BEHAVIOR REFLECTS THE (DEGREE OF) REALITY OF PHONOLOGICAL FEATURES IN THE BRAIN AS WELL

Holger Mitterer

Max Planck Institute for Psycholinguistics
holger.mitterer@mpi.nl

ABSTRACT

To assess the reality of phonological features in language processing (vs. language description), one needs to specify the distinctive claims of distinctive-feature theory. Two of the more far-reaching claims are compositionality and generalizability. I will argue that there is some evidence for the first and evidence against the second claim from a recent behavioral paradigm.

Highlighting the contribution of a behavioral paradigm also counterpoints the use of brain measures as the *only* way to elucidate what is "real for the brain".

The contributions of the speakers exemplify how brain measures can help us to understand the reality of phonological features in language processing. The evidence is, however, not convincing for a) the claim for underspecification of phonological features—which has to deal with counterevidence from behavioral as well as brain measures—, and b) the claim of position independence of phonological features.

1. WHAT IS SPECIAL ABOUT PHONOLOGICAL FEATURES?

Listeners respond categorically differently to sentences such as "hand me your gun" and "hand me your gum". Any model of speech processing must be able to explain this fact. Episodic models without a role for phonological features [9] differ from distinctive-feature theory most prominently with regard to the assumptions of compositionality and generalizability. According to distinctive-feature theory, the difference between *gun* and *gum* is a difference in place of articulation and hence the same as the following differences:

- *run* and *rum* (generalizing over words)
- *nail* and *mail* (generalizing over positions)
- *coat* and *cope* (generalizing over manner)

This last point also highlights the perhaps most far-reaching claim of distinctive-feature theory, the compositionality of distinctive features, which allows one to explain the phonological system of a language with just a few features.

A middle ground is a segmental (phonemic, demi-syllabic, moraic, etc.) model, in which the difference between /d/ and /b/ is independent from the difference between /n/ and /m/.

2. EVIDENCE FOR AND AGAINST FEATURES

At one point in time, it seemed that the "selective adaptation paradigm" [4] might provide the key to show the psychological reality of phonological features. Subsequent research, however, showed that selective adaptation can occur for any kind of feature, such as auditory or "ad-hoc" task-dependent features, not just phonological ones. This decreased the attractiveness of the paradigm to show the reality of distinctive features in language processing [28].

A new paradigm may provide more insights whether phonological features have a psychological reality. Norris, McQueen, and Cutler [22] showed that listeners use lexical knowledge to retune their perceptual categories, and by inference, their distinctive features. Importantly, they showed that this effect is not easily explained by low-level adaptation or acoustic contrast.

Crucially, the same authors showed that learning generalizes over words [18], providing evidence for the generalizability of distinctive features. Moreover, Kraljic and Samuel [15] showed that learning about voicing in alveolar stops generalizes to labial stops. This seems to suggest some psychological reality of the compositionality of a phonological voicing feature.

Not all investigations, however, paint such a favorable picture for the psychological reality of phonological features. Jesse and McQueen found that generalization over positions may be limited [13], which is problematic for the claim of position-independence of phonological features.

3. BRAIN?

The special session asks whether phonological features have any reality for the brain. I have, however, not yet talked about "brain data". From a psycholinguistic point of view, the difference is not so much between "brain" and "behavior", but rather between language description and language processing.

This point of view does not arise from a rejection of the claim that studying the brain is useful for cognitive science, but rather from the wholehearted endorsement of it. If we agree that the brain is crucial to guide behavior, we must also agree that behavior in well-designed experiments reflects the workings of the brain. Anything else would be Cartesian dualism, in which behavior reflects the working of the mind, and measures from EEG (electro-encephalogram), MEG (magnetoencephalogram), or fMRI (functional magnetic resonance imaging) reflect the working of the brain.

