# NORMAL AND APHASIC PROCESSING OF SENTENCE STRUCTURE AND INTONATION[1]

Aita Salasoo
Department of Psychology
State University of New York
Binghamton, NY 13901 USA

Rita Sloan Berndt
Department of Neurology
University of Maryland Medical School
Baltimore, MD 21201 USA

## ABSTRACT

The contributions of intonation contour and memory load to performance in an auditory grammaticality judgment (AGJ) task were investigated. College students, agrammatic and other aphasics, and control subjects judged the grammaticality of vocoded utterances with original and flat fundamental frequency (F0) contours. For normal and aphasic listeners, sensitivity and bias differences between seven syntactic structures outweighed smaller benefits from intonation information and few memory demands. Differences emerged according to the severity and selective linguistic deficits among the aphasics. The observed strategic components of syntactic processing have implications for previous interpretations of AGJ data in relation to normal and aphasic language behavior.

## INTRODUCTION

While little is known about how syntactic knowledge is used during speech processing, a task that involves explicit grammaticality judgments has been used increasingly to address issues about the language deficits of individuals classified as agrammatic aphasics.[1] The present study attempts to provide baseline data from normal language users and to probe the role of extra-syntactic factors in the task. These goals are prerequisites to the use of the auditory grammaticality judgment (AGJ) task in the study of aphasia. In a recent study[1] four agrammatic patients who failed to use syntactic devices successfully in their language production and in comprehension tasks showed great sensitivity to the violations of those syntactic structures, when asked to judge whether spoken sentences were grammatical. Poor performance was found only in conditions using tag questions and reflexive pronouns. It was concluded that agrammatic aphasics do not have a general syntactic deficit, but that they fail to use syntactic structure in more demanding tasks. The account of the poor performance structures given is in terms of poor semantic encoding of lexical features that cannot support dependent syntactic analyses.

The present study aims to strengthen these conclusions by ensuring that the pattern of results will not generalize to nonagrammatic populations and by addressing possible confounds in stimulus materials. Two factors that may contribute to AGJ performance are intonation and memory load. Normal listeners can use prosodic cues including the pitch or intonation perceived from the fundamental frequency (F0) contour of a sentence in many listening tasks. Agreement does not exist about the dominant source of information when intonation and syntax conflict. [2] [3] Since syntax guides the F0 contours in speech production,[4] listeners may be able to use this information to perceive spoken language. The role of intonation in grammaticality judgments may be studied by removing the information carried by F0 and asking listeners to judge the grammaticality of the resultant utterances. To the extent that performance is worse for these stimuli, intonation cues to grammaticality are implicated.

Syntactic knowledge is relational. In the temporal course of fluent speech, syntactic violations may increase memory demands as the duration (or intervening information load) between violating segments is increased. Adjacent violating elements have fewer memory demands than distant ones. Increased memory demands, in turn, may decrease the detectability of syntactic violations.

The present investigation examines the influence of these properties (available to normal and possibly aphasic listeners) in the AGJ task. Specifically, violations to seven syntactic structures are used to investigate the effects of intonation contour and memory load on performance. In Experiment 1, a group of normal listeners is studied, and in Experiment 2, five aphasics with left-hemisphere lesions and three control subjects are tested.

## EXPERIMENT 1

### Method

**Subjects.** The subjects were 48 students from SUNY Binghamton, who were native English speakers aged 18-23 with no known speech or hearing problems.

**Materials and design.** Seven syntactic violation types were selected to include six which presented few problems to patients in the Linebarger et al. study[1], and one type – verb copying in tag questions – on which patients performed poorly. In each violation type, stimulus pairs were generated whose members differed in grammaticality, and as little as possible on other properties. A variety of lexical and transformational rules are represented in the violation types; they are described fully elsewhere.[5]

Approximately equal numbers of the Linebarger[1] sentences and new sentences were used. The additional sentences employed medium- to high-frequency words to increase the vocabulary of the stimulus set. Sentence length and violated constituent size were controlled. Memory load was operationalized as **violation location**, defined as the point at which the utterance could no longer be completed as well-formed. Nongrammatical stimuli were classified as having **early** (first three words), **middle**, or **late** (last two words) violation locations. Only late violations may have a wide range of distance (in number of words) between the disagreeing sentential elements. In all 156 critical pairs, and 24 practice pairs were employed. The mean length of the grammatical and nongrammatical utterances is 7.88 and 7.86 words, respectively. Overall, 26 pairs have early violations, 67 middle violations, and 63 late violations. Table 1 summarizes the stimuli.

