Agrammatic output in non-fluent, including Broca’s, aphasia as a rational behavior

Evelina Fedorenko, Rachel Ryskin & Edward Gibson

To cite this article: Evelina Fedorenko, Rachel Ryskin & Edward Gibson (2022): Agrammatic output in non-fluent, including Broca’s, aphasia as a rational behavior, Aphasiology, DOI: 10.1080/02687038.2022.2143233

To link to this article: https://doi.org/10.1080/02687038.2022.2143233

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 18 Nov 2022.

Submit your article to this journal

View related articles

View Crossmark data
Agrammatic output in non-fluent, including Broca’s, aphasia as a rational behavior

Evelina Fedorenko a,b,c, Rachel Ryskin d and Edward Gibson a

a Massachusetts Institute of Technology, Brain & Cognitive Sciences Department; b Massachusetts Institute of Technology, McGovern Institute for Brain Research; c Speech and Hearing in Bioscience and Technology program at Harvard University; d University of California, Merced, Cognitive & Information Sciences Department

ABSTRACT

Background: Speech of individuals with non-fluent, including Broca’s, aphasia is often characterized as “agrammatic” because their output mostly consists of nouns and, to a lesser extent, verbs and lacks function words, like articles and prepositions, and correct morphological endings. Among the earliest accounts of agrammatic output in the early 1900s was the “economy of effort” idea whereby agrammatic output is construed as a way of coping with increases in the cost of language production. This idea resurfaced in the 1980s, but in general, the field of language research has largely focused on accounts of agrammatism that postulated core deficits in syntactic knowledge.

Aims: We here revisit the economy of effort hypothesis in light of increasing emphasis in cognitive science on rational and efficient behavior.

Main contribution: The critical idea is as follows: there is a cost per unit of linguistic output, and this cost is greater for patients with non-fluent aphasia. For a rational agent, this increase leads to shorter messages. Critically, the informative parts of the message should be preserved and the redundant ones (like the function words and inflectional markers) should be omitted. Although economy of effort is unlikely to provide a unifying account of agrammatic output in all patients—the relevant population is too heterogeneous and the empirical landscape too complex for any single-factor explanation—we argue that the idea of agrammatic output as a rational behavior was dismissed prematurely and appears to provide a plausible explanation for a large subset of the reported cases of expressive aphasia.

Conclusions: The rational account of expressive agrammatism should be evaluated more carefully and systematically. On the basic research side, pursuing this hypothesis may reveal how the human mind and brain optimize communicative efficiency in the presence of production difficulties. And on the applied side, this construal of expressive agrammatism emphasizes the strengths of some patients to flexibly adapt utterances in order to communicate in spite of grammatical difficulties; and focusing on these strengths may be more effective than trying to “fix” their grammar.

CONTACT Evelina Fedorenko evelina9@mit.edu

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
1. Setting the stage: expressive agrammatism and evidence against a unifying account in terms of syntactic knowledge loss

Non-fluent aphasia, including Broca’s aphasia, is characterized by difficulties in language production (e.g., Broca, 1861; Goodglass, 1976, 1993; Luria, 1964; see de Bleser, 1987 for a historical overview; see Gorno-Tempini et al., 2004 and Mesulam, 2001 for discussions of a similar profile in the non-fluent variant of primary progressive aphasia). This production impairment extends from relatively mild—where individuals can produce some phrases and sentences—to severe, where an individual may be able to utter at most 1-2 word utterances (what Luria called “serial naming”; Luria, 1970) or perhaps only a single syllable (as in Broca’s first reported case of his patient ‘Tan’). In milder cases, the output is typically morpho-syntactically impoverished, with short, structurally simple utterances and frequent omission or incorrect use of function words (e.g., articles, prepositions) and functional morphological markers (e.g., case and number markers on nouns, or person and tense markers on verbs) (e.g., Deleuze, 1819, as cited in Goodglass & Menn, 1985; Goodglass et al., 1972; Goodglass, 1997; Isserlin, 1922; Kean, 2013; Luria, 1970; Monakow, 1914; Pick, 1898, 1931/1973; Pitres, 1898; Tissot et al., 1973). This kind of output is consequently often described as “agrammatic”, i.e., without/ lacking grammar (Steinthal, 1871, as cited in Kolk et al., 1985; also, Kussmaul, 1876, as cited in Tsvetkova & Glozman, 1975).

