



# LUND UNIVERSITY

## A dual complexity gradient theory of speech perception

Roll, Mikael

*Published in:*  
Proceedings of Fonetik 2018

2018

*Document Version:*  
Publisher's PDF, also known as Version of record

[Link to publication](#)

*Citation for published version (APA):*

Roll, M. (2018). A dual complexity gradient theory of speech perception. In Å. Abelin, & Y. Nagano- Madsen (Eds.), *Proceedings of Fonetik 2018* (pp. 55-59). University of Gothenburg.  
[https://flov.gu.se/digitalAssets/1696/1696270\\_proceedingsfonetik2018.pdf#page=61](https://flov.gu.se/digitalAssets/1696/1696270_proceedingsfonetik2018.pdf#page=61)

*Total number of authors:*  
1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# A Dual Complexity Gradient Theory of Speech Perception

Mikael Roll

Center for Languages and Literature, Lund University

## Abstract

*This article proposes a Dual Complexity Gradient theory of speech perception in the brain. The theory unifies a previously proposed phonetic-to-phonological complexity gradient along the ventral auditory processing stream in the temporal lobe with a recently suggested cortical structure complexity gradient. Findings supporting the theory are discussed as well as its predictions.*

## Introduction

To improve methods of language teaching and rehabilitation in a society with increasing linguistic diversity and a growing elderly population, we need fundamental information about the way native speakers process language. However, to date, the very fundamental process of how the brain integrates acoustic information into phonetic features, phonemes, syllables, and words is still not completely understood. We propose a theoretical framework, the Dual Complexity Gradient theory of speech perception, which can be used to generate new hypotheses about the relation between brain structure and phonetic processing. The Dual Complexity Gradient theory unifies a previously proposed phonetic-to-phonological complexity gradient along the ventral auditory perception stream with a recently suggested cortical structure gradient for phonetic-to-phonological processing. As preliminary evidence, it takes recent findings of correlation between native phonological proficiency and cortical thickness in different brain areas.

## Auditory perception streams

Speech processing in the brain proceeds in two different streams (Saur, Kreher, Schnell, Kümmerer, Kellmeyer, Vry, Umarova, Musso, Glauche, Abel, Huber, Rijntjes, Hennig, & Weiller, 2008) most strongly represented in the left hemisphere (Peelle, 2012). Both streams begin in primary auditory cortex. The *dorsal stream* then extends to the superior temporal gyrus and through the parietal lobe via the superior longitudinal fasciculus to frontal cortex. The connection between auditory regions in temporal cortex and motor cortices in the frontal lobe is crucial for language learning, subvocal rehearsal

(Buchsbaum, Olsen, Koch, & Berman, 2005), effortful processing, and predictive processes (Rauschecker & Scott, 2009; Roll, Söderström, Frid, Mannfolk, & Horne, 2017) possibly involving emulation (Grush, 2004). A parallel pathway through the arcuate fasciculus has been argued to be more involved in combinatorial syntactic parsing (Friederici, Chomsky, Berwick, Moro, & Bolhuis, 2017). The primary functions of the dorsal stream are sound localization and auditory sensorimotor integration (Rauschecker & Scott, 2009).

The *ventral stream* extends laterally and frontally from Heschl's gyrus to the anterior superior temporal gyrus and even the anterior superior temporal sulcus. From the anterior temporal lobe this stream continues to the ventral part of the frontal lobe through the extreme capsule. The basic function of the ventral stream is auditory object identification based on increasingly complex analysis of acoustic features (Rauschecker & Scott, 2009).

## Spatial complexity gradient

A meta-analysis involving over 100 imaging experiments found a spatial phonetic complexity gradient along the ventral stream (DeWitt & Rauschecker, 2012). Thus, the further away one travels from primary auditory cortex along the ventral stream, the more complex the phonetic-to-phonological representations become, involving phonetic features, phonemes, syllables, words, and even short phrases. Primary auditory cortex has a rather detailed representation of the spectrotemporal characteristics of sounds. It is even tonotopically organized, meaning that different center frequencies map to different locations of the cortex (Humphries, Liebenthal, & Binder, 2010; Merzenich & Brugge, 1973).

