i$ *dal cuoco*

'Point to the magician$_i$ that t$_i$ is pushed t$_i$ by the cook'

(7) *Indica i maghi$_i$ che* t$_i$ *spingono il vecchietto*

'Point to the magicians$_i$ that t$_i$ push the oldie'

(8) *Indica il soldato$_i$ che i vecchietti spingono* t$_i$

'Point to the soldier$_i$ that the oldies push t$_i$'

Figure 1. Example of pictures for (a) subject, object and passive relatives with singular DPs, (b) subject relatives with number mismatch, and (c) object relatives with number mismatch.

2.4 Experimental procedures

Comprehension of relatives was tested using a sentence-to-picture matching task based on character selection. The protocol included 80 sentences with relative clauses: 24 subject relatives, 24 object relatives, and 12 passive relatives. Specifically, 12 subject, 12 object, and 12 passive relatives featured only singular DPs (examples 4-6), while the remaining 12 subject and 12 object relatives involved number mismatch (examples 7-8). Participants listened to each sentence and selected the correct referent in a picture showing three characters or groups of characters (see Figure 1).

Production was tested using a sentence priming task with the same type of stimuli used in the comprehension task. The protocol included 12 subject, 12 object, and 6 passive relatives. To facilitate priming, different types of relatives were tested in separate blocks. Participants were presented with a picture and a corresponding model sentence; then, they were asked to describe a series of pictures by using the same syntactic structure.

The protocol yielded ceiling performance in the control group, supporting its validity.

2.5 Statistical analyses

Given the limited number of participants, we conducted comparisons across experimental conditions using the nonparametric Wilcoxon signed-rank test. However, since no comparisons reached statistical significance, we provide instead individual results and group-level tendencies using descriptive statistics. To estimate the magnitude of contrasts relevant to the study, we report both differences in percentage accuracy (Δ) and Cohen's h coefficient (Cohen, 1988).

3. RESULTS

3.1 Standardised aphasia assessment

Table 2 reports individual results from the standardised assessment. Performance on the linguistic tests showed interindividual variability. Overall, all participants reported poor performance in short-term and working memory tasks.

Table 2. Performance of the participants at standardised assessment.

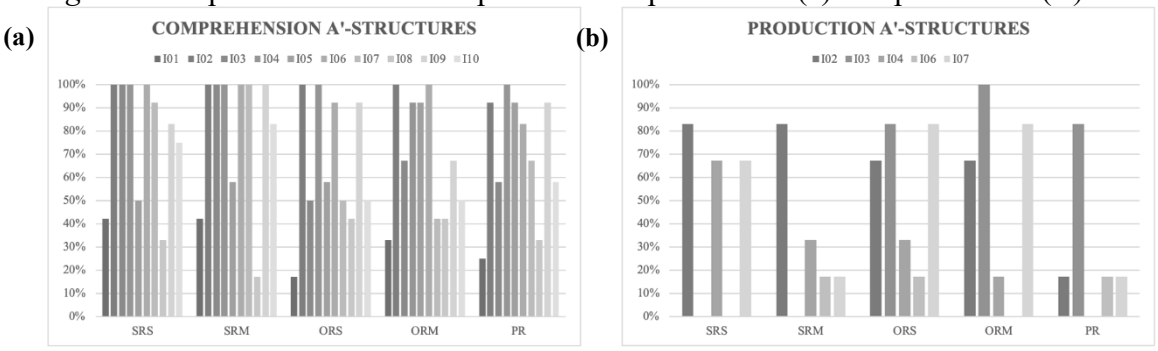
ID	I01	I02	I03	I04	I05	I06	I07	I08	I09	I10
AAT										
Communication	2	2	2	2	1	3	2	1	2	2
Articulation	3	2	2	2	1	3	3	2	2	1
Automated lang.	3	3	4	3	1	4	4	4	3	4
Semantic struct.	3	4	3	3	0	4	4	1	3	4
Phonemic struct.	4	4	2	2	0	3	4	1	4	3
Syntactic struct.	2	1	1	1	0	2	1	1	1	0
Token test	56	57	43	62	56	55	55	43	48	49
Repetition	46	49	56	45	39	52	56	47	50	53
Written lang.	42	48	55	47	41	40	60	51	57	60
Naming	45	52	62	54	40	50	61	47	52	54
Comprehension	51	55	63	52	66	38	58	41	55	58
H. Profile	46.56	50.94	55.64	50.25	43.88	48.15	58.08	46.95	52.18	54.72
BADA										
Gram Jud	20	6	7	5	1	8	2	3	13	0
Memory	0	0	5	0	0	3	3	0	3	1
DIGIT SPAN										
Forward	3	3	3	3	3	4	3	3	3	3
Backward	2	2	3	3	2	2	2	2	2	2

Notes. H. Profile=averaged level of performance. Gram Jud=BADA grammaticality judgements. na=not administrable.