---

[1] This work was supported by NINCDS grants NS-21054 to the University of Maryland Medical School and K04-NS-00851 to R. Berndt. A. Salasoo is now at Bell Communications Research, Inc., MRE 2G-356, 435 South St, Morristown, NJ 07960, USA.

**Table 1: Example Grammatical and Nongrammatical Stimuli (*) and Violation Location Distribution of 7 Syntactic Violation Types.** (E, M, and L refer to early, middle, and late violations, and N indicates the total number of stimulus pairs per condition.)

| Syntactic Violation Type | Grammatical and Nongrammatical Examples | E | M | L | N |
|---|---|---|---|---|---|
| **1. Verb control complements** | I let Harry pay for the birthday cake.<br>* I let Harry to pay for the birthday cake. | 2 | 16 | 6 | 24 |
| **2. Missing verb arguments** | Before the guests come I'll put the toys in the closet.<br>* Before the guests come I'll put the toys. | 3 | 11 | 10 | 24 |
| **3. Subject-auxiliary inversion** | Hasn't Fred walked the dog yet?<br>* Hasn't Fred hasn't walked the dog yet? | 8 | 4 | 0 | 12 |
| **4. Verb copying in tag questions** | The pies aren't very large, are they?<br>* The pies aren't very large, do they? | 0 | 0 | 24 | 24 |
| **5. Missing noun phrase elements** | I doubt I could afford a month's vacation in Greece.<br>* I doubt could afford a month's vacation in Greece. | 8 | 7 | 9 | 24 |
| **6. Wh-movement: Fronted noun phrases** | Whose broken fence will Gary fix next week?<br>* Whose broken will Gary fix next week fence? | 5 | 11 | 8 | 24 |
| **7. Gapless relative clauses** | She washed the windows that needed cleaning.<br>* She washed the windows that the floors needed cleaning. | 0 | 18 | 6 | 24 |

A male speaker read a random order of the stimuli at normal speed and with normal intonation. To prevent the occurrence of hesitations and other abnormalities in the F0 contour for nongrammatical stimuli,[4] model sentences were employed. The model was a grammatical sentence matched in number of syllables, and in the actual words as much as possible, to the following nongrammatical stimulus. When possible, the grammatical sentence served as the model. Recordings were digitized at 10 kHz, low-pass filtered at 4.8 kHz, and stored on a computer. After signal processing -- an LPC vocoder extracted 14 coefficients from overlapping 30-ms windows of the speech signal – each natural digitized utterance yielded a pair of vocoded stimuli, one with the **natural** F0 contour ranging from 75 - 175 Hz, and the other with a **flat** 90-Hz F0 contour.

Eight experimental tapes were generated for both the natural-F0 and flat stimuli. Each tape contained 3 practice and 39 critical stimuli with equal numbers of grammatical and nongrammatical utterances, and approximately equal representation of each violation type. A warning signal of three 1000-Hz tones preceded each stimulus. It was followed by two presentations of the stimulus, separated by an intersentence interval of one second. The intertrial interval was six seconds.

**Procedure.** Eight random groups of 6 subjects each listened to 2 natural-F0 and 2 flat-F0 tapes. A subject heard either the grammatical or nongrammatical version of each stimulus pair in either its natural-F0 or flat condition. The order of tapes and intonation conditions was counterbalanced between groups. Subjects were instructed to listen to the two presentations of each sentence and to decide whether or not the sentence was grammatical, recording their judgment in a response booklet. Stimuli were presented at a comfortable listening level from a speaker situated 1-2 meters away from the subjects.