A more implicit argument that may lie at the heart of the debate is that brain measures may more directly reflect the workings of the brain than behavioral measures. This may be true in an obvious sense; however, the point is that a measure must be able to separate conflicting theories. Especially the blood flow measure in fMRI is far removed from actually revealing what kind of algorithms and processing are *really* going on in the brain. It remains a matter of debate to what extent fMRI will actually be able to separate conflicting theories in cognitive sciences [25].

I would argue that there is no general rule for which methods will be most informative. Instead, the best measure to separate conflicting theories will be different from case to case. Any measure of brain processing, from single-cell recordings to behavioral experiments, requires a chain of inference from the empirical observation to its putative cause, an algorithm implemented in the brain. From a psychological point of view, it is remains important not to be bedazzled by the technical (and financial) sophistication of brain measures when evaluating how reasonable this inference is.

4. BRAIN MEASURES AND PHONOLOGICAL FEATURES

The target papers show how brain measures can be used to investigate the workings of phonological features. At the very least, they show the reality of non-linear auditory-phonetic relations already in auditory cortex. However, there are no strong tests yet that would back up the strong claims of generalization or underspecification of phonological features.

4.1. Feature categorization in auditory cortex

The experiments by Phillips and colleagues [26] as well as Kazanina and colleagues [14] (see the contribution of Isardi [12]) clearly show a non-linear translation from acoustics to phonetics that is already achieved in auditory cortex. The high spatial resolution of MEG gives this data an additional appeal. Specifically, it makes claims [27] that "mirror neurons" or the "motor cortex" are necessarily involved in speech perception rather unlikely. This fits well with the data on the neurobiology of speech perception by Sophie Scott and colleagues [30], which also highlights the contributions of auditory and other temporal areas.

4.1.1. What kind of features?

The debate on auditory vs. (pre-)motor cortex contributions to speech perception leads to the discussion about the nature of features: acoustics or gestures (see also [12])? Quite often, this is difficult to establish because of the obvious acoustic-gesture correlation. Finding evidence that auditory cortex encodes higher-order acoustic invariants (F2-F1 [24], or formant ratios [12]) is a huge step forward. Paradoxically, this can be interpreted as evidence for the direct-perception gestural account [6]—higher-order invariants are after all at the heart of direct realism—as well as auditory theories [3].

A possible way forward is to test more extreme examples of auditory-gestural non-linearity, such as the case of rhotics. Take for instance the Berber language, which distinguishes alveolar and uvular trills. Will the alveolar trill form a natural class with the articulatory similar alveolar stop, or with the acoustically similar uvular trill? Mismatch negativity designs along the lines the "natural class" experiment proposed by Isardi [12] can be extremely useful here, as well as "phoneme localization" studies along the line of [23].

4.1.2. Features or segments or allophones?

The data on the early categorization of voicing in the auditory cortex is in line with the assumption that phonological features are "real for the brain". However, the data also fit a segmental model, which makes a categorical distinction between /d/ and /t/. Isardi [12] proposes the crucial experiment in which standards [b d g] are contrasted with deviants [p t k].

However, in this case, there is still a strong acoustic constancy of the distinctive feature. The English language presents an even more telling test case due to the position-dependent cues to stop voicing. Phonological voicing is most reliably coded by VOT in onset position, but by vowel duration in coda position. Proponents of phonological features should live dangerously and try to show a common coding of these distinct acoustic cues for the same feature in auditory cortex. While I remain skeptical about this aspect of features, this kind of evidence would be very convincing.

4.2. No features ?

Eulitz [5] claims that MMN data show the underspecification of phonological features in the brain. The data seem solid: The strong prediction of an asymmetric MMN—when the roles of standard and deviant are reversed—is confirmed. If the standard has a putatively underspecified feature, the deviant only triggers an auditory MMN. In the reverse case, the standard is phonologically specified, and an additional phonological MMN occurs. In line with this hypothesis, the MMN is larger in cases where the standard contains a specified feature.