Secondary auditory cortex (Fig. 1) has neurons that are sensitive to specific constellations of the spectrotemporal characteristics detected in primary cortex. These correspond to distinctions in manner of articulation—further organized into larger groups of sonorants and obstruents—and place of articulation. Secondary cortex also combines formants, giving rise to perception of vowel distinctions (Mesgarani, Cheung, Johnson, & Chang, 2014). Secondary auditory cortex further has normalized neural representations of tone and intonation patterns independent of absolute F0 level (Tang, Hamilton, & Chang, 2017). Tertiary auditory cortex, anteriorly located in the superior temporal gyrus, is sensitive to categorical auditory object perception in both humans (Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002) and rhesus monkeys (Tsunada, Lee, & Cohen, 2011). For speech, this region seems sensitive to recognition of units at the word level. Slightly ventrally, in the frontal part of the superior temporal sulcus, even shorter phrases seem to be processed.

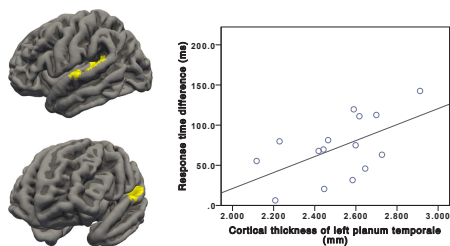


Figure 1. Cortical thickness of left secondary auditory cortex (planum temporale) in superior temporal gyrus correlates with proficiency in native language word accent processing (from Schremm, Novén, Horne, Söderström, Westen, & Roll, 2018).

## Structural complexity gradient

The other part of the Dual Complexity Gradient theory is a *structural complexity gradient*, a generalization emanating from recent findings about the relation between brain structure and behavior. The structural gradient assumes that complex cognitive functions are aided by thicker cortex in brain areas modulating associative processing, whereas effective pruning and cortical myelination help lower-level cognitive functions, yielding instead an advantage of thinner cortex in primary processing areas (Novén, Schremm, Nilsson, Horne, & Roll, submitted).

Previously, gray-matter volume was the most commonly used structural brain metric. How-

ever, cortical volume indistinctly hides two measures that recent research suggests might be rather independently associated with cognitive abilities: cortical thickness and cortical surface area (Vuoksimaa, Panizzon, Chen, Fiecas, Eyler, Fennema-Notestine, Hagler, Franz, Jak, Lyons, Neale, Rinker, Thompson, Tsuang, Dale, & Kremen, 2016). Cortical surface area is strongly genetically defined (Vuoksimaa, Panizzon, Chen, Fiecas, Eyler, Fennema-Notestine, Hagler, Fischl, Franz, Jak, Lyons, Neale, Rinker, Thompson, Tsuang, Dale, & Kremen, 2015). Areas early in the auditory stream, primary and secondary auditory cortex, have greater surface area in the left hemisphere than in the right. Left primary auditory cortex is also thinner (Meyer, Liem, Hirsinger, Jäncke, & Hänggi, 2014). The reason is thought to be that a larger surface of well-organized cortical columns with more myelinated axons increases processing speed but reduces cortical thickness. This would be advantageous for the kind of rapid categorization involved in phonetic feature processing in left primary and possibly secondary auditory cortices (Long, Wan, Roberts, & Corfas, 2018; Warrior, Wong, Penhune, Zatorre, Parrish, Abrams, & Kraus, 2009). In line with this idea but outside the language domain, amusic individuals have been found to have thicker cortex than control persons in right primary auditory areas and inferior frontal gyrus (Hyde, Lerch, Zatorre, Griffiths, Evans, & Peretz, 2007). In the same vein, cortical thickness in right inferior frontal gyrus has been found to correlate negatively with pitch discrimination proficiency, which can be argued to be a low-level acoustic task (Novén et al., submitted).

Cortical thickness increases in response to mental training (Román, Lewis, Chen, Karama, Burgaleta, Martínez, Lepage, Jaeggi, Evans, Kremen, & Colom, 2016) and is negatively associated with aging (Thambisetty, Wan, Carass, An, Prince, & Resnick, 2010) and cognitive decline in degenerative diseases (Gerrits, van Loenhoud, van den Berg, Berendse, Foncke, Klein, Stoffers, van der Werf, & van den Heuvel, 2016). Although maturation in children in general implies cortical thinning (Porter, Collins, Muetzel, Lim, & Luciana, 2011), regions in higher-level areas in the left ventral speech processing stream rather increase in thickness during childhood (Sowell, Thompson, Leonard, Welcome, Kan, & Toga, 2004).

