Journal of Communication Disorders 73 (2018) 34–49

Contents lists available at ScienceDirect

Journal of Communication Disorders

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

Perceived listener effort as an outcome measure for disordered speech☆

Kathleen F. Nagle*, Tanya L. Eadie

Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, 98105, USA

ARTICLE INFO

Keywords:
Perceived listening effort
Electrolarynx
Acceptability
Intelligibility
Speech outcomes

ABSTRACT

Purpose: Perceived listening effort is a perceptual dimension used to identify the amount of work necessary to understand disordered speech. The purpose of this study was to investigate the utility of perceived listening effort to provide unique information about disordered speech. The relationships between perceived listening effort and two current outcome measures (speech acceptability, intelligibility) were examined for listeners rating electrolaryngeal speech, along with their reliability and intra-rater agreement.

Methods: Ten healthy male speakers read low-context sentences using an electrolarynx. Twenty-five inexperienced listeners orthographically transcribed and rated the stimuli for perceived listening effort and speech acceptability using a visual analog scale. Strict reliability and agreement criteria were set.

Results: Perceived listening effort was moderately to strongly correlated with intelligibility ($r = -0.76$) and acceptability ($r = -0.80$), each of which contributed uniquely to ratings of perceived listening effort. However, only 17 listeners met stringent reliability and agreement criteria.

Conclusions: Ratings of perceived listening effort may provide unique information about the communicative success of individuals with communication disorders. There is great variability, however, among inexperienced listeners' perceptual ratings of electrolaryngeal speech. Future research should investigate variables that may affect perceived listening effort specifically and auditory-perceptual ratings in general.

1. Introduction

A thorough evaluation of an individual with a communication disorder involves the use of different types of instrumental and non-instrumental measures. In particular, auditory-perceptual evaluation is of critical importance for characterizing disordered voice and speech or for measuring treatment success. Assessment of alaryngeal speech is no exception to this approach; common outcomes include ratings of speech acceptability (i.e. "acceptability") and naturalness, which often complement measures of speech intelligibility (Doyle & Eadie, 2005). All of these measures are meant to gauge the impact of the speech disorder on unfamiliar communication partners.

The role of the listener when measuring acceptability or intelligibility has been recognized (Kent, 1996; Kreiman, Gerratt, Kempster, Erman & Burke, 1993), but little attention has been paid to the cognitive-perceptual processes used by listeners as they attempt to decode or judge disordered voice or speech (Evitts & Searl, 2006). These processes may be independent of the phonetic/

☆ Portions of this study were presented at the American Auditory Society's 42nd Annual Scientific & Technology Meeting, March 2015, Scottsdale, AZ.

* Corresponding author at: Department of Speech – Language Pathology, Seton Hall University, South Orange, NJ, USA.
E-mail address: kathleen.nagle@shu.edu (K.F. Nagle).

https://doi.org/10.1016/j.jcomdis.2018.03.003
Received 27 January 2017; Received in revised form 5 March 2018; Accepted 12 March 2018
Available online 13 March 2018
0021-9924/

linguistic content in the speech signal, and are not synonymous with auditory-perceptual outcomes that describe the signal. In other words, the attributes a listener perceives from the speech signal (e.g., voice quality, articulatory precision) may differ from the amount of cognitive effort needed to complete this processing.

The predominant model of cognitive processing posits a limited supply of cognitive resources that must be shared among activities such as perception, attention and memory (Kahneman, 1973). These resources may be divided among concurrent tasks, but sharing them has consequences on performance. For example, to complete a listening task involving degraded speech, or in a noisy environment, the listener must re-allocate cognitive resources to processing the "noise." Lemke and Besser (2016) describe the specific reallocation of resources in challenging listening tasks as "listening effort," which serves as an umbrella term for two related but distinct concepts. *Processing effort,* as it relates to listening, reflects the cognitive or processing load allocated to a listening task and the extra resources required and diverted to perform a challenging listening task (Pichora-Fuller, Kramer, Eckert, Edwards, Hornsby, Humes, 2016). On the other hand, *perceived "listening" effort* (PLE) provides a subjective estimation of how taxing it was to engage in a listening task (Lemke & Besser, 2016). Whereas processing effort is captured relatively directly and contemporaneously with a listening task, PLE ratings must be obtained after the listener has at least finished hearing a stimulus.