### Results and discussion

For each subject, the mean proportion of hits (correctly accepting a grammatical sentence) and false alarms (FAs – incorrectly accepting a nongrammatical utterance as grammatical) in each experimental condition were computed. Nonparametric signal detection techniques were applied to the measurement of sensitivity to grammaticality and performance bias with the **A′** (A-prime) and **B′** (B-prime) statistics, respectively.[6] If $y = p(Hit)$ and $x = p(FA)$ , then:

$$A' = 0.5 + \frac{(y-x)(1+y-x)}{4y(1-x)} \qquad 0.5 \leq A' \leq 1.0$$

$$B' = \frac{y(1-y) - x(1-x)}{y(1-y) + x(1-x)} \qquad -1.0 \leq B' \leq 1.0$$

In the present task, higher values of **A′** indicate greater sensitivity to grammaticality. **B′** values of 0 indicate optimal criterion that maximizes hits and minimizes false alarms, i.e., no bias. Negative non-extreme values indicate a lax criterion; positive non-extreme values indicate a strict criterion. Table 2 shows mean hit and false alarm rates, **A′** and **B′** values for normal listeners making grammaticality judgments as a function of violation type and intonation condition. The reported findings are based on analyses of variance (ANOVAs) and post hoc comparisons and are statistically significant beyond the level of **p** < .01 unless otherwise indicated.

**Sensitivity to grammaticality.** Overall, listeners were very sensitive to syntactic structure. An ANOVA of **A′**s with violation type and intonation as factors found main effects of both variables. Large differences were observed between the seven violation types, **F**(6,42)=17.95. Post-hoc comparisons revealed three clusterings: Sensitivity was greatest for violation types 3 and 6, intermediate for violations types 2, 4, and 7, and lowest for types 1 and 5. Several accounts of this pattern need exploration, including the possibility that greater F0 (and syntactic) discontinuity existed for type 3 and 6 violations.

**Table 2: Grammaticality Judgment Data for Normal Listeners as a Function of Violation Type and F0 Contour.**

| Violation | F0 Contour | p(Hit) | p(FA) | A′ | B′ |
|---|---|---|---|---|---|
| 1 | natural | .891 | .143 | .928 | -.116 |
|  | flat | .877 | .185 | .909 | -.166 |
| 2 | natural | .906 | .077 | .953 | .009 |
|  | flat | .878 | .074 | .946 | .220 |
| 3 | natural | .972 | .010 | .990 | .466 |
|  | flat | .895 | .040 | .961 | .420 |
| 4 | natural | .860 | .025 | .957 | .663 |
|  | flat | .831 | .047 | .941 | .516 |
| 5 | natural | .871 | .118 | .929 | .004 |
|  | flat | .866 | .130 | .924 | .001 |
| 6 | natural | .930 | .026 | .975 | .440 |
|  | flat | .919 | .027 | .972 | .478 |
| 7 | natural | .871 | .087 | .940 | .172 |
|  | flat | .869 | .088 | .939 | .173 |

Despite the very high levels of performance, stimuli with preserved intonation contours were judged with greater sensitivity than flat stimuli, **F**(1,47)=21.20, and importantly, intonation and violation type failed to interact statistically, F(6,42)=1.15, **p**>.35. Thus, while the presence of F0 information increased sensitivity to grammaticality, it did not differentially affect that sensitivity for the various syntactic structures tested.

Next, the effects of memory load were examined through the violation location variable. A main effect of location was observed, **F**(2,46)=6.97, but this variable again failed to interact with intonation, **F**<1.0, **p**>.70. Post-hoc tests revealed that sensitivity was greatest for early violations (mean **A**′=.957) and did not differ for midsentence and late violations (mean A′s of .939 and .931, respectively). The result suggests that memory load contributes to sensitivity in grammaticality judgments: The violating segments are either adjacent or separated by a single word in the early conditions and thus may be encoded and stored together in memory, decreasing the memorial demands of the judgment task. Another possible interpretation is in terms of primacy, namely that sentence-initial constituents are attended and encoded better than constituents in later positions, and hence early violations are detected better.

