Title

Disfluency Deafness: Graceful Failure in the Recognition of Running Speech

Permalink

https://escholarship.org/uc/item/9vc946hf

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 20(0)

Authors

Bard, Ellen Gurman
Lickley, Robin

Publication Date

1998

Disfluency Deafness: Graceful Failure in the Recognition of Running Speech

Ellen Gurman Bard (ELLEN@LING.ED.AC.UK)
Robin Lickley (ROBIN@LING.ED.AC.UK)
Human Communication Research Centre
University of Edinburgh, UK

Abstract

Models of perceptual systems customarily characterize their maximally efficient operation in optimal circumstances. Another engineering consideration – graceful failure - is usually ignored. Three experiments on spontaneous speech show that on-line speech recognition fails gracefully by making us deaf to the words in reparanda, the items which must be expunged to restore disfluent utterances to fluency. Experiment 1 uses word-level gating of fluent and disfluent utterances to show that disfluencies principally disrupt normal late recognition (Bard, Shillcock & Altmann, 1988) of words in reparanda. Experiment 2 shows that in more natural listening conditions, attention to continuing material and additional effects of repetition deafness (Miller & Mackay, 1996) make recall of the same words even more unlikely. Experiment 3 shows that the results are not attributable to the clarity of the lost words. Finally the relationships among late recognition and various kinds of disfluency deafness are discussed.

Introduction

Assuming that perceptual systems evolve to cope efficiently with their characteristic input, we usually devote our attempts to understand such systems to cases where they operate with maximal efficiency. A second consideration in designing robust systems, often ignored in the study of human cognition, is graceful failure: failures, if they must happen, should occur in such a way that recovery is relatively easy. This paper deals with the mechanisms which promote graceful failure in the recognition of words in running spontaneous speech.

The occurrence of such failures is all too obvious to anyone who has had to transcribe or code spontaneous speech. Disfluencies occur at a rate of about one every three utterances in normal speech. Transcribing disfluent speech verbatim is inordinately difficult: the contents of the disfluency seem strangely evanescent. Without many replays of the material, even the location of the disfluency is difficult to ascertain. Asking subjects to monitor for disfluency, Martin and Strange (1968) found that instructions to increase attention to the task essentially increased bias to report disfluencies but did not improve accuracy in locating them. Though some disfluencies can be spotted (Duez, 1995), many present problems.

Graceful failure ought to be centred on words actually in the reparandum, the speech that must be expunged to create a fluent utterance. Fox Tree (1995) has shown that words in reparanda enter and affect the lexical access system, but we have no indication of what happens next.

The behavior of automatic speech recognition systems indicates that these words are not necessarily indecipherable. Unlike human listeners, ASR systems seem to have similar records of words in fluent and in disfluent speech. The difficulty arises when the system attempts to rank alternate candidate sequences of words in the light of its language model. Since disfluencies almost always make strings ungrammatical, the best guess – usually the most nearly grammatical – strings provide very unsatisfactory transcriptions. The chief need for such systems is a way of expunging disfluencies on the fly. If human perceptual failure is more graceful here, then it is due to some characteristic of human perception which is not mirrored in ASR systems.

This paper shows that human failure to perceive disfluent speech is graceful and that it can be attributed to two characteristics of human perception: the dependence of on-line word recognition on subsequent as well as prior context and the recall phenomena usually described as repetition deafness.