In contrast to their production difficulties, the comprehension abilities of individuals with non-fluent aphasia appear relatively spared (e.g., Goodglass, 1976; Luria, 1964). But already at the beginning of the 20th century, it was observed that patients with expressive agrammatism have difficulty understanding certain kinds of sentences (see de Bleser, 1987 and Kolk et al., 1985 for reviews of such observations dating back to Bonhoeffer, 1902, Kleist, 1916, and Salomon, 1914). This observation gained prominence when in the 1970s, Zurif and colleagues (Caramazza & Zurif, 1976; Zurif et al., 1972) provided experimental evidence that patients with expressive agrammatism have difficulty understanding sentences whose meaning cannot be inferred from word meanings and world knowledge, and instead requires reliance on syntactic cues, like word order, function words, and/or inflectional morphology (for additional evidence of comprehension deficits in this population, see Ansell & Flowers, 1982; Bates et al., 1987; Berndt & Caramazza, 1980; Caplan et al., 2006; Heilman & Scholes, 1976; Schwartz et al., 1980 inter alia).

In light of the evidence that non-fluent aphasia is apparently characterized by both a) a reduction of syntactic information in production and b) a deficit in interpreting sentences where the meaning hinges on syntax, some researchers argued for a core deficit in syntactic representations in this population (e.g., Berndt & Caramazza, 1980; Caramazza, Berndt, et al. 1981; Cornell et al., 1993; Friedmann, 2006; Friedmann & Grodzinsky, 1997; Grodzinsky, 1986, 1995a, 1995b, 2000; Hickok & Avrutin, 1995; Kean, 1985; Mauner et al., 1993; Zurif et al., 1993 inter alia; for earlier discussions, see Jakobson, 1956, 1964). However, interpretations of expressive agrammatism in terms of syntactic knowledge loss have long been questioned (e.g., Bates & Goodman, 1997; Caplan et al., 1996; Carpenter et al., 1994; Dick et al., 2001; Friederici & Frazier, 1992; Kolk & van Grunsven, 1985; Miyake et al., 1994). We here highlight two key arguments against the core-syntactic deficit accounts.
**First**, patients with expressive agrammatism exhibit high variability in their syntactic comprehension abilities (e.g., Badecker & Caramazza, 1985; Berndt, 1987; Caplan et al., 2013; Goodglass & Menn, 1985; Howard, 1985; Parisi, 1987). Indeed, some exhibit no difficulties whatsoever in understanding even complex structures (e.g., Caramazza & Hilis, 1989; Isserlin, 1922; Kolk et al., 1982, 1985; Miceli et al., 1983; Nespoulous et al., 1988; see Berndt, 1991, for a review of dissociations between agrammatic production and agrammatic comprehension). The fact that some patients with expressive agrammatism have intact syntactic comprehension rules out a core syntactic deficit as the explanation of their agrammatic output. The existence of such patients shows that expressive agrammatism can be observed in the presence of intact syntactic knowledge.

And **second**, even in patients with agrammatism who exhibit sentence comprehension impairments, such impairments often do not appear to reflect damage to syntactic representations (see e.g., Dick et al., 2001 for discussion). In particular, many such patients show preservation of much syntactic knowledge, including sensitivity to even subtle distinctions, as evidenced by their intact performance on acceptability judgment tasks (e.g., Devescovi et al., 1997; Isserlin, 1922; Linebarger et al., 1983; Shankweiler et al., 1989; Wulfeck, 1988; Wulfeck & Bates, 1991). Furthermore, performance of patients with agrammatism on syntactic comprehension tasks is affected by a) the paradigm (e.g., better performance on act-out tasks compared to sentence-picture matching tasks; Caplan & Futter, 1986; Caplan & Hildebrandt, 1988; Hildebrandt, 1986) and b) task demands, like the number of picture choices in the sentence-picture matching paradigm (Cupples & Inglis, 1993) or utterance length (Mitchum et al., 1995). The fact that some patients exhibit good syntactic comprehension performance with a change in paradigm or task setup suggests that syntactic knowledge is preserved and the observed ‘deficits’ result from non-linguistic task demands. In line with these findings, comprehension difficulties like those observed in some patients with expressive agrammatism have been shown to be inducible in neurologically intact adults—who obviously do not suffer from damaged syntactic representations—under cognitive load (e.g., Bates & Wulfeck, 1989; Bayer et al., 1987; Blackwell & Bates, 1995; Dick et al., 2000; Miyake et al., 1994), and are also observed in other patients with aphasia, without expressive agrammatism (e.g., Caplan et al., 1996; Goodglass et al., 1970, 1979; Kurowsky, 1981; Wilson & Saygin, 2004 inter alia). Finally, additional evidence for preserved syntactic knowledge in some patients with expressive agrammatism comes from cross-linguistic differences in how agrammatic output manifests. In particular, many patients’ output is characterized by features that are specific to their native language (e.g., Bates et al., 1991; Menn et al., 1990), which suggests that their production still obeys language-specific grammatical constraints (see Bates et al., 1991; Bates & Wulfeck, 1989; and Menn et al., 1990 for reviews).