3.2 Comprehension of relatives

Results of the comprehension task for subject, object, and passive relatives are presented in Table 3 and Figure 2a. Performance is characterised by high interindividual variability. Overall, three out of ten participants (I02, I04, I06) performed almost at ceiling with all types of relatives, while five (I01, I03, I07, I09, I10) performed better on subject than on object relatives. Number mismatch in object relatives was associated with improved performance in two participants (I03 and I05), and with decreased performance in one participant (I09).

Figure 2. Proportion of correct response in comprehension (a) and production (B) tasks.



Notes. SRS=subject relatives with singular arguments; SRM=subject relatives with number mismatch; ORS=object relatives with singular arguments; ORM=object relatives with number mismatch; PR=passive relatives.

Table 3. Individual performance in the comprehension task for relatives.

ID	I01	I02	I03	I04	I05	I06	I07	I08	I09	I10	Mean(SD)
SRS	42%	100%	100%	100%	50%	100%	92%	33%	83%	75%	78%(26)
SRM	42%	100%	100%	100%	58%	100%	100%	17%	100%	83%	80%(30)
ORS	17%	100%	50%	100%	58%	92%	50%	42%	92%	50%	65%(29)
ORM	33%	100%	67%	92%	92%	100%	42%	42%	67%	50%	68%(26)
PR	25%	92%	58%	100%	92%	83%	67%	33%	92%	58%	70%(26)
TOT SR	42%	100%	100%	100%	54%	100%	96%	25%	92%	79%	79%(28)
TOT OR	25%	100%	58%	96%	75%	96%	46%	42%	79%	50%	67%(26)
Δ SR M-S	0%	0%	0%	0%	8%	0%	8%	-17%	17%	8%	3%
***h* SR M-S**	0	0	0	0	0.167	0	0.586	-0.390	0.841	0.206	0.061
Δ OR M-S	17%	0%	17%	-8%	33%	8%	-8%	0%	-25%	0%	3%
***h* OR M-S**	0.390	0	0.340	-0.586	0.818	0.586	-0.167	0	-0.645	0	0.071
Δ SR-OR	17%	0%	42%	4%	-21%	4%	50%	-17%	13%	29%	12%
***h* SR-OR**	0.356	0	1.403	0.411	-0.440	0.411	1.243	-0.356	0.362	0.623	0.273
Δ PR-OR	0%	-8%	0%	4%	17%	-13%	21%	-8%	13%	8%	3%
***h* PR-OR**	0	-0.586	0	0.411	0.462	-0.430	0.423	-0.172	0.362	0.167	0.072

Notes. SRS=subject relatives with singular DPs; SRM=subject relatives with number mismatch; ORS=object relatives with singular DPs; ORM=object relatives with number mismatch; PR=passive relatives; M=number mismatch; S=all singular DPs; Δ=difference in accuracy (%); *h*=Cohen's *h*.

3.3 Production of relatives

Results of the production task for subject, object, and passive relatives are presented in Table 4 and Figure 2b. Five out of ten participants did not complete the task due to severe extra-syntactic deficits. Of the remaining five participants, two (I02 and I04) performed better on subject relatives than on object relatives, while two others (I03 and I07) showed the opposite pattern. Additionally, one participant (I03) reported better performance in the presence of number mismatch, while two participants (I04 and I07) showed the inverse pattern.

Table 4. Individual performance in the production task for relatives.