Normally, listeners depend on both preceding and subsequent context to recognize words in running speech (Bard, Shillcock, & Altmann, 1988; Connine, Blasko & Hall, 1991; Grosjean, 1985). While most words can be recognized as soon as they and their prior contexts have been heard, some are not identified until a prosodic or constituent boundary occurs up to several words later in the utterance (Shillcock, Bard & Spensley, 1988). The more prior context a word has, however, the more likely immediate recognition is. When disfluencies interrupt speech, they interrupt both contexts on which listeners depend. As Table 1 illustrates, the fifth word in the fluent utterance, (*further*), has four words of prior context which can be construed together. In the disfluent examples, words after I, the interruption point (*a, not*), have shorter continuously construable prior contexts. Because they reduce the supporting prior context, disfluent interruptions create conditions where subsequent context should be important to the recognition process. Whether recognition can recover from the interruptions depends on how much of the utterance follows. For words before interruption points, particularly words in reparanda, like (*bit*), the discontinuity truncates subsequent context, usually before the next expected prosodic or constituent boundary. For these words, disfluencies may remove or delay those sites where late recognition would normally occur.

Unlike effects of context, repetition deafness (Miller and Mackay, 1996) has not yet been shown to influence online recognition of spontaneous speech. Repetition deafness and blindness (Kanwisher, 1987) are inabilities to distinguish in recall two very similar stimuli witnessed close together in time, particularly in presentation modes (e.g. rapid list intonation, time-compressed speech) which make perception and encoding very difficult. Disfluent speech may provide naturally-occurring circumstances for repetition deafness: words whose right context is delayed and which are repeated may be harder to recall as subsequent input becomes the focus of attention.

Experiment 1 – word-level gating

Experiment 1 uses word-level gating to determine whether the disruptive effects of disfluency create failures of late recognition particularly for reparanda. The gating method increments the presented portion of an utterance by one word on each successive trial, giving listeners information about the number and location of words which they would have to discern for themselves in more normal circumstances. The experiment therefore compares listeners' recognition of words in disfluent utterances and in length-matched fluent utterances (Table 1) under conditions which should maximise rates of recognition.

Table 1: Example Stimuli: Repetition, recast and fluent control. OU = Original Utterance, RM = Reparandum, I = Interruption point RR = Repair, CO = Continuation

TYPE	OU	RM	I	RR	CO
Repetition	it's just	*a bit*	.	a bit	further up
Recast	it's just	*a bit*	.	not very	far up
Fluent	it's just	a bit	.	further up	the road

If disfluencies interfere with contextual support for running speech recognition, several predictions should be fulfilled. First, there should be more failures to identify words from disfluent items than from fluent items of the same total length. Second, the difficulties should cluster about the interruption point. The greatest rate of outright failure should be in the reparandum, where subsequent context is truncated. Following the interruption point, continuous prior context is initially minimal and the disfluent interruption should not support immediate recognition as well as in the uninterrupted control. Third, eventual recognition of of pre-interruption words should depend on the type of disfluency. If the disfluency is a recast, the context which would permit late recognition of reparandum words may never arrive and the repair words will lack prior context. If the disfluency is a repetition, then the pertinent later context is merely postponed. Repetitions should therefore support more successful late word recognitions than recasts.

Method

Materials. All materials were spontaneous utterances from the HCRC Map Task Corpus (Anderson *et al.* 1991). A subsection of the corpus was word segmented and coded for disfluencies via Entropic xwaves and xlabel software using waveform and spectrographic representations. Twenty-eight disfluent repetitions and 28 recasts containing no repetition were selected as disfluent stimuli. Half of each group ended in whole words and half in word-fragments. Each disfluent utterance was paired with a fluent utterance of the same length in words, produced by the same speaker. These 112 test utterances and 56 fluent fillers were distributed among 4 tapes by Latin square and blocked by speaker (N = 11).

Subjects and Procedure. Subjects were 16 members of the Edinburgh University community, all native speakers of English with no known hearing loss. Four subjects heard each tape. Each subject heard all 56 filler utterances and 56 test utterances, 14 from each cell of the design. Subjects were told that they would hear utterances beginning with the first word and then including one additional word until the utterance was complete. Their task was to identify each new word as soon as they had heard it, writing it on an answer sheet which allowed one block for each word presented on each trial. They were encouraged to guess and allowed to alter their transcription for any word on the line corresponding to the trial where they changed their mind, but not to alter previous lines.