In summary, accounts whereby agrammatic production is due to damage to syntactic knowledge representations do not seem tenable for many patients with expressive agrammatism. More generally, one important theme in the aphasia literature since the 1980s, when research on agrammatism flourished, has been the extreme heterogeneity of patients with expressive agrammatism (e.g., Badecker & Caramazza, 1985; Berndt, 1997; Bruns et al., 2019; Fromm et al., 2021; Goodglass & Menn, 1985; Hofstede, 1992; Kolk & van Grunsven, 1985; Rochon et al., 2000 inter alia). So, given that research to date has not converged on a single unifying explanation of agrammatic output—such convergence
seems generally unlikely—we here revisit one of the early ideas about expressive agrammatism that has largely fallen out favor but that we believe has promise.

The hypothesis dates back to Pick (1913/1973; see also Isserlin, 1922 and Salomon, 1914) and was later revived and developed in the 1980s and 1990s (e.g., Bates & Goodman, 1997; Heeschen, 1985; Kolk et al., 1985; Kolk, 1987; Kolk & Heeschen, 1990). The idea is that at least in some patients (cf. Section 5), agrammatic output does not result from the inability to produce grammatically well-formed utterances but is instead an adaptation to the increase in the cost of generating linguistic output. In the presence of difficulties with language generation, producing agrammatic output is a rational strategy for maximizing communicative efficacy (i.e., getting the intended message across). Given the growing emphasis in cognitive science, including language research, on rational and efficient behavior (e.g., Chater & Manning, 2006; Gershman et al., 2015; Gibson et al., 2019; Griffiths et al., 2010; Lieder & Griffiths, 2020; Rescorla, 2021), we believe this account deserves some new attention.

In the remainder of the paper, we summarize the basic tenets of Shannon’s (1948) communication theory as applied to human language processing (Section 2). This theory provides a mathematical description of efficient information transmission. We then introduce the idea of agrammatic output as a rational response to increased production costs (Section 3) and connect this idea back to earlier, related proposals (Section 4). Finally, we discuss some outstanding issues (Section 5).

2. Communication theory applied to natural language processing

Under communication theory (Shannon, 1948), any information exchange consists of i) a transmitter, ii) a noisy channel over which messages are passed, and iii) a receiver. In linguistic communication, the producer (e.g., speaker or writer) has some intended meaning they want to pass on to a comprehender (e.g., listener or reader). They produce a message to express this meaning, and the comprehender then attempts to recover the intended meaning from the perceived message. Communicative success is achieved when the recovered meaning is identical to the producer’s intended meaning.

The information content, or entropy, of a signal in this framework is construed as the freedom of choice in selecting a message (e.g., a word or an utterance in linguistic communication) from among the set of possible messages. In particular, elements that are highly predictable—based on the statistical rules that govern the use of the relevant symbols—carry little information. For example, the word “old” in the phrase “I am 4 years old” or “to” in “I am going to . . .” carry little information, whereas the word “Lana” in “my name is Lana” carries a lot of information. This formal framework provides a powerful way to precisely quantify information and characterize various properties of its transmission in any communication system, including human language.

