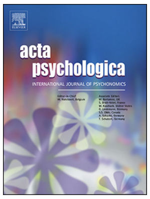
Acta Psychologica 191 (2018) 69–75

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy



Auditory (dis-)fluency triggers sequential processing adjustments[☆]

Thomas Dolk[a,*], Claudia Freigang[b], Johanna Bogon[a], Gesine Dreisbach[a]

[a] *Department of Psychology, University of Regensburg, Regensburg, Germany*
[b] *Hearing Diagnostics Limited, Edinburgh, United Kingdom*

ARTICLE INFO

Keywords:
Aversive signals
Effort mobilization
Hearing
Perceptual fluency
Sequential processing adjustments

ABSTRACT

An increasing amount of studies indicates that experiencing increased task demands, triggered for example by conflicting stimulus features or low perceptual fluency, lead to processing adjustments. While these demand-triggered processing adjustments have been shown for different paradigms (e.g., response conflict tasks, perceptual disfluency, task switching, dual tasking), most of them are restricted to the visual modality. The present study investigated as to whether the challenge to understand speech signals in normal-hearing subjects would also lead to sequential processing adjustments if the processing fluency of the respective auditory signals changes from trial to trial. To that end, we used spoken number words (one to nine) that were either presented with high (clean speech) or low perceptual fluency (i.e., vocoded speech as used in cochlear implants—Experiment 1; speech embedded in multi-speaker babble noise as typically found in bars—Experiment 2). Participants had to judge the spoken number words as smaller or larger than five. Results show that the fluency effect (performance difference between high and low perceptual fluency) in both experiments was smaller following disfluent words. Thus, if it's hard to understand, you try harder.

1. Introduction

“*Oh, I missed that last sentence; could you repeat, please?*”

There are several instances during the day that challenge our understanding: It typically starts in the morning when the news from the radio are hard to understand due to the sound of the coffee machine, the croaking of the kids playing around, or the traffic sounds from the street and so on. While we get used to this or similar situations over time, there are several other quite demanding and often very annoying instances such as a bad connection via telephone/skype or a poor signal-to-noise ratio at the railway station that require immediate behavioral adjustments to handle the challenge at hand appropriately. The ease with which information is processed is referred to as (perceptual) fluency. Derived from previous research showing that visual (dis-)fluency can be used to predict processing efforts and corresponding behavioral adjustments (Dreisbach & Fischer, 2011; Song & Schwarz, 2008), we aim to show that the cognitive system dynamically adjusts to changing fluency experiences also in the auditory domain.

In the lab, flexible behavioral adjustments to contextual challenges can be observed in classical stimulus-response-compatibility tasks like the Eriksen flanker, the Simon, and the Stroop task. Here, the response conflict is reflected in increased response times and error rates for incompatible trials when compared to compatible trials. The response conflict is typically reduced after the presentation of an incompatible stimulus relative to a compatible stimulus—an effect known as the conflict adaptation effect or Gratton effect (Gratton, Coles, & Donchin, 1992; for a review, see Egner, 2014). According to *conflict monitory theory* (CMT; Botvinick, Braver, Barch, Carter, & Cohen, 2001; for a review, see Botvinick, Cohen, & Carter, 2004), the interference by task-irrelevant information of an incompatible stimulus is reduced as a consequence of cognitive control already being up-regulated in the trial following an incompatible (conflicting) trial.

While the initial notion of CMT proposed that it is the conflict as such that triggers control adjustments, more recent attempts in the field suggest that it is the aversive character of the conflict signal that originally motivates processing adjustments (Botvinick, 2007; Dreisbach & Fischer, 2012, 2015, 2016; Inzlicht, Bartholow, & Hirsh, 2015; Saunders, Lin, Milyavskaya, & Inzlicht, 2017). Evidence in favor of this claim first came from studies showing that in fact, conflict stimuli are experienced as aversive (Dreisbach & Fischer, 2012; Fritz & Dreisbach,

---

[☆] We would like to thank Anja Klaßen and Sophia Geisenhofer for data collection.

