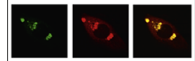


Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Brain Research



Research Report

The perceptual chunking of speech: A demonstration using ERPs

Annie C. Gilbert^{a,b,*}, Victor J. Boucher^a, Boutheina Jemel^{b,c}^aLaboratoire de Sciences Phonétiques, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, QC, Canada H3C 3J7^bLaboratoire de Recherche en Neurosciences et Électrophysiologie Cognitive, Hôpital Rivière-des-Prairies, 7070, boul. Perras, Montréal, QC, Canada H1E 1A4^cCentre de recherche Fernand-Seguin, École d'Orthophonie et d'Audiologie, Université de Montréal, 6128, Succursale Centre-Ville, Montréal, QC, Canada H3C 3J7

ARTICLE INFO

Article history:

Accepted 17 January 2015

Available online 28 January 2015

Keywords:

Chunking

Perceptual chunking

ERP

Positive shifts

Speech processing

Short-term memory

Grouping effects

Domain-general processes

Prosody

ABSTRACT

In tasks involving the learning of verbal or non-verbal sequences, groupings are spontaneously produced. These groupings are generally marked by a lengthening of final elements and have been attributed to a domain-general *perceptual chunking* linked to working memory. Yet, no study has shown how this domain-general chunking applies to speech processing, partly because of the traditional view that chunking involves a conceptual recoding of meaningful verbal items like words (Miller, 1956). The present study provides a demonstration of the perceptual chunking of speech by way of two experiments using evoked Positive Shifts (PSs), which capture on-line neural responses to marks of various groups. We observed listeners' response to utterances (Experiment 1) and meaningless series of syllables (Experiment 2) containing changing intonation and temporal marks, while also examining how these marks affect the recognition of heard items. The results show that, across conditions — and irrespective of the presence of meaningful items — PSs are specifically evoked by groups marked by lengthening. Moreover, this on-line detection of marks corresponds to characteristic grouping effects on listeners' immediate recognition of heard items, which suggests chunking effects linked to working memory. These findings bear out a *perceptual chunking* of speech input in terms of groups marked by lengthening, which constitute the defining marks of a domain-general chunking.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Unlike alphabetic text where spaces divide units like words, speech entails fleeting series of connected sounds. In order to

interpret these changing signals, listeners need to buffer incoming sequences in short-term memory (STM, or “working memory”, Baddeley, 2010). Moreover, since STM is limited, one has to assume that sequential information is somehow segmented into

*Corresponding author. Present address: Neurolinguistics Laboratory, School of Communication Sciences and Disorders, McGill University, 2001, McGill College, 8th floor, Montréal, QC, Canada H3A 1G1.

E-mail addresses: annie.c.gilbert@mcgill.ca (A.C. Gilbert), victor.boucher@umontreal.ca (V.J. Boucher), boutheina.jemel@umontreal.ca (B. Jemel).

chunks that fit this limited store (Kurby and Zacks, 2008). Such segmentation, which would logically operate at a perceptual stage before the interpretation of sequences, is essential to understanding how a domain-general chunking process can apply to speech processing. But this idea departs from the conventional view of chunking, which stands as a central concept of cognitive psychology.

Traditionally, chunking is often conceptualized by reference to the recall of verbal lists. For instance, when recalling series of digits or syllables, even unstructured lists, speakers create groups (e.g., Boucher, 2006; Ryan, 1969a, 1969b; Thrope and Rowland, 1965; Wilkes, 1975). In explaining this behavior, the most often cited work is Miller (1956). According to this author, although STM has a limited capacity, large numbers of items can be learned if they are semantically “recoded” into chunks before being stored in long-term memory (LTM). A classic example is the letter series I B M F B I C I A I R S, which can be learned by recoding items according to the acronyms I B M., F B I., C I A., I R S. This view of chunking continues to prevail not only in studies of verbal memory but also in neurophysiological research on sequence learning (for recent examples, see Boyd et al., 2009; Jin et al., 2014; Tremblay et al., 2009). In fact, many neuroscientists see chunking as a means by which subjects organize sequences so as to overcome limitations in STM. In this light, chunking has been a foundational notion in analyzing the organization of sequential behaviors, from simple motor functions such as reaching to more complex activities like speaking (Gallistel, 1980; Lashley, 1951). However, critics note that Miller’s idea of chunking as a conceptual recoding hardly applies to chunks observed in learning lists of nonsense syllables or motor sequences (e.g. Terrace, 2001). Such observations indeed point to a more *domain-general* principle.