Over the last decade, these classic ideas have permeated models of language comprehension. Early comprehension models had all assumed the linguistic input to the sentence processor to be pristine (e.g., Gibson, 1998; Hale, 2001; Jurafsky, 1996; Levy, 2008a). In contrast, Levy (2008b; see also Levy et al., 2009) proposed that the processes of language production and comprehension could be noisy (see Ferreira et al., 2002 and Goldberg & Ferreira, 2022 for a related proposal). In particular, Levy proposed that in order to interpret an utterance, we combine our knowledge of what is likely to be
communicated (the prior probability of a message) with our knowledge of how messages can get corrupted by noise during transmission (the noise model). Gibson et al. (2013) provided support for this hypothesis across diverse English constructions. They showed that an implausible message is likely to be interpreted as its plausible alternative when comprehenders can easily explain how the message got corrupted from a plausible to an implausible version (e.g., ‘The mother gave the candle the daughter’ likely resulted from ‘The mother gave the candle to the daughter’ via the deletion of ‘to’). In contrast, in the absence of a close plausible alternative, comprehenders rely on the literal string for interpretation, according to the compositional rules of the language. Further, the reliance on the prior (i.e., rate of plausibility-based inferences) vs. literal interpretation is strongly modulated by the statistical properties of the input context: when the surrounding input is noisy (i.e., contains many errors), readers make more plausibility-based inferences, and when the input includes a high proportion of implausible sentences, readers are more likely to endorse the literal interpretation.

This “noisy-channel” proposal has found extensive empirical support across diverse materials and languages (e.g., Gibson et al., 2017; Keshet & Meltzer-Asscher, 2021; Liu et al., 2020; Nathaniel et al., 2018 March; Poliak et al., 2022; Poppels & Levy, 2016; Ryskin et al., 2018, 2021; Zhan et al., 2017; see Gibson et al., 2019 and Traxler, 2014 for reviews). This framework has also been applied to aphasic comprehension (e.g., Gibson et al., 2015; Ryskin et al., 2019; Warren et al., 2017): given that individuals with aphasia plausibly assume a higher amount of noise in the input that their language system receives, they are predicted to rely more strongly on the prior probability of a message, i.e., on plausibility cues, which is indeed the case (e.g., Caramazza & Zurif, 1976; Gibson et al., 2015; Saffran et al., 1998). Next, we show how this information-theoretic framework can be applied to expressive agrammatism.

3. Agrammatic output as a rational response to increased production costs

As noted above, information in Shannon’s theory of communication is defined as the freedom in selecting a particular element from a set of all elements. The parts of the message that are not under the producer’s control and are instead determined by the statistical rules of the relevant code are said to be redundant. For example, English is at least 50% redundant (Shannon, 1951; see Guerrero, 2009, for a higher estimate), and the estimates are similar for other natural languages (e.g., Newman & Waugh, 1960). This means that half or more of the sounds or morphemes or words in any given message can often be omitted and the message would still be complete. Linguistic codes have plausibly evolved to be highly redundant in order to make messages more robust to noise (e.g., Campbell, 1982; Gibson et al., 2013; Mahowald, Diachek et al., 2022; Nubold & Turner, 1987; Piantadosi et al., 2011) and/or easier to learn (e.g., Tal & Arnon, 2022).

Let us now consider the cost of message encoding and/or transmission during language production. The producer has to expend a certain amount of energy to plan and utter messages. Let’s say there is a fixed cost per unit of linguistic output under typical conditions. For a rational agent, an increase in this cost—perhaps above some threshold—should lead to a reduction in message length. Critically, given that different parts of the message differ in how informative they are, a rational agent would preserve the parts that carry the most information (these are also likely the elements that are activated earlier
because utterance construction is driven by meaning; e.g., Ferreira, 2008) and omit the most redundant elements. The latter happen to be function words and inflectional morphological markers (e.g., Shannon, 1948; see Mahowald, Diachek et al., 2022 for empirical evidence). This strategy would thus result in precisely the kind of output that we observe in some individuals with expressive agramatism: short utterances with frequently omitted function words and morphological markers. Thus, agrammatic output can be thought of as a rational and communicatively optimal behavior under the conditions of increased production cost (see e.g., Gallée et al., 2021 for evidence that output of individuals with non-fluent primary progressive aphasia has similar informativeness as that of neurologically intact controls, as assessed in a picture description task; see also Beeke, 2013 and van Lancker, 2001 for additional discussion).