However, at least two data sets exist that do not find that outcome. I found a symmetric MMN with nasals [19]: A labial deviant with a putatively underspecified alveolar standard elicited the same MMN as a the alveolar deviant with the specified labial standard. The same results were obtained by Bonte et al. [1] with fricatives. Two things are noteworthy here: First of all, finding a symmetric MMN may seem duller than finding the opposite, which makes positive results more likely to be published than negative results. The empirical basis for the claim of underspecification may hence be overestimated due to the 'file-drawer problem' [29]. Secondly, the data by Bonte et al. provide an alternative explanation for the positive results. Prototypical deviants give rise to stronger MMNs than less prototypical ones. Because high-frequency phonemes, such as alveolars, are also assumed to be underspecified, the evidence assumed to show underspecification may just reflect a difference between high- and low-frequency phonemes.

The claim of underspecification has also lost its theoretical appeal as a theory of word recognition that allows pronunciation variants to be recognized [16]. With the very same material that failed to show an effect of underspecfciation, I [20, 21] showed that there are perceptual mechanisms that allow pronunciation variants to be recognized with full specification, converging with findings from at least three other, independent, labs [2, 8, 10].

4.3. Audiovisual specification of features

Hertrich and Ackermann [11] investigate audio-visual fusion. Their sophisticated analysis techniques are well-suited to speak to the long debate about early-interaction [6] vs. late-integration [17] accounts of audiovisual speech perception and they provide clear data: Audiovisual percepts arise by late integration. Early effects are purely attention modulation and do not depend on the phonological features present in the visual stream.

They also provide a phonological analysis for the context of audiovisual fusion in which they allude to the special status of the underspecified alveolar place of articulation. An alternative is a simple Bayesian inference model, in which the likelihood of labial event given no visible closure is close to zero— overriding the acoustic evidence for a labial. The phonological analysis is problematic for the reverse case: With a special status for coronals, the same percept should arise for a visual labial and a acoustic dorsal. In contrast, a Bayesian inference model correctly predicts that, because of the high likelihood of a labial given a visual closure, the percept must contain a labial; as the most frequent percept /bga/ does.

5. SUMMARY

In this discussion contribution, I have argued that the reality of phonological features for the brain should be investigated by both behavioral and brain measures. The target articles provided good examples of theory-driven use of MEG/EEG, in which cognitive models are used to make predictions about the data patterns.

Nevertheless, I remain skeptical about the underspecification of phonological features and other strong claims of the phonological-feature model. The theory of underspecified recognition has lost its appeal as a theory of language processing, because alternative models exist to explain the recognition of pronunciation variation [7, 20]. The claim that a single feature such as [±voice] is used to distinguish stops in onset and coda position seems unlikely, given the diverse acoustic cues. However, MMN designs are well suited to test this claim and either support or reject it.

6. REFERENCES

[1] Bonte, M., Mitterer, H., Zellagui, N., Poelmans, N., and Blomert, L. "Auditory cortical tuning to statistical regularities in phonology", Clinical Neurophysiology, 116: 2765, 2005.

[2] Darcy, I., Peperkamp, S., and Dupoux, E., "Bilinguals play by the rules: perceptual compensation for assimilation in late L2-learners". In: Cole J, Hualde J (eds) "Bilinguals play by the rules: perceptual compensation for assimilation in late L2-learners", p., in press.

[3] Diehl, R., Lotto, A. J., and Holt, L. L. "Speech perception", Annual Review of Psychology, 55: 149-179, 2004.

[4] Eimas, P. D. and Corbit, J. D. "Selective adaptation of linguistic feature detectors", Perception & Psychophysics, 4: 99-109, 1973.

[5] Eulitz, C. "Representation of phonological features in the brain: evidence from mismatch negativity" in Proceedings of Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 2007

[6] Fowler, C. A. "Listeners do hear sounds, not tongues", Journal of the Acoustical Society of America, 99: 1730-1741, 1996.

[7] Gaskell, G. M. "Modelling regressive and progressive effects of assimilation in speech perception", Journal of Phonetics, 31: 447-463, 2003.