[*] Corresponding author at: University of Regensburg, Department of Psychology, Universitätsstraße 31, 93053 Regensburg, Germany.
E-mail address: thomas.dolk@ur.de (T. Dolk).

https://doi.org/10.1016/j.actpsy.2018.08.021
Received 19 March 2018; Received in revised form 9 July 2018; Accepted 31 August 2018

2013, 2015; Schouppe et al., 2015; Braem et al., 2017). However, showing that conflict stimuli are in fact experienced as aversive signals does not allow inferring their function as a trigger for control adjustments. After all, the aversive character of conflict stimuli might only be an epiphenomenon of conflict processing. Indirect evidence supporting the idea that this aversive conflict signal is the driving source of conflict adaptation (and not the conflict proper), has first been provided by van Steenbergen, Band, and Hommel (2009, 2010, 2012): In one of their studies, random monetary gain cues in an arrow-version of the Eriksen flanker task eliminated conflict adaptation. This was taken as evidence that positive signals counteract the aversive signal and thus eliminate the trigger for control adjustments (van Steenbergen et al., 2009). Supporting this line of reasoning, recent behavioral and electrophysiological findings showed that the (metacognitive) subjective (and presumably unpleasant) experience of a given conflict modulates the conflict adaptation effect (Desender, Van Opstal, & Van den Bussche, 2014; Fröber, Stürmer, Frömer, & Dreisbach, 2017; Questienne, Van Opstal, van Dijck, & Gevers, 2016; Schouppe et al., 2015). For example, in the study by Fröber and colleagues (2017) participants were asked after every (Simon) trial, whether they had experienced the just executed task as rather pleasant or unpleasant. The lateralized readiness potential (LRP) was registered concurrently as a measure of conflict adaptation (Stürmer & Leuthold, 2003). Results showed that not only did participants rate Simon incompatible trials less pleasant than Simon compatible trials. Additionally, early LRP activation was found to be evoked by wrong responses following Simon incompatible trials that were rated as pleasant. Such LRP activation, however, is usually reduced as a result of conflict adaptation (see Stürmer & Leuthold, 2003; Stürmer, Leuthold, Soetens, Schröter, & Sommer, 2002; for reviews see, Duthoo, Abrahamse, Braem, Boehler, & Notebaert, 2014; Egner, 2014). It thus seems, that it is in fact the aversive conflict signal that serves as a trigger for control adaptations. Consequentially, one question that arises: Is the response conflict obligatory in order to find processing adjustments or would any increased task demand that is experienced as aversive result in sequential processing adjustments as well?

One study that directly addressed this question manipulated the fluency of processing of written number stimuli (Dreisbach & Fischer, 2011). In the absence of any response-conflict manipulating the fluency of processing can be used as means to manipulate the affective value of otherwise neutral stimuli: That is, the ease or difficulty of cognitive processing (i.e., the *perceptual fluency*) is modulated for example by modifying the audibility of the signal, the stimulus contrast, the font-type, or the presentation duration. Disfluent stimuli or disfluent presentation of information, appears to be more difficult, effortful, time-consuming, strange, less truthful and even more dangerous (Dragojevic & Giles, 2016; Lev-Ari & Keysar, 2010; Oppenheimer & Frank, 2008; Rhodes & Castel, 2008; Sanchez & Khan, 2016; Yue, Castel, & Bjork, 2013; for reviews see, Oppenheimer, 2008; Reber & Greifeneder, 2016; Schwarz, 2004; Winkielman, Schwarz, Fazendeiro, & Reber, 2003). Yet, increasing the demand of information processing does not only lead to affective reaction in the long run (Song & Schwarz, 2008), it immediately motivates sequential processing adjustments (Dreisbach & Fischer, 2011).