ID	I01	I02	I03	I04	I05	I06	I07	I08	I09	I10	Mean(SD)
SRS	na	83%	0%	67%	na	0%	67%	na	na	na	43%(40)
SRM	na	83%	0%	33%	na	17%	17%	na	na	na	30%(32)
ORS	na	67%	83%	33%	na	17%	83%	na	na	na	57%(30)
ORM	na	67%	100%	17%	na	0%	83%	na	na	na	53%(43)
PR	na	17%	83%	0%	na	17%	17%	na	na	na	27%(32)
TOT SR	na	83%	0%	50%	na	8%	42%	na	na	na	37%(34)
TOT OR	na	67%	92%	25%	na	8%	83%	na	na	na	55%(37)
Δ SR M-S	na	0%	0%	-33%	na	17%	-50%	na	na	na	-13%
***h* SR M-S**	na	0	0	-0.680	na	0.841	-1.070	na	na	na	-0.278
Δ OR M-S	na	0%	17%	-17%	na	-17%	0%	na	na	na	-3%
***h* OR M-S**	na	0		-0.390	na	-0.841	0	na	na	na	-0.067
Δ SR-OR	na	17%	-92%	25%	na	0%	-42%	na	na	na	-18%
***h* SR-OR**	na	0.390	-2.556	0.524	na	0	-0.897	na	na	na	-0.370
Δ PR-OR	na	-50%	-8%	-25%	na	8%	-67%	na	na	na	-28%
***h* PR-OR**	na	-1.070	-0.255	-1.047	na	0.255	-1.459	na	na	na	-0.586

Notes. SRS=subject relatives with singular DPs; SRM=subject relatives with number mismatch; ORS=object relatives with singular DPs; ORM=object relatives with number mismatch; PR=passive relatives; M=number mismatch; S=all singular DPs; Δ=difference in accuracy (%); *h*=Cohen's *h*.

4. DISCUSSION

The results of the study revealed substantial variability among participants in both comprehension and production tasks. This variability is consistent with the heterogeneity of the sample in terms of aphasia severity, although all participants presented a profile compatible with nonfluent agrammatic aphasia. As anticipated in the introduction, performance variability may also reflect inter-individual fluctuations due to clinical factors, attentional variability, or task-induced cognitive load. Furthermore, as will be discussed in detail below, differences in short-term memory resources and in the preservation versus impairment of morphological processing may, at least partially, account for the observed variations in performance.

Consequently, neither the predictions of the Trace Deletion Hypothesis (Grodzinsky, 2000) nor those of the Generalized Minimality account (Grillo, 2008) appear to be fully confirmed by overall group trends, at least at first glance. In the following sections, we will examine the results in greater detail to identify relevant patterns within this variability.

4.1 Subject advantage in comprehension

In the comprehension task, the prediction of better performance with subject relatives is overall confirmed. asExcluding the three participants with ceiling accuracy across all structures, five out of seven participants showed a subject advantage. This result is compatible with both the Trace Deletion Hypothesis and with Generalized Minimality. However, the variability in the data suggests that the latter approach, combining structural and resource-based components, is more convincing. Indeed, cases of

good performance with both subject and object relative in agrammatism have been previously reported in the literature (e.g., Nerantzini et al., 2014 for Greek).

As for the two participants who presented the unexpected response pattern, with higher accuracy for object relatives, an individualised explanation can be attempted. For I08, the difference is very small, and the accuracy is below chance for both structures. Conversely, I05 performed at chance on all structures except for object relatives with number mismatch (92%). Given her severely impaired repetition (AAT) and short-term memory skills (BADA and Digit Span), it is plausible that I05 relied on a partial-processing strategy, using subject-verb agreement cues at the end of the sentence to infer the correct referent.[22] While this superficial heuristic could, in principle, be employed by other, less impaired, patients, the performance patterns suggest it was uniquely adopted by I05.

4.2 Number mismatch in comprehension

As noted above, three of the ten participants performed with very high accuracy across all structures, making it impossible to assess any mismatch-related facilitation. Among the remaining seven, only two participants showed a potential facilitation effect. I03 performed at chance level with object relatives without number mismatch but above chance when mismatch was present (67%), while performing at ceiling on subject relatives. I05 also showed improved performance with mismatched object relatives; however, as discussed earlier, this behaviour may rely on superficial strategies rather than full syntactic comprehension. Conversely, I09 exhibited the opposite pattern: higher accuracy with object relatives without mismatch (92%) than with mismatch (67%). Given I09's poor performance on BADA's grammaticality judgment task, this pattern may reflect a morphological deficit that impairs processing of number features, rendering number mismatch disruptive rather than helpful. The remaining four participants showed no substantial effects of number mismatch. For two of them, below-chance accuracy across all structures suggests that aphasia severity may have precluded any facilitation. For the other two, no clear reason emerges, and we can speculate that the morphological deficit has nullified any potential benefit of number mismatch.

Interestingly, despite variability, the results rule out superficial morphological cues as the source of number mismatch effects, since no participant showed general facilitation across subject and object relatives, challenging non-structural similarity accounts (see Terzi & Nanousi, 2018 for Greek).

[22] Note that, based on examples (4-8), this strategy yields the correct response only in the case of object relatives with number mismatch, as in (8).