The literature of hearing science has a long history of examining how noise and hearing loss can increase the amount of effort needed for listeners to process speech in various contexts. This line of research favors a dual task paradigm, in which processing effort is indexed by changes in either behavioral performance or physiological function. Performance accuracy or reaction time to a secondary challenging task is compared to a simultaneously presented primary speech recognition task (Fisk, Derrick & Schneider, 1986). A decrease in accuracy or an increase in reaction time to a stimulus suggests a reallocation of resources. Likewise, changes in central nervous system activity during effortful processing have been measured using functional magnetic resonance imaging, electroencephalography and event-related potentials (e.g., Zekveld, Heslenfeld, Johnsrude, Versfeld & Kramer, 2014; Bernarding, Strauss, Hanneman, Seidler & Corona-Strauss, 2017; Romero-Rivas, Martin & Costa, 2015, respectively). Effects on the autonomic system have similarly been used to index processing effort during challenging listening tasks. For example, pupillometry is a popular measure of focused attention or concentration derived from autonomic changes in processing load (Winn & Edwards, 2013; Zekveld, Festen & Kramer, 2013). Changes in pupil dilation can be complicated to interpret, but when pupil size is normalized, relative changes in pupil size seem to indicate the degree of cognitive effort required to process speech in noise (Zekveld, Kramer & Festen, 2010).

Subjective, post-listening measures of PLE capture the effects of investing increased processing resources for a given task. In what may be considered both a feature and a bug, PLE is thus affected by the individual characteristics of the listener providing the ratings. Behavioral and physiological processing effort outcomes are frequently reported in the literature without supporting measures of PLE (e.g., Houben, van Doorn-Bierman & Dreschler, 2013), but ratings of PLE may complement measures of processing effort by providing a window into the listener's attitude, motivation and tolerance for listening in adverse conditions.

Investigations of the impact of disordered speech or voice on listening effort have almost exclusively involved post-listening self-ratings of "attention allocation" (Beukelman, Childes, Carrell, Funk, Ball & Pattee, 2011; Beukelman, Gillespie, Fager & Ullman, 2014), "mental effort" (Panico & Healey, 2009) or PLE (Nagle & Eadie, 2012a; Whitehill & Wong, 2006) as primary outcomes. These measures complement other perceptual or acoustic measures of disordered speech stimuli, and may be more relevant to *communicative success* for individuals with speech or voice disorders. That is, the decision to initiate or maintain a conversation with a speaker with a speech or voice disorder may be equally influenced by perception and reality; for some listeners, a perception of increased listening effort may outweigh the reality of expending additional processing resources in a conversation. Because it may affect the ultimate success of a communicative interaction, understanding how PLE differs from traditional perceptual measures of speech, such as intelligibility and acceptability, is important for improving functional speech outcomes.

1.1. PLE & speech intelligibility

The value of PLE as an outcome measure depends on its ability to provide unique information not captured by other measures. For this reason, early investigations of PLE for disordered speech compared these subjective ratings to measures of intelligibility. Intelligibility is usually defined as the accuracy with which a listener is able to decode an utterance (Kent, Weismer, Kent & Rosenbek, 1989). It might be expected that as intelligibility increases, the amount of effort required to listen to it would correspondingly decrease. However, there is a known trade-off between signal quality and intelligibility that drives multidimensional assessment of communication disorders (Brons, Houben & Dreschler, 2013; Brons, Houben & Dreschler, 2014) and studies investigating the association between PLE and the intelligibility of dysarthric speech have found that the relationship is not so predictable (Beukelman et al., 2011; Whitehill & Wong, 2006).

Whitehill and Wong (2006) defined PLE as the "effort needed to understand the speaker" in their investigation of the perceptual factors contributing to listening effort. Twenty inexperienced listeners transcribed decontextualized sentences and rated PLE using an undifferentiated 10 cm visual analog scale (VAS). The authors grouped 33 Cantonese dysarthric speakers by etiology (i.e., Parkinson's disease, cardiovascular accident, cerebral palsy and "other"), representing hypokinetic, spastic, athetoid, and mixed dysarthrias. Apart from intelligibility scores, no additional perceptual measures of severity of disorder were provided. Sentence intelligibility scores and ratings of PLE were strongly negatively correlated for their dysarthric speech samples (Spearman's $R = -0.95$). However, three speakers with high intelligibility (greater than 85%) received relatively high mean ratings of PLE, indicating a distinction between transcribed intelligibility and the perceived effort needed to understand some speech. No particular perceptual features seemed to differentiate the low-effort from the high-effort speakers; in fact, "slurred speech" was among the top three features selected at all levels of perceived listening effort.