Two points are important in this construal of agrammatic output. First, although in discussions of rational human behaviors, terms like ‘strategy’ and ‘choice’ are commonly invoked, it is important to emphasize that these processes need not be subject to conscious awareness or require explicit reasoning and decisions about how to maximize behavioral utility, i.e., to achieve desired outcomes. Ample evidence shows that the human mind implements unconscious Bayesian inferences across diverse domains of perception (e.g., Knill & Richards, 1996; Körding et al., 2007; Weiss et al., 2002), motor control (e.g., Haith & Krakauer, 2013; Wolpert, 2007), and cognition (e.g., Baker & Tenenbaum, 2014; Chater et al., 2010; Gopnik & Bonawitz, 2015; Griffiths & Tenenbaum, 2009; Levy, 2008b) (see e.g., Chater & Manning, 2006; Gershman et al., 2015; Gibson et al., 2019; Griffiths et al., 2010; Lieder & Griffiths, 2020; Rescorla, 2021 for reviews)².

And second, the cost of language production can increase for any number of reasons and thus affect different stages of the production process, from conceptual processing (thinking), to high-level linguistic planning and encoding (converting abstract semantic representations into word sequences), to retrieving phonological word-forms, to lower-level motor planning and encoding, to executing the physical movements of the effectors (i.e., articulators when speaking, or fingers/hands when writing or signing). The general idea of agrammatic output as an adaptation to increased production costs is compatible with different sources of the cost (although the predictions would differ somewhat depending on the stage at which difficulties arise: for example, late-stage difficulties should selectively affect production in the relevant modality). Further, this idea can be applied to any population, including neurologically intact individuals for whom the cost of production is suddenly higher than normal for whatever reason (from cognitive demands—such as individuals under extreme physical or emotional duress—to having a sore throat or talking while having one’s mouth full, or having to pay for each word when sending a telegram or each letter when sending a text)³.

4. The origins of the “economy of effort” hypothesis, some evidence for it, and some arguments against it

As noted above, related hypotheses about agrammatic output have been put forward before. The earliest form of this idea, termed the “economy of effort” hypothesis, appears to be due to Pick (1931/1973) and Salomon (1914; see also Isserlin, 1922). In the 1980s and 1990s, ideas related to the original economy of effort hypothesis were advocated by Heeschen and Kolk. In both Heeschen’s “avoidance-correctness hypothesis” (1985; see
also Goldstein, 1948) and Kolk’s “adaptation hypothesis” (Kolk, 1987, 2006; Kolk et al., 1985; Kolk & Heeschen, 1990), agrammatic output is construed as a reaction to the deficit and not a deficit in and of itself. To quote Heeschen (1985, p. 234), non-fluent speakers with aphasia omit certain grammatical markers “not because they are unable to emit them [. . .], but because they do not want to produce them because of the risk that they might go wrong or because producing them would cost too much effort”. And Kolk et al. (1985), who attribute the core of their proposal to Isserlin (1922), write (p. 196), “aphasics adapt by simplifying messages. They do so under pragmatic constraint. Therefore, relatively uninformative elements should have a high chance of being omitted.”

Heeschen, Kolk and their colleagues provide both informal observations and some experimental evidence in support of their proposals. For example, Heeschen describes a patient who produced classic agrammatic output in spontaneous conversations. However, in more clinical settings (e.g., when his speech was being recorded on a tape recorder), he abandoned his telegraphic style and attempted to speak (with great effort) in well-formed utterances. This case suggests that the ability to produce well-formed utterances can be preserved at least in some patients with expressive agrammatism, but because of the effort associated with producing such utterances, they resort to agrammatic speech in spontaneous interactions, presumably because it allows for greater fluency (see Isserlin, 1922, for a similar case; but also note that this behavior is generally rare—Varley, personal communication).

Heeschen and Kolk also discuss some experimental evidence, where under certain conditions, patients with expressive agrammatism in spontaneous speech show the ability to produce words (including function words) and constructions (e.g., passive voice) that they typically avoid (e.g., Hartsuiker & Kolk, 1998; Saffran & Martin, 1990). Such data again suggest that expressive agrammatism does not index a fundamental inability to generate certain words/structures but rather their avoidance.