4.3 Production of relatives

Five out of ten participants were not able to undertake the task, two presented the expected pattern of subject advantage, two reported the inverted pattern, and one showed no difference. Large discrepancies between comprehension and production abilities are not entirely surprising. In particular, for participants for whom the production task was not feasible, lexical access and/or articulatory difficulties, commonly associated with agrammatism, made it impossible to produce sentences of the type included in the protocol. This often led to task abandonment or refusal to attempt the production trials.

The marked interindividual variability and the overall lower performance in production in comparison with comprehension may support an explanation in terms of modality-specific syntactic processing with impaired representation of the higher projection of the syntactic tree in production in agrammatism (Tree Pruning Hypothesis; Friedmann, 2001). However, I02 and I04 showed the expected subject advantage in production, alongside ceiling-level comprehension performance. This pattern suggests that the two modalities may rely on a shared underlying syntactic representation, with comprehension supporting production.[23] Furthermore, I02's high production accuracy in subject relatives challenges the idea of truncation of the syntactic tree.

Moving to the inverse pattern of performance, this may be explained by the order of testing, as subject relatives were always presented first. I03 and I07 may have initially struggled to engage with the sentence priming task, independently of structural complexity. Indeed, priming might operate on superficial similarities rather than on full syntactic processing (Branigan, 2007). Furthermore, lexical, phonological or articulation impairments may have contributed to the overall poor production, masking potential syntactic effects.

In accordance with both the inconsistent response patterns in production and the methodological limitations noted above, no effect of number mismatch emerged. This fact may also reflect selective morphological deficits in a subgroup of individuals with agrammatism (Beber et al., 2024), which could block any mismatch-related facilitation.

[23] This pattern is consistent with findings from a rehabilitation study on agrammatism in German, which also reported that preserved comprehension supported production outcomes (Schröder et al., 2015).

4.4 The performance with passive relatives

In typical language development, passive relatives were found easier than object relatives by Belletti and Contemori (2010), who propose that children may avoid minimality effects in relatives via smuggling (Collins, 2005). However, for our participants, passive relatives proved particularly challenging, especially in production, suggesting that language acquisition and agrammatism are different in this respect and that agrammatism is characterized by a specific deficit with passives (see also Gilardone et al., 2023), possibly due to a disrupted representation of the smuggling operation.

4.5 Limitations and future developments

This study is limited by the small sample size; therefore, our findings need to be confirmed through group analyses involving a larger number of participants and experimental items. It is also limited by the relatively limited characterization of participants from neurological (absence of neuroimaging data), neuropsychological, and linguistic perspectives: all of which could help account for the observed variability. In particular, a standardized assessment of morphological processing, which is currently unavailable in Italian, could further clarify the patterns of performance. Indeed, the facilitation induced by grammatical number mismatch presupposes correct encoding and decoding of morphophonological information both upstream (in comprehension) and downstream (in production) of morphosyntactic processing. Furthermore, future research could also extend cross-linguistically the analysis of these phenomena and investigate competing hypotheses on passives through targeted manipulations of morphosyntactic features.

5. CONCLUSION

This study explored the subject-object asymmetry in the comprehension of relatives among Italian speakers with agrammatism, confirming an overall subject-relative advantage despite marked interindividual variability. Number mismatch occasionally facilitated object-relative comprehension; the effect was inconsistent, yet incompatible with facilitation based on superficial morphological cueing. In production, no reliable pattern emerged, likely due to extra-syntactic deficits and task-related limitations. The findings also point to specific difficulties with passivisation in agrammatism. Overall, linguistic theory proved essential in interpreting these complex data, with Generalized Minimality offering the better, yet not unchallenged, account.

REFERENCES

Adani, F., Van der Lely, H. K., Forgiarini, M., & Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier. *Lingua*, *120*(9), 2148-2166. https://doi.org/10.1016/j.lingua.2010.03.018

Adelt, A., Burchert, F., Adani, F., & Nicole, S. (2020). What matters in processing German object relative clauses in aphasia-timing or morpho-syntactic cues?. *Aphasiology*, *34*(8), 970-998. https://doi.org/10.1080/02687038.2019.1645290

Adelt, A., Stadie, N., Lassotta, R., Adani, F., & Burchert, F. (2017). Feature dissimilarities in the processing of German relative clauses in aphasia. *Journal of Neurolinguistics*, *44*, 17-37. https://doi.org/10.1016/j.jneuroling.2017.01.002

Beber, S., Bontempi, G., Miceli, G., & Tettamanti, M. (2024). The Neurofunctional Correlates of Morphosyntactic and Thematic Impairments in Aphasia. *Neuropsychology Review*, 1-34. https://doi.org/10.1007/s11065-024-09648-0

Belletti, A., & Contemori, C. (2010). Intervention and attraction: On the production of Subject and Object Relatives by Italian (young) children and adults. In *Language Acquisition and Development, Proceedings of Gala 2009*. Cambridge Scholars Press.