[8] Gaskell, G. M. and Marslen-Wilson, W. D. "Phonological variation and inference in lexical access", Journal of Experimental Psychology: Human Perception and Performance, 22: 144-158, 1996.

[9] Goldinger, S. "Echoes of echoes? An episodic theory of lexical access", Psychological Review, 105: 251–279, 1998.

[10] Gow, D. W. "Does English coronal place assimilation create lexical ambiguity", Journal of Experimental Psychology: Human Perception and Performance, 28: 163-179, 2002.

[11] Hertrich, I. and Ackermann, H. "Phonological aspects of audiovisual speech perception" in Proceedings of 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 2007

[12] Isardi, W. "Some MEG correlates for distinctive features" in Proceedings of 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 2007

[13] Jesse, A. and McQueen, J. M. "Prelexical adjustments to speaker idiosyncrasies: Are they position-specific?" in Proceedings of Interspeech 2007, Antwerp, Belgium, 2007

[14] Kazanina, N., Phillips, C., and Isardi, W. "The influence of meaning on the perception of speech sounds", PNAS, 103: 11381-11386, 2006.

[15] Kraljic, T. and Samuel, A. G. "Generalization in perceptual learning for speech", Psychonomic Bulletin and Review, 13: 262-268, 2006.

[16] Lahiri, A. and Reetz, H., "Underspecified recognition". In: Gussenhoven C, Warner N (eds) "Underspecified recognition", p. 637-676. Mouton de Gruyter, Berlin, 2002.

[17] Massaro, D., "Perceiving talking faces: from speech perception to a behavioral principle", MIT Press, Cambridge, MA, 1998.

[18] McQueen, J. M., Cutler, A., and Norris, D. "Phonological abstraction in the mental lexicon", Cognitive Science, 30: 1113-1126, 2006.

[19] Mitterer, H., "Understanding "gardem bench": Studies on the perception of assimilation word forms [dissertation]", Universiteit Maastricht, Maastricht, The Netherlands, 2003.

[20] Mitterer, H. and Blomert, L. "Coping with phonological assimilation in speech perception: Evidence for early compensation", Perception & Psychophysics, 65: 956-969, 2003.

[21] Mitterer, H., Csépe, V., Honbolygo, F., and Blomert, L. "The recognition of assimilated word forms does not depend on specific language experience", Cognitive Science, 30: 451–479, 2006.

[22] Norris, D., McQueen, J. M., and Cutler, A. "Perceptual learning in speech", Cognitive Psychology, 47: 204-238, 2003.

[23] Obleser, J., Lahiri, A., and Eulitz, C. "Magnetic brain response mirrors extraction of phonological features from spoken vowels", Journal of Cognitive Neuroscience, 16: 31-39, 2004.

[24] Ohl, F. W. and Scheich, H. "Orderly cortical representation of vowels based on formant interaction", PNAS, 94: 9440-9444, 1997.

[25] Page, M. P. A. "What can't functional neuroimaging tell the cognitive psychologist?" Cortex, 42: 428-443, 2006.

[26] Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., and Roberts, T. "Auditory cortex accesses phonological categories: an MEG mismatch study", Journal of Cognitive Neuroscience, 12: 1038-1055, 2000.

[27] Pulvermüller, F., Huss, M., Kheri, F., Moscoso del Prado Martin, F., Hauk, O., and Shtyrov, Y. "Motor cortex maps articulatory features of speech sounds", PNAS, 103: 7865-7870, 2006.

[28] Remez, R. E., "Neural models of speech perception: a case history". In: Harnad S (ed) "Neural models of speech perception: a case history", p. 199-225. Cambridge University Press, Cambridge, Mass., 1987.

[29] Scargle, J. D. "Publication bias: The “file-drawer” problem in scientific inference", Journal of Scientific Exploration, 14: 91-106, 2000.

[30] Scott, S. K. and Wise, R. J. S. "The functional neuroanatomy of prelexical processing in speech perception", Cognition, 92: 13-45, 2004.