Results

Two faulty recast items and their fluent controls were discarded, leaving 14 repetitions and 12 recasts. Over all subjects and materials, the data comprise 4384 attempts, usually over multiple trials, to recognize spoken words, half in fluent and half in disfluent utterances. A word received an *immediate recognition* if correctly identified by a subject on its first presentation with only prior context, a *late recognition* if first recognized only after at least one additional word, and a *failed recognition* if never correctly transcribed.

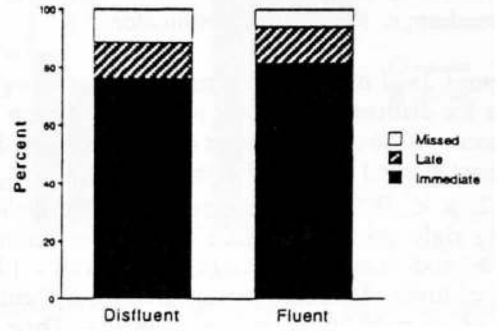

Figure 1: Experiment 1: Distribution of outcomes of all attempts (N = 4384) at recognizing words in disfluent items and their fluent controls.

Figure 1 shows that, as predicted, words in disfluent utterances are the more difficult to recognize ($\chi^2_{(4384)} = 48.82$, *df* = 2, $p < .0001$). Disfluent items yielded fewer immediate recognitions than fluent (76.1% v 81.7%) and more failures (11.5% v 5.6%), while late recognitions occurred at a similar rate in the two (12.4% v 12.7%). The largest component of χ^2 was contributed by the difference in rates of failure (2×2.2).

As predicted also, difficulties clustered around the interruption point, with failures peaking where they are most graceful. To test this proposal, each disfluent utterance was divided into 4 parts set out in Table 1: the *reparandum* (RM) immediately preceded the interruption point and contained words retraced or repeated; the *original utterance* (OU) preceded the RM; the *repair* (RR) immediately followed the interruption and consisted of either a genuine replacement for the RM or of a string of words equal in length to the RM; the *continuation* (CO) concluded the utterance. Fluent utterances were divided at the same points as their disfluent counterparts for purposes of comparison.

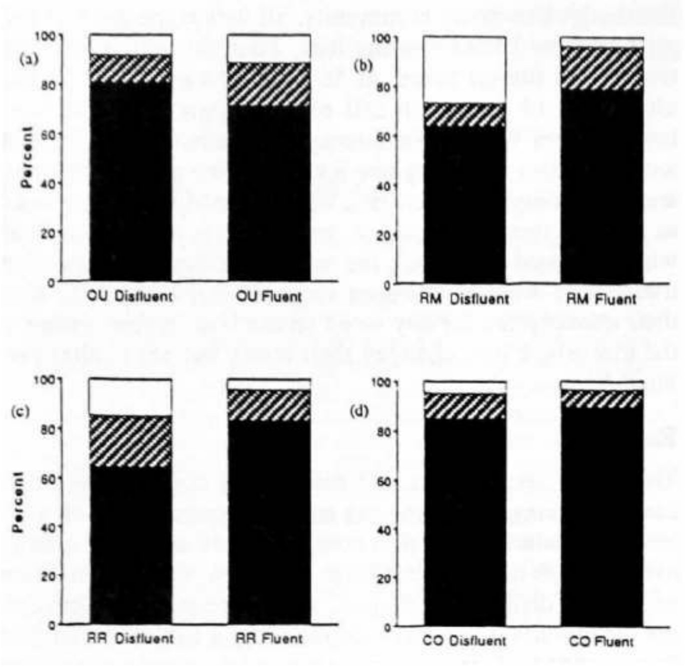

Figure 2: Experiment 1: Distribution of outcomes of all attempts at recognizing words in disfluent items and their fluent controls by part of utterance: a. Original Utterance; b. Reparandum; c. Repair; d. Continuation.

