NeuroImage 48 (2009) 541–553

Contents lists available at ScienceDirect

NeuroImage

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

Time course and functional neuroanatomy of speech segmentation in adults

Toni Cunillera [a], Estela Càmara [b,c], Juan M. Toro [a], Josep Marco-Pallares [b], Nuria Sebastián-Galles [d], Hector Ortiz [e,f], Jesús Pujol [e,f], Antoni Rodríguez-Fornells [b,g,*]

[a] Department of Basic Psychology, Faculty of Psychology, University of Barcelona, 08035, Barcelona, Spain
[b] Department of Physiology, Faculty of Medicine, Campus de Bellvitge – IDIBELL, University of Barcelona, 08907, Barcelona, Spain
[c] Department of Neuropsychology, Otto-von Guericke University, 39106, Magdeburg, Germany
[d] Brain and Cognition Unit, Universitat Pompeu Fabra, 08018 Barcelona, Spain
[e] Institut d'Alta Tecnologia-PRBB, CRC Corporació Sanitària, Barcelona, Spain
[f] Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Barcelona, Spain
[g] Institució Catalana de Recerca i Estudis Avançats (ICREA), Spain

ARTICLE INFO

Article history:
Received 21 January 2009
Revised 2 June 2009
Accepted 28 June 2009
Available online 4 July 2009

ABSTRACT

The present investigation was devoted to unraveling the time-course and brain regions involved in speech segmentation, which is one of the first processes necessary for learning a new language in adults and infants. A specific brain electrical pattern resembling the N400 language component was identified as an indicator of speech segmentation of candidate words. This N400 trace was clearly elicited after a short exposure to the words of the new language and showed a decrease in amplitude with longer exposure. Two brain regions were observed to be active during this process: the posterior superior temporal gyrus and the superior part of the ventral premotor cortex. We interpret these findings as evidence for the existence of an auditory–motor interface that is responsible for isolating possible candidate words when learning a new language in adults.

Introduction

The present study was focused on understanding one of the initial processes in language learning: speech segmentation. The difficulty of the task lies in the lack of reliable acoustic cues that indicate where a word begins and ends. Thus, unlike the blank spaces that appear between printed words, the spoken acoustic signal could be considered, in many respects, continuous. Listeners, therefore, must parse the speech signal in order to start learning the new language. Notice that this initial process of isolating words is mandatory for subsequent language processes; for example, associating the possible lexical trace with a specific meaning.

To segment the auditory stream into words, the learner could exploit different acoustical cues such as allophonic variation, stress patterns, prosody, or/and distributional cues such as phonotactic regularities and transitional probabilities of syllable combinations (see for a review Jusczyk et al. (1999)). In relation to the distributional cues, several models have proposed the existence of a powerful statistical learning mechanism in order to explain speech segmentation (Brent, 1999). Supporting this idea, several experiments have shown that listeners are able to exploit the distributional properties of the input, regardless of whether it consists of syllables, tones or visual information (Fiser and Aslin, 2001; Saffran et al., 1999), and it likely proceeds in an incidental fashion (Saffran et al., 1997; Toro et al., 2005; Turk-Browne et al., 2005). Thus, statistical learning is understood as a domain-general mechanism that profits from the regularities of the environment to drive learning.

The computational implementation of this hypothesis has been successfully applied by connectionist models (Christiansen et al., 1998; Elman, 1990; for a different account see Brent, 1997). The underlying idea is that word boundaries are in locations where the transitional probabilities between two sounds are low. In other words, word boundaries can be inferred based on the fact that transitional probabilities are higher for word-internal than for word-external pairs of syllables. For instance, in the sentence "look, a balloon" the string "ba/lloon" is more likely to occur together across other sentences than the string "a/ba" as the latter string would not be heard in phrases such as "the balloon", "the red balloon", etc. In fact, it has been demonstrated in infants and adults that computing the transitional probabilities between syllables is sufficient for isolating new words that are embedded in an artificial continuous speech stream when no acoustical cues are available (Saffran et al., 1996a,b). Remarkably, this process is also accomplished by other animal species (Hauser et al., 2001; Toro and Trobalon, 2005), highlighting its generality.

Interestingly, an alternative computational proposal suggests that the formation of syllabic chunks might be the only process required to isolate possible words (Perruchet and Vinter, 1998). Accordingly to the authors, a chunk can be considered that brings together the elements that are at the attentional focus at a particular moment. In subsequent

* Corresponding author. Department of Ciencias Fisiològiques, Faculty of Medicine, Campus de Bellvitge, IDIBELL, Feixa Llarga s/n, 08907. L'Hospitalet (Barcelona), Spain. Fax: +34 934024268.
E-mail address: antoni.rodriguez@icrea.es (A. Rodríguez-Fornells).

1053-8119/$ – see front matter
doi:10.1016/j.neuroimage.2009.06.069

Poldrack, R.A., Wagner, A.D., Prull, M.W., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage 10, 15–35.