In that latter study, participants had to judge number words from one to nine except five presented randomly either fluent (i.e., Arial, black, 12 pt.) or disfluent (Mistral, light gray, 18 pt) as either smaller or larger than five. Despite a fluency effect showing faster responses to fluent trials as compared to disfluent ones, the results showed that participants invested more effort to read the stimuli as indicated by smaller fluency effects after disfluent compared to fluent trials. Note, that these sequential behavioral adjustments to manipulations of processing fluency occurred in the absence of any (stimulus-stimulus or stimulus-response) conflict providing support for the idea that aversive stimulus features alone can trigger processing adjustments (Dreisbach & Fischer, 2016).

Taken together, while sequential processing adjustments have been shown for different tasks and paradigms, most of them are restricted to the visual modality. Given, however, that auditory challenges on the daily basis are more frequent as opposed to visual ones and as such perhaps even more ecologically valid, the present study aimed at providing further evidence for the aversive character of processing demands and investigate, whether sequential adjustments to disfluency can also be found for the auditory modality (i.e., hard to understand spoken words). While there are many high-level similarities between the processes required for reading and listening (Bemis & Pylkkänen, 2012; Hagoort & Brown, 2000), there are significant differences in the processing of visual and auditory information that do not allow for a straight-forward adoption of the fluency findings in the visual to the auditory system: Written text is represented in a temporally stable spatial map on the retina and the visual cortex, and does not require the ability to integrate visual information in a temporally dynamic manner. In contrast to this, speech is fleeting and consists of dynamically varying temporal information, which requires the auditory system to store the acoustic information such that acoustic information can be integrated over time. Thus, auditory processing of speech relies on intact memory processes such as perceptual and working memory to allow for successful speech understanding (Rönnberg et al., 2013; Rönnberg, Rudner, Lunner, & Zekveld, 2010). Considering this significant difference in processing of written and spoken words, it is not immediately clear whether the sequential processing effect to visually disfluent words would be evoked in the auditory domain for disfluent spoken words. Thus, the present study in itself presents a novel approach on elucidating the effect of disfluent acoustic stimuli on the processing of speech.

To that end, we adapted the experimental design of the initial study by Dreisbach and Fischer (2011) from the visual to the auditory modality. Accordingly, spoken number, that had to be judged as words smaller or larger than five, were either presented at high/fluent (i.e., clean speech in quiet) or low/disfluent perceptual fluency (i.e., vocoded speech as used in cochlear implants presented in quiet [Experiment 1] or undistorted speech presented in a multi-speaker babble [Experiment 2]).[1]

Based on the literature reviewed above, we hypothesize that the auditory disfluency of the vocoded and the multi-speaker babble speech is harder to understand and should therefore result in worse performance (slower RTs and/or higher ERs) as compared to clean fluent speech (i.e., Fluency effect in Trial $N = RT_{disfluent} - RT_{fluent}$ or $ER_{disfluent} - ER_{fluent}$, respectively). Moreover, we assume that the aversive experience of vocoded or speech in multi-speaker babble will

---

[1] The stimulus materials in Experiment 1 and Experiment 2 were chosen so as to capture the effect of disfluency on the two main categories of acoustic information that the auditory system is challenged with most frequently: (*i*) frequency processing (vocoded stimulus in Experiment 1), and (*ii*) signal in noise processing (speech-in-noise in Experiment 2).

The stimulus in condition (*i*) was chosen so as to mimic the condition of disfluent letters in the visual task in the study from Dreisbach and Fischer (2011). The sensory epithelium of the retina represents visual information in a spatially organized manner, and disfluent words therefore create a blurry spatial map of the word. The cochlea, on the other hand, is organized in a frequency-selective manner and can be viewed as a series of bandpass filters. Thus, disfluent acoustic frequency information is the closest method for creating acoustic stimuli that correspond to disfluent written words.

The speech in noise condition (*ii*) was chosen in order to investigate whether the disfluency effect can be found for complex acoustic signals in which a target sound is competing with background sounds—a situation that very closely resembles everyday communication environments, and the processing of which is the most challenging task that the auditory system is faced with. The processing of speech in noise depends on the ability of the auditory system to separate a target sound from the rest. This ability relies on the frequency-selectivity as well as temporal processing and memory capacity, as the speaker-specific temporal dynamics need to be integrated over time in order to allow for a successful segregation of speakers.