Figures 2a–d display the distributions of recognition outcomes for disfluent and fluent items within each part of the utterance. All four components show significant differences. The most marked are found in the reparanda ($\chi^2_{(768)} = 84.00$, $df = 2$, $p < .0001$). As we predicted, the disfluent RMs, lacking right context, produce many more failures (26.8% v 3.6%) and fewer late recognitions (9.1% v 17.7%) than fluent controls. Disfluent repairs are also difficult to recognize ($\chi^2_{(864)} = 46.16$, $df = 2$, $p < .0001$). Their effectively truncated left context results in fewer immediate recognitions (64.4% v 83.6%) and more late recognitions (20.8% v 12.3%) and failures (14.8% v 4.2%) than the corresponding parts of fluent controls. The disproportionate rate of failures makes the largest contribution to χ^2 for both RMs and RRs, though the rate is higher for RMs than for RRs (26.8% v 14.8%). Disfluent original utterances, with abbreviated right context, produce fewer late recognitions than their fluent counterparts (11.2% v 17.5%: $\chi^2_{(1144)} = 14.20$, $df = 2$, $p < .0008$). Finally, continuations yield more late recognitions than are needed in the final portions of their fluent controls (10.3% v 7.1%: $\chi^2_{(1608)} = 8.04$, $df = 2$, $p < .02$).

Finally, recognition outcomes also depend on the relationship between what precedes and what follows the interruption point. Repetition disfluencies, where the two are more likely to be parsable as a single sequence once extra tokens of repeated words are removed, are more successfully recognized than recast disfluencies, where reconstructing an utterance would be more difficult. As figure 3 shows, recasts produce more failures to recognize words (13.3% v 9.9%: $\chi^2_{(2192)} = 6.86$, $df = 2$, $p < .04$).

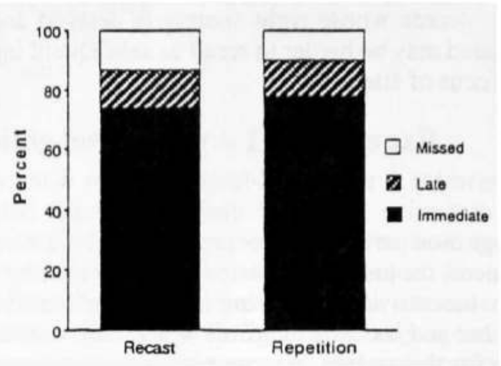

Figure 3: Experiment 1: Distribution of outcomes of all attempts at recognizing words in recast ($N = 1004$) and repetition ($N = 1188$) disfluencies.

Experiment 2 – transcription

Experiment 1 indicated that disfluent utterances are subject to disruption even in a paradigm that provides optimal conditions for successful word recognition. By identifying word boundaries, the word-level gating method removes one of the major problems in recognizing running speech. The second experiment gives listeners the less artificial task of reporting words larger chunks of speech. A large-scale verbatim transcription task was designed with two purposes. First, it checked for recognition failures in the more natural task. Second, it tested the hypothesis that repetition deafness will make recall even worse for disfluencies which contain repeated words than for those which do not. To compare effects of repetition per se with other characteristics of disfluency, we presented both recast disfluencies and repetition disfluencies up to the end of the reparandum, up to the end of the first word of the repair where the disfluency should first be noticeable (Lickley, Shillcock, & Bard, 1991), up to the end of the repetition or to an equivalent position in recasts, or in their entirety. Marked decrease in report of reparandum words from one chunk to the next should indicate which kind of additional material is implicated in recognition failure. If interruption is responsible, deficits should appear as soon as interruptions are encountered; if disfluent repetitions are to blame, only stimuli including them will suffer. If the problem is processing pressure, deficits should increase as more right context is included.

Method

Materials Eighty simplex disfluencies, each containing a single contiguous reparandum, included 30 recasts and 50 with repetitions. The remaining 16 disfluencies were complex, containing either multiple attempts to repeat or replace the reparandum or a series of different disfluencies. For 6 of these, the ultimate repair did not repeat any word in the reparandum, while for the other 10, repetition was involved. For each disfluent utterance, there was a fully fluent control utterance matching it for speaker and length in words.