Secondary auditory cortex has shown seemingly paradoxical results for cortical thickness. Thus, a low-level auditory experiment found

increased electrophysiological response for *thinner* cortex (Liem, Zaehle, Burkhard, Jäncke, & Meyer, 2012). For more complex tone-suffix association, however, processing speed was seen to augment with *thicker* cortex (Schremm et al., 2018) (Fig. 1). Further, intense language training increases cortical thickness in the anterior portion of secondary auditory cortex (Mårtensson, Eriksson, Bodammer, Lindgren, Johansson, Nyberg, & Lövdén, 2012). In line with the proposed structural gradient, this likely indexes acquisition of a new phonology with novel combinations of phonetic features. The structural complexity gradient provides an explanation for the apparent paradox of cortical thickness advantage or disadvantage if secondary auditory cortex is an intermediate area between primary auditory cortex and higher-level cognition areas, as its myelination patterns seem to indicate (Glasser & Van Essen, 2011). Thus, as mentioned above, increased cortical myelination reduces cortical thickness and is related to faster processing of low-level features. However, higher-level processing requires increased association between information types, and is expected to be related to increased cortical thickness. Accordingly, primary cortices are highly myelinated whereas secondary cortices show an intermediate degree of myelination, and association areas, a low degree of myelination (Glasser, Coalson, Robinson, Hacker, Harwell, Yacoub, Ugurbil, Andersson, Beckmann, Jenkinson, Smith, & Van Essen, 2016).

Anterior to secondary auditory cortex, tertiary auditory cortex is thicker on the left side (Meyer et al., 2014). This is in accordance with accumulation of knowledge regarding high-level phonological patterns during language acquisition resulting in increased number of neurons, synapses, and/or glial cells (Zatorre, Fields, & Johansen-Berg, 2012). Thickness of tertiary auditory cortex is positively associated with speed of processing word accents in real words (Novén, Schremm, van Westen, Horne, & Roll, in preparation), possibly further indicating improved whole form storage in line with Schremm et al. (2018). When listeners were forced to use combinatorial processing to access tone-suffix associations in pseudowords, cortical thickness of Broca's area in the left inferior frontal gyrus rather correlated with speed of access (Fig. 2).

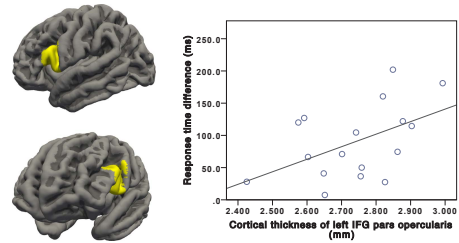


Figure 2. Cortical thickness of pars opercularis of left inferior frontal gyrus (Broca's area) with proficiency in native language word accent processing when combinatorial processing is forced by using pseudowords (from Schremm et al., 2018).

## Native speaker proficiency

Previously brain structure correlates of phonetic or phonological proficiency had mostly been measured along the dorsal stream (Golestani, 2012). This is not strange since proficiency measures are mostly relevant for language learners. As mentioned above, the sensorimotor integration functions of the dorsal stream are crucial for language learning. However, recently, the speed of associating word accent tones with word endings in native speakers of Swedish begun to be measured (Roll, Söderström, & Horne, 2013; Roll, Söderström, Mannfolk, Shtyrov, Johansson, van Westen, & Horne, 2015; Söderström, Horne, Mannfolk, Westen, & Roll, 2017; Söderström, Horne, & Roll, 2017; Söderström, Roll, & Horne, 2012). Since this phonological association seems to be important for online prediction and facilitation in speech processing (Söderström, Horne, Frid, & Roll, 2016), dominating it can be seen as an indication of increased language proficiency, even in native speakers. This has made it possible to assess the relation between cortical thickness and a tentative measure of native phonological “proficiency” at different levels, giving rise to correlations along the ventral speech processing stream (Schremm et al., 2018).

## Conclusions

This article has reviewed some works showing evidence for a previously suggested phonetic-to-phonological *spatial complexity gradient* in the ventral stream of auditory processing. This gradient proceeds in anterior direction from primary auditory cortex through secondary and tertiary cortices in the superior temporal gyrus and anterior superior temporal sulcus. The article has also taken up a recently suggested *structural*