Furthermore, it is worth mentioning another paper by Miller (1962), which suggested a very different chunking concept. In that paper, he argued that listeners are not interpreting speech by making decisions on a sound-by-sound basis. Instead, they are delaying decisions over some “perceptual unit”, which can span a number of elements (Miller, 1962). This presents a view of chunks that does not involve a recoding but instead refers to a basic perceptual segmentation. Such a principle has been described by Terrace (2001; see also Gobet et al., 2001) who emphasized the need to distinguish between a domain-general *input* chunking and *output* chunking involving LTM. According to this view, output chunks reflect a process where learned forms in LTM underlie the recognition of units in sequences. On the other hand, input chunks relate to a perceptual unitization of sequential stimuli that conforms to the capacity limit of STM and the “focus of attention” (Cowan, 2000). Both types of chunking would apply in processing utterances and the problem of separating these processes may explain the lack of research on the perceptual chunking of speech. Yet, there are defining marks that serve to distinguish perceptual chunking from output chunking across behaviors.

On this point, Terrace (2001) notes that the uncritical application of the concept of chunking in various disciplines has led to neglect the fact that perceptual or sensori-motor chunking presents observable marks. These relate to latency variations in produced sequences or what is called

“inter-response times” (IRTs). Typically, delays in IRTs create temporal groups or chunks marked by a lengthening of group-final elements. Such marks can be observed in animal and human behaviors using various methods (e.g., Graybiel, 1998; Terrace et al., 1996; Yin and Knowlton, 2006). For instance, in studies of sequence learning using key-pressing responses, chunks are identified by a lengthened delay between two key presses interrupting a succession of fast key presses (Bo and Seidler, 2009; Kennerley et al., 2004; Sakai et al., 2003; Verwey et al., 2009; Verwey and Eikelboom, 2003; Wymbs et al., 2012). Interestingly, delays in IRTs marking chunks in non-verbal sequences can correspond to a “boundary lengthening”, which appears across spoken languages (Vaissière, 1983). The correspondence can be striking in that this lengthening creates groups in speech much like delays in IRTs create groups associated with a domain-general sensori-motor chunking. To illustrate this, we refer to Fig. 1, which shows the acoustic patterns of sequences of digits recited in three languages.

In producing such lists, chunks usually emerge in terms of groups of two to four items (similar groupings arise in recalling unstructured lists of meaningless syllables; see Boucher, 2006). A look at the acoustic patterns clarifies the grouping marks. In the figure, intonation is shown by fundamental frequency (F0), while the duration of elements (monosyllables) is shown by the lines over the spectrograms. One can see that, while F0 contours vary extensively, a relative lengthening of digits systematically occurs at the end of groups (i.e., digits 1, 2, 3, and 8 are longer in group-final position than within groups). These subtle timing changes, which can be accompanied by brief pauses, create groupings much as varying IRTs create groups in non-verbal sequences. As for the role of these patterns, a body of work has shown that grouping items in verbal lists benefits recall (e.g., Wickelgren, 1964; Broadbent and Broadbent, 1973; Frankish, 1989, 1995; Hitch et al., 1996; Reeves et al., 2000; Maybery et al., 2002; Chen and Cowan, 2005). In this work, observations of optimal effects for groups of three are seen to reflect capacity limits of STM (for a review, see Cowan, 2000). It should be noted, however, that such grouping effects have recently been shown to extend to speech processing. In particular, a study by Gilbert et al. (2014) using EEG showed that temporal groups of three or four items in utterances variably affect the memory trace of heard items as measured by amplitude changes in the N400 (with a better trace and thus smaller N400 appearing for optimal groups of three). In other words, this study provided original evidence that a perceptual chunking of utterances in temporal groups links to immediate memory of heard elements.

One way to demonstrate the perceptual chunking of speech is to monitor listeners’ on-line responses upon hearing utterances and sequences of syllables using Event-Related Potentials (ERPs) so as to determine whether there are specific responses to lengthening marks that constitute defining attributes of input chunks. In considering the possible neural correlates of chunking, it is essential to note that, in research using ERPs, certain components have served to capture chunk-like units in speech processing. Though the components bear varying names, they operate in similar ways. Specifically, research has shown that listening to speech generally elicits negative rising potentials across a number of elements interrupted by Positive Shifts that can be evoked by varying marks (e.g., Bögels et al., 2011, 2013;

Download English Version:

<https://daneshyari.com/en/article/4323852>

Download Persian Version:

<https://daneshyari.com/article/4323852>

[Daneshyari.com](https://daneshyari.com)