Table 2: Stimuli for two kinds of disfluency. Reparanda in bold, repairs in italics. Interruption points ({IP}) were not indicated to subjects in any way

CHUNK	DISFLUENCY TYPE	
	REPETITION	RECAST
a	Right **there's a** {IP}	**There's about** {IP}
b	Right **there's a** {IP} *there's*	**There's about** {IP} *You've*
c	Right **there's a** {IP} *there's a*	**There's about** {IP} *You've got*
d	Right **there's a** {IP} *there's a* line about half way down	**There's about** {IP} *You've got* a yacht club right

As table 2 illustrates, for each of the 96 disfluent utterances, four substrings were prepared. All began at the beginning of the utterance. The first (chunk a), ran up to the interruption point, the second (b) ran up to the first word of the repair (at which point listeners can usually detect that a disfluency has taken place , the third (c), to the end of any repetition, or, for recasts, to the end of the next stressed word, and the fourth (d), to the end of the utterance. Control utterances were segmented at the corresponding serial positions. The substrings of each utterance were distributed by Latin square among four listener groups to give substring comparison between subjects and fluency comparisons within subjects.

Subjects and procedure Subjects were University of Edinburgh students, with no known hearing loss. Nine were assigned to each listener group. Listeners were instructed to transcribe everything they heard into real words in the standard orthography and to be as accurate as possible even though some of the stimuli were difficult or odd. They were not told how many words any stimulus contained. Stimuli were presented three times in succession via high quality headphones. A transcription was required after each presentation.

Results

We report analyses of first pass attempts at recall and transcription, the most natural listening condition.

As gating results would predict, listeners had great difficulty in reporting words from reparanda (Figure 4)

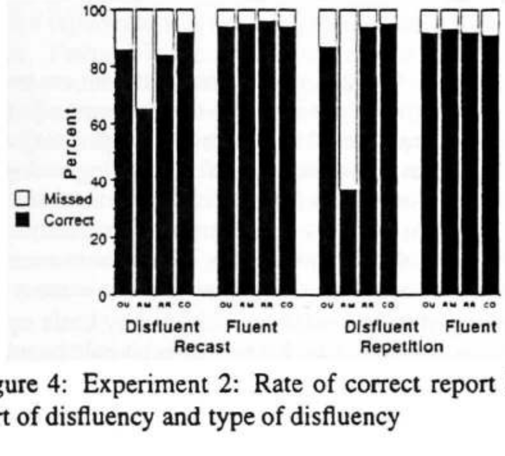

Figure 4: Experiment 2: Rate of correct report by fluency, part of disfluency and type of disfluency

Identification of words in reparanda was markedly worse than in Experiment 1. Control materials showed slight, insignificant improvement with longer stimuli. Recall of words in reparanda was actually worse in the longest strings, where the completion of the utterance could have allowed late recognition, than it was in the shortest strings, where only immediate recognition was possible (Figure 5). Scoring whole reparanda and corresponding control words as right or wrong, the interaction between fluency and chunk-length (a-d) was highly significant ($F_1(3,105) = 122.10, p < .0001$; $F_2(3,282) = 49.47, p < .0001$). All fluent outcomes were significantly better than any disfluent (Scheffés at $p < .01$). Recall for disfluent reparanda was significantly better in the stimuli (a) which stopped at the interruption point than at *any* of the longer substrings (at $p < .01$). The difference was not merely the effect of encountering a discontinuity at the point of interruption: chunk-d, the whole utterance, gave significantly worse recall than chunk-b, which contained the first word of the repair.