Heeschen and Kolk differ in their views of the underlying impairment in non-fluent aphasia: Kolk argues that the deficit is due to the slowing of language production (due to delayed activation and/or faster decay of representations in memory; Kolk, 1987, 2006; see also Bates & Goodman, 1997), but Heeschen (1985) maintains that syntactic representations themselves are affected (“although admittedly in a still unclear way”, p. 247). Both emphasize, though, that agrammatic output is well-formed: it uses a different (simplified) register than typical speech, but it follows its own set of conventions. For example, Heeschen (1985) writes, “agrammatics are so effective in avoiding danger points that all that is left in their speech [. . .] is basically correct”, and Kolk (1987) writes (p. 380), “aphasic telegraphic speech is produced in the same way as baby talk, foreigner talk and context-elliptical utterances” (see also Kolk, 2001; Kolk & Heeschen, 1992). In other words, the agrammatic ‘register’ is always available to language producers and can be chosen by neurologically intact individuals in cases of temporary increases in production costs, as discussed above, or for the benefit of the comprehender (e.g., when speaking to individuals without full command of the language, like young children or non-native speakers; e.g., Ferguson & DeBose, 1977; see also work on the syntax of newspaper headlines, or ‘headlines’, which uses a similar register; e.g., Halliday, 1969; Mårdh, 1980; van Dijk 1988).

Another proposal that it is worth mentioning is Kean’s (1977, 1979) proposal that non-fluent speakers with aphasia omit function words and inflectional morphological markers because these elements are typically phonologically unstressed. Kean’s idea of deficient
phonological representations as a unifying account of expressive agrammatism is unlikely for similar reasons that a core syntactic deficit is unlikely, e.g., because of robust dissociations between production and comprehension in some patients. However, phonological reduction is correlated with informativity, such that highly predictable elements are often reduced (e.g., Bell et al., 2009; Gahl, 2008; Gahl et al., 2012; Lieberman, 1963). As a result, Kean’s account would, in many cases make similar predictions about the nature of the agrammatic output as the economy of effort account that we advocate for here.

One common argument against the economy of effort idea is that some patients with expressive agrammatism repeat their utterances (sometimes in an apparent attempt to correct the grammar) (e.g., Goodglass et al., 1972; Luria, 1970). Another argument that is sometimes made is that the output of some non-fluent speakers with aphasia contains frequent “empty” words and expressions (like “y’know”; e.g., Goodglass & Menn, 1985). However, these arguments only apply to formulations of the economy of effort hypothesis where the cost arises at the relatively late (articulation) stage. If, instead, the cost arises at an earlier stage (such as linguistic planning or encoding), then these behaviors are compatible with economy of effort. In such cases, patients may avoid redundancy during message/utterance assembly, but the cost of phonological encoding may be low, similar to that of a neurologically intact individual. As a result, these patients may both i) repeat their utterance in a slightly different way, to the best of their ability, if they feel they are not getting their message across (much like neurologically intact individuals sometimes do in an effort to clarify what they mean), and ii) produce filler phrases like “y’know”—a common strategy in language production used to minimize pauses and increase fluency (e.g., Beeke, 2003). The cost of producing such filler phrases is likely minimal in terms of linguistic planning and access.

Another apparent challenge for the economy of effort idea is the relative paucity of verbs in the output of patients with agrammatism (e.g., Berndt et al., 1997; Hillis et al., 2006; Miceli et al., 1984; Saffran et al., 1980; Zinsinger et al., 1990). Verbs carry a lot of information so should be preserved to maximize message informativity. Two points are important here. First, in many cases, a verb’s arguments (noun phrases) strongly constrain its identity (e.g., Levin, 1993). As a result, if the event participants have already been mentioned or are clear from the context, the verb may be redundant. And second, to the extent that speakers with agrammatism do produce verbs, these verbs tend to be semantically rich (e.g., Berndt et al., 1997; Breedin et al., 1998; Gordon & Dell, 2003; Kohn et al., 1989; Zimmerer et al., 2020), in line with trying to maximize informativity. However, additional factors—other than maximizing informativity—almost certainly come into play in the utterance construction of patients with expressive agrammatism. For example, the ease of accessing different linguistic elements (words or constructions), which strongly depends on their frequencies, has been shown to play an important role in language production in aphasia (e.g., Bruns et al., 2019; Duffield, 2016; Gahl & Menn, 2016; Menn, 2009; Menn & Bastiaanse, 2016; Rezaii et al., 2022).