*complexity gradient*, by which processing of low-level acoustic features is facilitated by a thinner and more myelinated cortex, where straightforward choices can rapidly be processed. Higher-level phonological processing is rather aided by a higher number of associations in a more complex network yielding thicker cortex. Further evidence is needed to corroborate the structural complexity gradient along the ventral speech perception stream, but unification of the two complexity gradients into one framework, a Dual Complexity Gradient theory, gives a number of testable predictions. Thus, proficiency at different levels of processing in one's native language should correlate with cortical thickness in different ways and in different brain areas. For lower-level processing and in primary areas, a negative correlation would be expected. This has been found to a certain degree in the right hemisphere for proficiency in non-speech pitch discrimination. Proficiency in higher-level processing would be thought to correlate positively with cortical thickness. This is what has been found for proficiency in word accent processing in secondary auditory cortex and, when involving forced combinatorial processing in pseudowords, in inferior frontal gyrus. Many points along the ventral pathway and levels of phonetic-to-phonological complexity need to be tested for the theory to be considered to be corroborated.

## Acknowledgments

This work was supported by Knut and Alice Wallenberg Foundation (grant number 2014.0139) and Marcus and Amalia Wallenberg Foundation (grant number 2014.0039). I am grateful to Merle Horne, Mikael Novén, Andrea Schremm, Pelle Söderström and Sabine Gosselke Berthelsen for insightful discussion.

## References

- Buchsbaum BR, Olsen RK, Koch P, & Berman KF (2005). Human dorsal and ventral auditory streams subservise rehearsal-based and echoic processes during verbal working memory. *Neuron*, 48(4): 687-697.
- DeWitt I, & Rauschecker JP (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, 109(8): E505-E514.
- Friederici AD, Chomsky N, Berwick RC, Moro A, & Bolhuis JJ (2017). Language, mind and brain. *Nature Human Behaviour*, 1(10): 713-722.
- Gerrits NJHM, van Loenhou AC, van den Berg SF, Berendse HW, Foncke EMJ, Klein M, Stoffers D, van der Werf YD, & van den Heuvel OA (2016). Cortical thickness, surface area and subcortical volume differentially contribute to cognitive heterogeneity in Parkinson's disease. *PLoS ONE*, 11(2): 1-14.
- Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Uğurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, & Van Essen DC (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536: 171-181.
- Glasser MF, & Van Essen DC (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *Journal of Neuroscience*, 31(32): 11597-11616.
- Golestani N (2012). Brain structural correlates of individual differences at low-to high-levels of the language processing hierarchy: A review of new approaches to imaging research. *International Journal of Bilingualism*, 18(1): 6-34.
- Grush R (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27(3): 377-396.
- Humphries C, Liebenthal E, & Binder JR (2010). Tonotopic organization of human auditory cortex. *Neuroimage*, 50(3): 1202-1211.
- Hyde KL, Lerch JP, Zatorre RJ, Griffiths TD, Evans AC, & Peretz I (2007). Cortical Thickness in Congenital Amusia: When Less Is Better Than More. *Journal of Neuroscience*, 27(47): 13028.
- Liem F, Zaehle T, Burkhard A, Jäncke L, & Meyer M (2012). Cortical thickness of supratemporal plane predicts auditory N1 amplitude. *Neuroreport*, 23: 1026-1030.
- Long P, Wan G, Roberts MT, & Corfas G (2018). Myelin development, plasticity, and pathology in the auditory system. *Developmental Neurobiology*, 78(2): 80-92.
- Merzenich MM, & Brugge JF (1973). Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Research*, 50(2): 275-296.
- Mesgarani N, Cheung C, Johnson K, & Chang EF (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, 343: 1006-1010.
- Meyer M, Liem F, Hirsiger S, Jäncke L, & Hänggi J (2014). Cortical surface area and cortical thickness demonstrate differential structural asymmetry in auditory-related areas of the human cortex. *Cerebral Cortex*, 24(10): 2541-2552.
- Mårtensson J, Eriksson J, Bodammer NC, Lindgren M, Johansson M, Nyberg L, & Lövdén M (2012). Growth of language-related brain areas after foreign language learning. *Neuroimage*, 63: 240-244.
- Novén M, Schremm A, Nilsson M, Horne M, & Roll M (submitted). Cortical thickness of Broca's area and right homologue predict grammar learning aptitude and pitch discrimination proficiency.
- Novén M, Schremm A, van Westen D, Horne M, & Roll M (in preparation). Cortical thickness of planum polare in native tone perception.
- Patterson RD, Uppenkamp S, Johnsrude IS, & Griffiths TD (2002). The Processing of Temporal Pitch and Melody Information in Auditory Cortex. *Neuron*, 36(4): 767-776.
- Peelle JE (2012). The hemispheric lateralization of speech processing depends on what "speech" is: a hierarchical perspective. *Frontiers in Human Neuroscience*, 6: 309.
- Porter JN, Collins PF, Muetzel RL, Lim KO, & Luciana M (2011). Associations between cortical thickness and verbal fluency in childhood, adolescence,