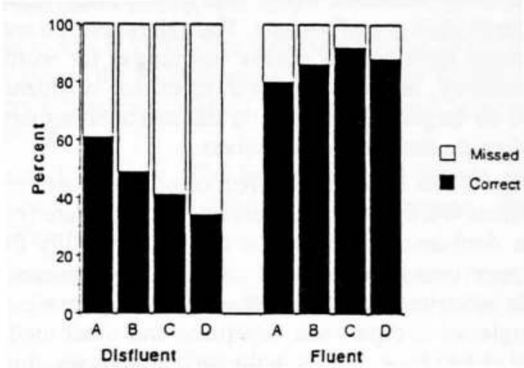

Figure 5: Experiment 2: Rate of correct report by substring length for words in reparanda of disfluent utterances and for corresponding words of fluent controls.

Four kinds of evidence bear on the second hypothesis, that repetition deafness helps to expunge disfluencies. Repetition itself did not produce disastrous reductions in the recall of the repeated word (at chunk-b for single-word reparanda, chunk-c for others). Instead, stress on memory or processing load seemed to promote special deficits for repetition disfluencies.

First, there is a repetition deficit. For the most vulnerable words, those just preceding the interruption point, report rate falls more sharply in repetition disfluencies than in others. In an analysis of recall loss, i.e., how much recall rates changed from chunk-a to the later chunks,, the down-turn for repetitions was particularly marked: (fluency x chunk x disfluency type: $F_2(1,367) = 450, p < .035$). Again, the fall-off in recall continued beyond the point where the repetition occurred (chunk-b/c) and so may be due to the memory load created by the additional words in chunk-d. Fluent controls showed no comparable trends.

Second, extensive exploration of the results by multiple regression analyses showed that recall results for repetition and recast disfluencies were subject to somewhat different influences. All words from disfluent stimuli were coded for dictionary characteristics of the words (raw frequency, functor/contentive word class), for their characteristics as uttered tokens (strength of following phonological boundary, from sentence boundary at 3 down to functor-contentive boundary at 0; word duration; duration of following pause; stress, from 2 for pitch accent to 0 for no stress), and for characteristics of their location in a disfluent utterance (length of chunk-a; distance of word from interruption point; number of words in RM; number of words in utterance). All words from fluent stimuli were coded for the same variables, with characteristics of the matched disfluent partner used for certain position variables.

In a set-hierarchical multiple regression, equations including characteristics of the disfluent utterance always accounted for significantly more of the variance in recall rate than equations lacking these variables. All words showed effects of the structure of their disfluency: proximity to the interruption and longer sequences of words before the interruption point made for worse recall. Words preceding more important prosodic and syntactic boundaries were reported better. After chunk-a, however, only the recall of recast words depended on length and frequency variables, which would have made these words more intelligible out of context. Though repetition and recast disfluencies have similar means and ranges for word length and frequency, recall of words in repetition disfluencies appears to be largely dependent on the surrounding structures, rather than on the words themselves.

Third, we can see a direct effect of prior context on results for the final word of the reparandum. We compare results for simplex disfluencies, where the utterance is fully fluent up to the interruption point, with complex disfluencies, where multiple interruptions disrupt the string (Figure 6). In simplex single-word reparanda, repetition and other disfluencies behaved alike ($F_2 < 1$). As in the earlier analyses, fluent control words were somewhat easier to report when more context was presented, while reparandum-final words were reported less accurately in longer strings (fluency x chunk: $F_2(2,68) = 14.06$, $p < .0001$). For the complex disfluencies, there is both a detrimental effect of longer stimuli ($F_2(2,68) = 10.88$, $p < .0003$) and an additional deficit for repeated words (disfluency type x fluency x chunk: $F_2(2,68) = 3.81$, $p < .035$). In other words, repetition disfluencies are significantly more forgettable than others when they occur in utterances which are already difficult to process because of multiple false starts.