**Response bias.** Arguments about sensitivity differences hold with greatest strength if subjects maintain a constant bias in performance. Evidence about this was derived from ANOVAs of the **B**′ measure analogous to those performed for sensitivity. The seven violation types differed in performance bias, **F**(6,42)=6.04. Only for stimuli in the verb control complement condition (where least sensitive performance was observed) was the mean **B**′ negative, suggesting a lax criterion: Subjects tended to accept both grammatical and nongrammatical utterances in that condition as grammatical. In contrast, a very strict criterion was observed for the tag questions condition. No main effect of intonation contour on bias was observed, **F**(1,47)=1.40, **p**>.24. Notably the bias differences between violation types are more dramatic than those related to intonation contour information. Examined by violation location, the most optimal criterion was found for the midsentence position (mean *B*′=.010), **F**(2,46)=5.21. Early and late violations did not differ statistically: Both exhibited strict criterion placement. Again, intonation did not interact with the location effect for performance bias, **F**(2,46)=2.20, **p**>.12.

In sum, Experiment 1 has provided five major findings about how normal listeners make grammaticality judgments. First, young adults are very sensitive to the grammaticality of spoken sentences. Second, in most cases, unbiased or conservative response criteria are adopted. Third, grammaticality is not unitary knowledge -- performance differed in both sensitivity and bias for different types of syntactic structures. Fourth, memory load as measured by violation locations may play a small part in sensitivity in the AGJ task. Finally, intonation contour information, even when optimally modelled for nongrammatical utterances, has a facilitatory effect on sensitivity to syntactic structure. These results suggest that grammaticality judgments offer valid evidence about the syntactic processes that occur during normal speech processing. The substantial differences in performance observed among our stimuli are not primarily attributable to intonational or memorial factors, but reflect, instead, strategies used by listeners. By placing strict response criteria, listeners are able to judge the grammaticality of sentences accurately; the only exception to this criterion placement resulted in poor judgment performance.

## EXPERIMENT 2

In Experiment 2 aphasic patients and control listeners make grammaticality judgments for the speech utterances used in Experiment 1. The study aims to replicate the Linebarger et al.[1] results with listeners diagnosed as agrammatic aphasics and to compare performance across various subject populations. Agrammatics are expected to perform well in all conditions except the tag questions. College students in Experiment 1 had no problems with that syntactic structure; analogous data do not exist for other aphasic, i.e., nonagrammatic, populations nor for normal listeners matched to agrammatic patients in age and education.

### Method

**Subjects.** Five aphasics with left-hemisphere lesions following cerebrovascular accidents participated. One of them, VS, had been tested by Linebarger et al.[1], and all of them had prior experience with the natural utterances on which the stimulus set was based. Full subject descriptions, including detailed reports of comprehension tasks and of performance with those natural speech stimuli may be found in Berndt et al.[5] For present purposes, the subjects are informally grouped as agrammatic (FM, VS), nonagrammatic aphasic (JD and JS have mild and severe auditory comprehension deficits, HY is anomic), and 3 normal control listeners of similar age and background.

**Materials and procedure.** The stimuli and tapes from Experiment 1 were used. All subjects were tested with stimuli preserving the F0 contour before flat stimuli over the course of 4-8 weeks. Subjects were tested individually and heard 2 tapes per session. The experimenter recorded the subject's verbal grammaticality judgment. Patient and control data are reported individually, since agrammatic and nonagrammatic labels may not reflect homogeneous deficits.