- and young adulthood. *Neuroimage*, 55(4): 1865-1877.
- Rauschecker JP, & Scott SK (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6): 718-724.
- Roll M, Söderström P, Frid J, Mannfolk P, & Horne M (2017). Forehearing words: Pre-activation of word endings at word onset. *Neuroscience Letters*, 658: 57-61.
- Roll M, Söderström P, & Horne M (2013). Word-stem tones cue suffixes in the brain. *Brain Research*, 1520: 116-120.
- Roll M, Söderström P, Mannfolk P, Shtyrov Y, Johansson M, van Westen D, & Horne M (2015). Word tones cueing morphosyntactic structure: neuroanatomical substrates and activation time course assessed by EEG-fMRI. *Brain and Language*, 150: 14-21.
- Román FJ, Lewis LB, Chen C-H, Karama S, Burgaleta M, Martínez K, Lepage C, Jaeggi SM, Evans AC, Kremen WS, & Colom R (2016). Gray matter responsiveness to adaptive working memory training: a surface-based morphometry study. *Brain Structure and Function*, 221(9): 4369-4382.
- Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry M-S, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, & Weiller C (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, 105(46): 18035-18040.
- Schremm A, Novén M, Horne M, Söderström P, Westen Dv, & Roll M (2018). Cortical thickness of planum temporale and pars opercularis in native language tone processing. *Brain and Language*, 176: 42-47.
- Sowell ER, Thompson PM, Leonard CM, Welcome SE, Kan E, & Toga AW (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, 24(38): 8223.
- Söderström P, Horne M, Frid J, & Roll M (2016). Pre-activation negativity (PrAN) in brain potentials to unfolding words. *Frontiers in Human Neuroscience*, 10: 1-11.
- Söderström P, Horne M, Mannfolk P, Westen Dv, & Roll M (2017). Tone-grammar association within words: Concurrent ERP and fMRI show rapid neural pre-activation and involvement of left inferior frontal gyrus in pseudoword processing. *Brain and Language*, 174: 119-126.
- Söderström P, Horne M, & Roll M (2017). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*, 46: 271-280.
- Söderström P, Roll M, & Horne M (2012). Processing morphologically conditioned word accents. *The Mental Lexicon*, 7(1): 77-89.
- Tang C, Hamilton LS, & Chang EF (2017). Intonational speech prosody encoding in the human auditory cortex. *Science*, 357: 797-801.
- Thambisetty M, Wan J, Carass A, An Y, Prince JL, & Resnick SM (2010). Longitudinal changes in cortical thickness associated with normal aging. *Neuroimage*, 52(4): 1215-1223.
- Tsunada J, Lee JH, & Cohen YE (2011). Representation of speech categories in the primate auditory cortex. *Journal of Neurophysiology*, 105(6): 2634-2646.
- Vuoksima E, Panizzon MS, Chen C-H, Fiecas M, Eyer LT, Fennema-Notestine C, Hagler DJ, Fischl B, Franz CE, Jak A, Lyons MJ, Neale MC, Rinker DA, Thompson WK, Tsuang MT, Dale AM, & Kremen WS (2015). The Genetic Association Between Neocortical Volume and General Cognitive Ability Is Driven by Global Surface Area Rather Than Thickness. *Cerebral Cortex*, 25(8): 2127-2137.
- Vuoksima E, Panizzon MS, Chen C-H, Fiecas M, Eyer LT, Fennema-Notestine C, Hagler DJ, Franz CE, Jak AJ, Lyons MJ, Neale MC, Rinker DA, Thompson WK, Tsuang MT, Dale AM, & Kremen WS (2016). Is bigger always better? The importance of cortical configuration with respect to cognitive ability. *Neuroimage*, 129: 356-366.
- Warrier C, Wong P, Penhune V, Zatorre R, Parrish T, Abrams D, & Kraus N (2009). Relating structure to function: Heschl's Gyrus and acoustic processing. *Journal of Neuroscience*, 29(1): 61-69.
- Zatorre RJ, Fields RD, & Johansen-Berg H (2012). Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nature Neuroscience*, 15: 528-536.