The final evidence for repetition deafness as a function of processing pressure is the difference between the results of Experiments 1 and 2. In gating, word boundaries are indicated, since each word forms the end of some stimulus, and subjects hear very little new material on each trial. Words in repetitions were recognized somewhat better as subsequent context accrued. In the transcription technique, where word boundaries were not marked, and recognition of many words was required on a single trial, the same disfluencies showed significantly more tendency to suffer from additional context (fluency x task: $F_2(1,388) = 10.16$, $p < .01$). Though they improved ultimately in gating, repetitions were worse than recasts in Experiment 2.

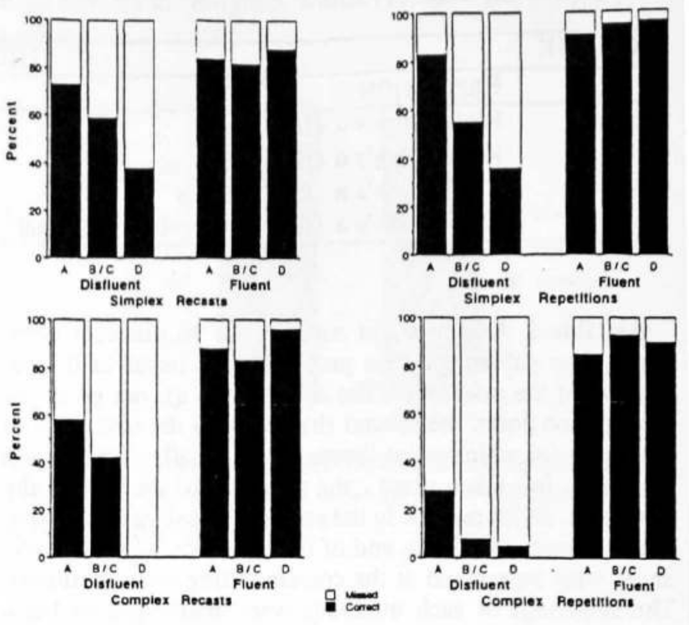

Figure 6: Experiment 2: Rate of correct report for final word of RM and for corresponding word of fluent controls by fluency, type of disfluency and complexity.

Experiment 3 – isolated word intelligibility

The transcription experiment showed that words in reparanda were perceptually vulnerable. But was this simply because they were less intelligible *per se* than the control words? To answer this question an intelligibility study was run on words isolated from the reparanda used in the previous experiment, words from the fluent controls and a further set of words from fluent utterances.

Method

Materials Eighty-four full words from disfluent reparanda used in the previous experiment and the corresponding 84 words from their length-matched controls (control set) were carefully isolated using speech waveform editing software. Since the control words were not matched for lexical identity with the reparandum words, a further set of 84 lexically identical words (matched set) was selected from fluent contexts uttered by the same speakers and prepared in the same way.

The resulting set of 252 words was divided by Latin square into two groups in such a way that no subject would hear both members of any pair from the disfluent and matched sets of stimuli while all subjects would hear the member of the control set corresponding to both other sets. A further division was then made, both to prevent too much repetition of different tokens of the same lexical item in any one set of data and to reduce the length of the experiment. This gave 4 sets of 84 stimuli.

Subjects and procedure Subjects were 20 students of the University of Edinburgh, all native speakers of English with no known hearing loss. Five subjects heard each of the four sets of stimuli. They were asked to try to give a full-word transcription of each word that they heard. Only one attempt was allowed at each stimulus.

Results

Recognition outcomes were classed as "correct" or "wrong" and analyses performed on percent correct responses per subject and per item.

Overall, the three sets of stimuli did not differ in intelligibility ($F_1(2,38) = 2.19$, $p > .1$; $F_2(2,166) = 0.46$, $p > .6$). There was some evidence that words from repetition reparanda were less clear than their controls : two-way ANOVAs revealed a source (disfluent, control, matched) by type (repetition, recast) interaction significant only by subjects (Source x type: $F_1(2,38) = 4.61$, $p < .02$; $F_2(2,164) = 1.19$, $p > .1$) (Table 3) but differences between the critical cells were not significant (Scheffés at $p < .05$).