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, *122*(10), 1053-1069. https://doi.org/10.1016/j.lingua.2012.02.007

Branigan, H. (2007). Syntactic priming. *Language and Linguistics Compass*, *1*(1-2), 1-16. https://doi.org/10.1111/j.1749-818X.2006.00001.x

Caplan, D., Waters, G., DeDe, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia. *Brain and Language*, 101(2), 103-150. https://doi.org/10.1016/j.bandl.2006.06.225

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension. *Brain and Language*, 3(4), 572-582. https://doi.org/10.1016/0093-934X(76)90048-1

Collins, C. (2005). A smuggling approach to the passive in English. *Syntax*, *8*(2), 81-120. https://doi.org/10.1111/j.1467-9612.2005.00076.x

Friedmann, N. (2001). Agrammatism and the psychological reality of the syntactic tree. *Journal of Psycholinguistic Research*, 30, 71-90. https://doi.org/10.1023/A:1005256224207

Friedmann, N. (2008). Traceless relatives: Agrammatic comprehension of relative clauses with resumptive pronouns. *Journal of Neurolinguistics*, *21*(2), 138-149. https://doi.org/10.1016/j.jneuroling.2006.10.005

Friedmann, N., Reznick, J., Dolinski-Nuger, D., & Soboleva, K. (2010). Comprehension and production of movement-derived sentences by Russian speakers with agrammatic aphasia. *Journal of Neurolinguistics*, *23*(1), 44-65. https://doi.org/10.1016/j.jneuroling.2009.08.002

Gilardone, G., Viganò, M., Costantini, G., Monti, A., Corbo, M., Cecchetto, C., & Papagno, C. (2023). The role of verbal short-term memory in complex sentence comprehension. *International Journal of Language & Communication Disorders*. https://doi.org/10.1111/1460-6984.12851

Grillo, N. (2008). *Generalized Minimality*. Utrecht Institute of Linguistics, OTS.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences*, *23*(1), 1-21. https://doi.org/10.1017/s0140525x00002399

Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua*, 85(2-3), 211-258. https://doi.org/10.1016/0024-3841(91)90022-W

Lukatela, K., Shankweiler, D., & Crain, S. (1995). Syntactic processing in agrammatic aphasia by speakers of a Slavic language. *Brain and Language*, *49*(1), 50-76. https://doi.org/10.1006/brln.1995.1021

Luzzatti, C., De Bleser, R., Scola, I., Frustaci, M., & Willmes, K. (2023). Update on the psychometric properties for the Italian version of the Aachen Aphasia Test (IT-AAT). *Aphasiology*, 37(4), 658-695. https://doi.org/10.1080/02687038.2022.2037501

Miceli, G., Laudanna, A., Burani, C. & Capasso, R. (1994). *BADA. Batteria per l'Analisi dei Deficit Afasici*. Università Cattolica del Sacro Cuore.

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data. *Neurological Sciences*, 34(5), 749-754. https://doi.org/10.1007/s10072-012-1130-x

Nerantzini, M., Varlokosta, S., Papadopoulou, D., & Bastiaanse, R. (2014). Wh-questions and relative clauses in Greek agrammatism. *Aphasiology*, *28*(4), 490-514. https://doi.org/10.1080/02687038.2013.870966

Rizzi, L. (2018). Intervention effects in grammar and language acquisition. *Probus*, *30*(2), 339-367. https://doi.org/10.1515/probus-2018-0006

Schröder, A., Burchert, F., & Stadie, N. (2015). Training-induced improvement of noncanonical sentence production does not generalize to comprehension. *Cognitive Neuropsychology*, 32(3-4), 195-220. https://doi.org/10.1080/02643294.2014.968535

Terzi, A., & Nanousi, V. (2018). Intervention effects in the relative clauses of agrammatics. *Glossa*, *3*(1). https://doi.org/10.5334/gjgl.274

Varlokosta, S., Nerantzini, M., Papadopoulou, D., Bastiaanse, R., & Beretta, A. (2014). Minimality effects in agrammatic comprehension. *Lingua*, 148, 80-94. https://doi.org/10.1016/j.lingua.2014.05.013