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado, M.F., Hauk, O., Shtyrov, Y., 2006. Motor cortex maps articulatory features of speech sounds. Proc. Natl. Acad. Sci. U.S.A. 103, 7865–7870.

Raichle, M.E., Fiez, J.A., Videen, T.O., MacLeod, A.M., Pardo, J.V., Fox, P.T., Petersen, S.E., 1994. Practice-related changes in human brain functional anatomy during nonmotor learning. Cereb. Cortex 4, 8–26.

Rizzolatti, G., Arbib, M.A., 1998. Language within our grasp. Trends Neurosci. 21, 188–194.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat. Neurosci. 2, 1131–1136.

Rugg, M.D., Coles, M.G.H., 1995. Electrophysiology of Mind. Event-Related Brain Potentials and Cognition. Oxford University Press, Oxford.

Saffran, J.R., Aslin, R.N., Newport, E.L., 1996a. Statistical learning by 8-month-old infants. Science 274, 1926–1928.

Saffran, J.R., Newport, E.L., Aslin, R.N., 1996b. Word segmentation: the role of distributional cues. J. Mem. Lang. 35, 606–621.

Saffran, J.R., Newport, E.L., Aslin, R.N., Tunick, R.A., Barrueco, S., 1997. Incidental language learning: listening (and learning) out of the corner of your ear. Psychol. Sci. 8, 101–105.

Saffran, J.R., Johnson, E.K., Aslin, R.N., Newport, E.L., 1999. Statistical learning of tone sequences by human infants and adults. Cognition 70, 27–52.

Sanders, L.D., Neville, H.J., 2003a. An ERP study of continuous speech processing I. Segmentation, semantics, and syntax in native speakers. Cogn. Brain Res. 15, 228–240.

Sanders, L.D., Neville, H.J., 2003b. An ERP study of continuous speech processing II. Segmentation, semantics, and syntax in non-native speakers. Cogn. Brain Res. 15, 214–227.

Sanders, L.D., Newport, E.L., Neville, H.J., 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat. Neurosci. 5, 700–703.

Schubotz, R.I., von Cramon, D.Y., Lohmann, G., 2003. Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. Neuroimage 20, 173–185.

Scott, S.K., Wise, R.J.S., 2004. The functional neuroanatomy of prelexical processing in speech perception. Cognition 92, 13–45.

Scott, S.K., Blank, C.C., Rosen, S., Wise, R.J., 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123 (Pt 12), 2400–2406.

Scott, S.K., Rosen, S., Lang, H., Wise, R.J.S., 2006. Neural correlates of intelligibility in speech investigated with noise vocoded speech — a positron emission tomography study. J. Acoust. Soc. Am. 120, 1075–1083.

Smith, E.E., Jonides, J., 1998. Neuroimaging analyses of human working memory. Proc. Natl. Acad. Sci. U. S. A. 95, 12061–12068.

Studdert-Kennedy, M., 1987. The phoneme as a perceptuomotor structure. In: Allport, D., MacKay, D., Prinz, W., Scheerer, E. (Eds.), Language Perception and Production. Academic Press, London, pp. 67–84.

Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., Weniger, D., 2002. Neural correlates for the acquisition of natural language syntax. NeuroImage 17, 700–709.

Toro, J.M., Trobalon, J.B., 2005. Statistical computations over a speech stream in a rodent. Percept. Psychophys. 67, 867–875.

Toro, J.M., Sinnett, S., Soto-Faraco, S., 2005. Speech segmentation by statistical learning depends on attention. Cognition 97, B25–B34.

Turk-Browne, N.B., Junge, J.A., Scholl, B.J., 2005. The automaticity of visual statistical learning. J. Exp. Psychol.-Gen. 134, 552–564.

van Petten, C., Luka, B.J., 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain Lang. 97, 279–293.

Warren, J.E., Wise, R.J.S., Warren, J.D., 2005. Sounds do-able: auditory–motor transformations and the posterior temporal plane. Trends Neurosci. 28, 636–643.

Watkins, K.E., Strafella, A.P., Paus, T., 2003. Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia 41, 989–994.

Wilson, S.M., Iacoboni, M., 2006. Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. Neuroimage 33, 316–325.

Wilson, S.M., Saygin, A.P., Sereno, M.I., Iacoboni, M., 2004. Listening to speech activates motor areas involved in speech production. Nat. Neurosci. 7, 701–702.

Wise, R.J.S., Scott, S.K., Blank, S.C., Mummery, C.J., Murphy, K., Warburton, E.A., 2001. Separate neural subsystems within 'Wernicke's area'. Brain 124, 83–95.

Wong, P.C., Perrachione, T.K., Parrish, T.B., 2007. Neural characteristics of successful and less successful speech and word learning in adults. Hum. Brain Mapp. 28, 995–1006.

Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited — again. Neuroimage 2, 173–181.

Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256, 846–849.

Zatorre, R.J., Meyer, E., Gjedde, A., Evans, A.C., 1996. PET studies of phonetic processing of speech: review, replication, and reanalysis. Cereb. Cortex 6, 21–30.