5. Concluding remarks

We have discussed some of the history of research on expressive agrammatism with the goal of bringing to the forefront a hypothesis that has largely fallen out of favor. In particular, we have argued that the old “economy of effort” idea—whereby agrammatic
output is a reaction or adaptation to language production difficulties—deserves a re-examination. In particular, given that many aspects of human behavior approximate that of a (resource-rational) Bayesian ideal observer (e.g., Chater & Manning, 2006; Gershman et al., 2015; Gibson et al., 2019; Griffiths et al., 2010; Lieder & Griffiths, 2020; Rescorla, 2021), we argue that expressive agrammatism can be construed as another example of a rational behavior: shortening linguistic messages by omitting uninformative elements in an effort to maximize communicative efficiency in the presence of production difficulties.

We conclude by highlighting several important points.

First, a fully unified account of expressive agrammatism seems unlikely (see also Bates & Goodman, 197; cf. Kolk 1987, 2006, who argues that his adaptation hypothesis may provide such an account). Patients with expressive agrammatism are a highly heterogeneous population, varying in the nature, scope, and severity of their symptoms, in both production and comprehension (e.g., Badecker & Caramazza, 1985; Goodglass & Menn, 1985; Kolk & van Grunsven, 1985). However, a rational adaptation account of agrammatic production appears to provide a plausible explanation for a large subset of the reported cases of expressive aphasia.

Second, at least some patients with aphasia (including those with expressive agrammatism) likely do suffer from damage to syntactic representations. After all, linguistic (including syntactic) knowledge has to be stored somewhere in the brain. The most likely place is the fronto-temporal language-selective network (e.g., Fedorenko et al., 2011). It follows then that damage to this network should, in some cases, affect linguistic representations. Given that syntactic processing does not appear to be carried out focally within a particular part of the language network and is instead distributed across its inferior frontal, posterior temporal, and anterior temporal components (e.g., Bautista & Wilson, 2016; Blank et al., 2016; Fedorenko et al., 2010, 2020; Shain, Blank et al., 2020; Shain et al., 2022), syntactic representation loss may require extensive perisylvian damage affecting both frontal and temporal components of the language network, as in some cases of global aphasia (e.g., Varley et al., 2005; Bek et al., 2010; cf. Wilson et al., 2022 for evidence that posterior temporal damage may suffice for long-term syntactic comprehension difficulties).

Importantly, in order to argue that a given patient suffers from a core syntactic knowledge impairment, it is necessary to demonstrate that the deficit a) is present in both production and comprehension, b) generalizes across spoken and written modalities, diverse linguistic materials, and experimental paradigms, and c) cannot be explained by low-level perceptual/motor difficulties or non-linguistic factors (e.g., executive limitations). This is a high bar, which is not met in the vast majority of published studies that argue for core syntactic impairments in patients with expressive agrammatism.

Third, as is the case with much psycholinguistic and neurolinguistic work in general, the bulk of past work on agrammatism has been conducted on English and other Germanic languages. If we are to make claims about the language system in general, it is critical to examine a wider range of typologically diverse languages (e.g., see Heeschen, 1985, for a discussion of how a failure to examine morphologically rich languages has led to a misguided characterization of the output of patients with fluent/Wernicke’s aphasia as grammatically well-formed). Bates, MacWhinney and their colleagues (e.g., Bates & Wulfeck, 1989; Bates et al., 1991; MacWhinney et al., 1991) and Menn and colleagues (e.g., Menn et al., 1990) attempted to remedy the situation with their cross-linguistic
investigations of agrammatic aphasia, but the field remains focused on a handful of ‘dominant’ languages, most prominently, English.

**Finally,** the account of expressive agrammatism (due to stroke or degeneration) as a rational behavior may be helpful in both advancing basic research and informing/guiding aphasia therapies. On the basic research side, pursuing this explanation may reveal how the human mind and brain optimize communicative efficiency in the presence of (different kinds of) production difficulties, and thus illuminate core properties of the language system as well as of the domain-general mechanisms, which may help the speech and language mechanisms recover from damage.