Table 3: Experiment 3: Percent intelligibility for words in Disfluent, Control and Matched sets by disfluency type.

TYPE	DISFLUENT	CONTROL	MATCHED
Recast	58.25	49.53	58.01
Repetition	46.69	54.99	53.28

Discussion and Conclusions

Experiment 1 showed that disfluencies were subject to exactly those disruptions which are predicted from listeners' reliance on subsequent as well as prior context in the recognition of words in running speech. The disruption is worst where it is most graceful: in the reparandum. Disruption is also quite severe for the repair which follows the interruption point. This effect may be more graceful than it looks: in many disfluencies, the repair is not a fresh restart. Instead it must be interpreted with some part of the original utterance, and in some cases the interpretation has to be fairly loose. Mere excision of the reparandum is not enough to yield a fully fluent utterance. Perhaps less perfect recognitions of repairs allow the listener some room for reinterpretation in these cases.

Experiment 2 showed that the additional stress induced by operating in real time made graceful failure more severe. Repetition deafness also contributed to the effect: repetition disfluencies induced an additional penalty in longer stimuli. Experiment 3 showed that the words least often recognized in context were as intelligible in isolation as other tokens of the same words and as their controls in Experiments 1 and 2. Mere clarity of articulation was therefore not a major cause of perceptual failures.

What is striking is the similarity between the factors contributing to graceful failure in repetition and recast disfluencies. Although context was more important in the case of repetitions, both showed disruption effects in gating and a tendency to disappear from recall with longer presentations. It appears that perceptual difficulty is created by the disruption of context and that the unresolved items become harder and harder to convert into a form which can be lodged in memory as the listener's attention is diverted to new material. It seems likely that repetition deafness may be a special case of more general inability to operate under perceptual pressure.

Acknowledgements

This work was funded by ESRC grant R000235212 to the authors and by EPSRC SALT grant GR/L50280 to H. Thompson, S. Isard, and the first author.

References

Anderson, A.H., Bader, M., Bard, E.G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H.S. & Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech, 34*, 351–366.

Bard, E.G., Shillcock, R.C. & Altmann, G.T.M. (1988). The Recognition of Words after their Acoustic Offsets in Spontaneous Speech: Effects of Subsequent Context. *Perception and Psychophysics, 44(5)*, 395–408.

Connine, C.M., Blasko, D. & Hall, M. (1991). Effects of Subsequent Sentence Context in Auditory Word Recognition – Temporal and Linguistic Constraints. *Journal of Memory and Language, 30(2)*, 81–94.

Duez, D. (1995). Perception of Hesitations in Spontaneous French Speech. *Proceedings of the International Congress of Phonetic Sciences*, 498–501.

Fox Tree, J.E. (1995). The Effects of False Starts and Repetitions on the Processing of Subsequent Words in Spontaneous Speech. *Journal of Memory and Language, 43*, 709-738.

Grosjean, F. (1985). The Recognition of Words after their Acoustic Offset: Evidence and Implications. *Perception and Psychophysics, 38(4)*, 299–310.

Kanwisher, N.G. (1987). Repetition Blindness: Type Recognition without Token Individuation. *Cognition, 27*, 117–143.

Lickley, R.J., Shillcock R.C. & Bard, E.G. (1991). Processing Disfluent Speech: How and When are Disfluencies Found?. *Proceedings of Eurospeech 91*, 1499-1502.

Martin, J.G. & Strange, W. (1968). The Perception of Hesitation in Spontaneous Speech. *Perception and Psychophysics, 3(6)*, 427–438.

Miller, M.D. & MacKay, D.G. (1996). Relations between Language and Memory – the Case for Repetition Deafness. *Psychological Science, 7(6)*, 347–351.

Shillcock, R.C., Bard, E.G. & Spensley, F. (1988). Some Prosodic Effects on Human Word Recognition in Continuous Speech. *Proceedings of Speech '88: 7th FASE Symposium*