**Table 3: Grammaticality Judgment Data for 2 Agrammatic Listeners**

| Violation Type | F0 Contour | FM |  |  |  | VS |  |  |  |
|---|---|---|---|---|---|---|---|---|---|
|  |  | p(Hit) | p(FA) | A′ | B′ | p(Hit) | p(FA) | A′ | B′ |
| 1. Verb control complements | natural | .70 | .30 | .79 | .00 | .88 | .68 | .71 | -.35 |
|  | flat | .70 | .35 | .76 | -.04 | .96 | .70 | .78 | -.69 |
| 2. Missing verb arguments | natural | .83 | .39 | .81 | -.26 | .88 | .23 | .90 | -.25 |
|  | flat | .75 | .35 | .79 | -.10 | .96 | .41 | .88 | -.73 |
| 3. Subject-auxiliary inversion | natural | .67 | .18 | .83 | -.20 | .75 | .45 | .74 | -.14 |
|  | flat | .75 | 0.0 | .94 | - | .92 | .36 | .87 | -.52 |
| 4. Tag questions | natural | .91 | .83 | .64 | -.27 | .04 | .04 | .50 | 0.0 |
|  | flat | .87 | .83 | .57 | -.11 | 0.0 | .04 | .50 | - |
| 5. Missing NP elements | natural | .48 | .35 | .62 | .05 | .83 | .58 | .72 | -.27 |
|  | flat | .65 | .26 | .78 | .08 | .96 | .74 | .77 | -.67 |
| 6. Wh-movement: Fronted NPs | natural | .83 | .27 | .86 | -.17 | .83 | .25 | .87 | -.14 |
|  | flat | .78 | .24 | .85 | -.03 | .87 | .43 | .82 | -.37 |
| 7. Gapless relative clauses | natural | .70 | .14 | .86 | .27 | .91 | .38 | .86 | -.48 |
|  | flat | .65 | .36 | .72 | -.01 | .91 | .32 | .88 | -.45 |

**Table 4: Grammaticality Judgment Data for 3 Nonagrammatic Aphasic Listeners**

| Type | F0 Contour | JD |  |  |  | JS |  |  |  | HY |  |  |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  | p(Hit) | p(FA) | A′ | B′ | p(Hit) | p(FA) | A′ | B′ | p(Hit) | p(FA) | A′ | B′ |
| 1 | natural | .91 | .62 | .77 | -.48 | .65 | .50 | .63 | -.05 | .70 | .48 | .68 | -.08 |
|  | flat | .96 | .40 | .88 | -.72 | .48 | .40 | .58 | .02 | .78 | .50 | .73 | -.19 |
| 2 | natural | 1.0 | .42 | .90 | - | .46 | .35 | .60 | .04 | .64 | .26 | .78 | .09 |
|  | flat | 1.0 | .35 | .91 | - | .42 | .41 | .51 | 0.0 | .92 | 0.0 | .98 | - |
| 3 | natural | .92 | .45 | .84 | -.54 | .92 | .82 | .67 | -.33 | .83 | .36 | .83 | -.24 |
|  | flat | 1.0 | 0.0 | 1.0 | 0.0 | .83 | .64 | .69 | -.24 | .92 | 0.0 | .98 | - |
| 4 | natural | .22 | .04 | .75 | .63 | .91 | .74 | .71 | -.40 | .26 | .04 | .77 | .67 |
|  | flat | .09 | .04 | .65 | .36 | .65 | .78 | .30 | -.14 | .56 | .04 | .87 | .73 |
| 5 | natural | .96 | .63 | .81 | -.72 | .43 | .37 | .56 | .03 | .46 | .18 | .74 | .25 |
|  | flat | .96 | .68 | .79 | -.70 | .56 | .53 | .53 | -.01 | .74 | .26 | .82 | 0.0 |
| 6 | natural | .91 | .14 | .94 | -.19 | .70 | .52 | .66 | -.08 | .74 | .14 | .88 | .23 |
|  | flat | .96 | .14 | .95 | -.52 | .65 | .43 | .68 | -.04 | .74 | 0.0 | .94 | 1.0 |
| 7 | natural | .86 | .36 | .84 | -.31 | .65 | .41 | .69 | -.03 | .55 | .09 | .84 | .50 |
|  | flat | .91 | .41 | .85 | -.49 | .65 | .32 | .75 | .02 | .65 | .14 | .84 | .31 |

### Results and discussion

The control subjects performed well in the AGJ task. For them, mean sensitivity was computed and compared to the performance of young adults in Experiment 1, with the seven violation types as repeated measures. No difference in sensitivity was seen for the natural-F0 stimuli, **t**(6)=1.14, **p**>.10, but the older control subjects outperformed those in Experiment 1 with the flat stimuli, **t**(6)=5.60, **p**<.01. One reason for this result may be that in the present study improvements due to stimulus repetition are masking possible decrements in performance caused by the absence of intonation information in the flat stimuli. In fact, contrary to Experiment 1, these subjects showed no effect of intonation on sensitivity to grammaticality, **t**(6)=1.35, **p**>.10. Sensitivity was greatest for violation types 3 and 6, consistent with Experiment 1.