This construal of expressive agrammatism may also have applications for aphasia treatments. Historically, a lot of emphasis in aphasia evaluation and rehabilitation has been placed on standardized meta-linguistic tasks, like grammaticality judgments, picture naming, or sentence-picture matching. Such tasks can be critical during the *evaluation* stage, for understanding the relative strengths and weaknesses of a given patient, although they should arguably be supplemented with paradigms that help evaluate functional communication abilities (e.g., see Manochiopinig et al., 1992 for a review of such assessments) and/or with detailed analyses of naturalistic conversation exchanges (the Conversation Analysis approach; e.g., Beeke, 2013; Killmer et al., 2022). However, the goal of aphasia *therapy* is clearly to improve the patients’ ability to interact with others in their daily lives, not to raise their performance on some linguistic task (e.g., Armstrong & Ferguson, 2010; Beeke, 2013; Doedens & Meteyard, 2020; Holland, 1980; Sarno, 1969; van Lancker, 2001; Volkmer et al., 2020). Here, we have emphasized the strengths of some patients with expressive agrammatism to adapt their utterances in order to communicate in spite of language production difficulties. These strengths can and should be leveraged during therapy—by working with both the patient and their communication partner—to help strengthen the communication strategies that already work well (focusing on words that have high informativity in everyday contexts; e.g., Edmonds, 2016), minimize ones that are not effective (e.g., trying to produce complex syntactic structures), and introduce new ones (e.g., tailored to the particular communicative needs of the individual), all to maximize the patient’s ability to have meaningful interactions with those around them (e.g., Berube & Hillis, 2019; Davis & Wilcox, 1985; Fridriksson, 2021; Green, 1984; Holland, 1991, 2021; Kagan, 1998; Volkmer et al., 2020; Holland, 2021).

**Notes**

1. It is worth noting that—although spoken production has historically been studied more than written production—at least in some cases, agrammatic output characterizes both spoken and written modalities (e.g., Heilbroner, 1906, as cited in de Bleser, 1987; Josephy-Hernandez et al., 2022; Kleist, 1916 and Pitres, 1898, as cited in Goodglass & Menn, 1985; Nespoulous et al., 1988).

2. Relatedly, Fedorenko & Shain (2021) have summarized evidence against the role of domain-general executive resources—resources that support reasoning and decision making (e.g., Duncan, 2010; Duncan et al., 2020)—in core linguistic processes like lexical access and semantic/syntactic composition, showing that such computations are instead carried out within the language-specific cortical areas (Fedorenko et al., 2011).

3. Can this framework be applied to patients with fluent aphasia? In contrast to difficulties in production in the presence of intact linguistic knowledge, as in many non-fluent patients that
we focus on here, fluent paragrammatic patients suffer from actual loss of linguistic knowledge (the mapping between linguistic forms and the associated meanings). As a result, such patients simply cannot be communicatively efficient. However, their thought processes are intact (e.g., Fedorenko & Varley, 2016) and so are their social skills and understanding of social situations (including, for example, an understanding that a response is required when someone asks you a question), and so they produce some linguistic output in an effort to follow the social conventions. But because of the loss of linguistic knowledge, they cannot understand others or evaluate their own productions as not communicating anything meaningful/relevant.

Acknowledgements

The authors are grateful to Brian MacWhinney, Lise Menn, Rosemary Varley, Stephen Wilson, Heather Wright, and two anonymous reviewers for constructive comments on the earlier version of the manuscript, members of the Fedorenko and Gibson labs for helpful discussions, Saima Malik-Moraleda for help with reviewing the literature, and Niharika Jhingan and Sara Swords for help with the references. EF was supported by NIH R01 awards DC016950 and DC016607 from NIDCD, U01 award NS121471 from NINDS, and research funds from the McGovern Institute for Brain Research, the Brain and Cognitive Sciences Department, and the Simons Center for the Social Brain.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the the National Institute on Deafness and Other Communication Disorders [DC016607,DC016950].

ORCID

Evelina Fedorenko  http://orcid.org/0000-0003-3823-514X

References


Duffield, C. J. (2016). Frequency, activation, and the production of verb inflection in agrammatic aphasia. *Aphasiology* 30(11 [https://doi.org/10.1080/02687038.2015.1105357](https://doi.org/10.1080/02687038.2015.1105357)).


Goodglass, H., & Menn, L. (1985). Is Agrammatism a Unitary Phenomenon?**The writing of this chapter and the authors’ cited research were supported in part by the Medical Research Service of the Veterans Administration, and in part by USPHS Grants NS06209 and NS07615. In M.-L. Kean (Ed.), Agrammatism (pp. 1–26). Academic Press. [https://doi.org/10.1002/wcs.1330](https://doi.org/10.1002/wcs.1330).