Table 3 shows AGJ performance of two agrammatic listeners. Both patients performed better than chance, but worse than the control listeners, and the subjects tested by Linebarger et al.[1] In particular, VS who was tested in the earlier report showed less sensitivity to syntactic structure in the present study. Several factors of our stimuli[5] may have contributed to this result. More importantly, both agrammatic patients showed least sensitivity to manipulations in tag questions, as predicted by the Linebarger et al.[1] results. Moreover, this syntactic structure elicited dramatic changes in decision criteria for both patients, but in opposite directions: FM accepted the majority of type 4 stimuli as grammatical, while VS rejected virtually all of them. Notably, FM also performed poorly with the two structures causing normal listeners most difficulty. Another difference between the two agrammatics is that FM, but not VS, showed normal effects of memory demand. AGJ performance was consistent across intonation conditions, suggesting that previous AGJ results with aphasics are probably not based on reliance on prosodic cues to syntax.

The final part of this study compares performance of other aphasic patients (see Table 4) to that of agrammatics. JS who has an asyntactic comprehension deficit but is fluent[5] could not perform the AGJ task. Only for the subject-auxiliary inversion manipulations was JS's performance above chance. Both JD and HY have milder deficits, and indeed, performed almost as well as normal listeners. Notably, JD had a lax decision criterion in all conditions except the tag questions -- there, he rejected most stimuli as nongrammatical, and showed least sensitivity. The similarity to the agrammatic pattern seen for VS exists, despite the overall high levels of performance. A similar bias shift was seen for the anomic HY with the tag questions, but it was not accompanied by decreased sensitivity. Instead, the verb control complement violations involving lexical rules were most difficult for HY. Also, JD and HY, but not JS, showed an advantage for early violations, suggestive of the normal pattern of memory effects. Most importantly, removal of F0 information did not alter the pattern among the syntactic structures for each patient.

## CONCLUSIONS

In sum, signal detection methods allowed detailed investigation of how aphasic listeners differ (from one another and from normal listeners) not only in selective sensitivity to syntactic structures, but also in response bias in the AGJ task. Experiment 2 has replicated and extended the results of Linebarger et al.[1] by reporting similar findings from 2 agrammatics and 1 nonagrammatic aphasic and different patterns for 2 other aphasics and for normal listeners. Intonation information did not influence the syntactic differences observed for both normal and aphasic listeners. For aphasics with nonsevere deficits and for normal listeners, performance suffered when memory load was increased. These results suggest a complex relationship between sentence intonation contour, memory demand, and response bias in grammaticality judgment tasks.

## *REFERENCES*

1. Linebarger, M.C., Schwartz, M.F., Saffran, E.M. (1983) Sensitivity to grammatical structure in so-called agrammatic aphasics. **Cognition, 13**, 361-392.
2. Berkovits, R. (1984) A perceptual study of sentence-final intonation. **Language and Speech**, **27**, 291-308.
3. Nespor, M. & Vogel, I. (1983) Prosodic structure above the word. In A. Cutler & D.R. Ladd (Eds.) **Prosody: Models and measurements** (pp.123-140). Berlin: Springer-Verlag.
4. Sorensen, J.M., & Cooper, W.E. (1980) Syntactic coding of fundamental frequency in speech production. In R.A. Cole (Ed.) **Perception and production of fluent speech** (pp. 399-440). Hillsdale, NJ: Erlbaum.
5. Berndt, R.S., Salasoo, A., Mitchum, C.C., & Blumstein, S.E. (1986) The role of intonation cues in the grammaticality judgment task. Submitted for publication.
6. Grier, J.B. (1971) Nonparametric indexes for sensitivity and bias: Computing formulas. **Psychological Bulletin, 75**, 424-429.