INTERNATIONAL JOURNAL OF RESEARCH IN COMMERCE, IT,

ENGINEERING AND SOCIAL SCIENCES

ISSN: 2349-7793 Impact Factor: 6.876., Volume: 16 Issue: 01 in January 2022

EFFECTS OF LINGUISTIC KNOWLEDGE ON SPEECH PERCEPTION

[1]S.O. Solijonov,
[1]Docent., Andijan State University
The head of English Phonetics Department

Article history:
Received: 12 th January., 2022
Accepted: 14 th January., 2022
Published: 16 th January., 2022

Abstract: *This article deals with the problems of phonology. It studies effects of linguistic knowledge on speech perception and the importance of teaching phonetics.*
Key words: *phoneme, speech perception, contextual variation, interpretation, phonological components.*

Speech signals are inherently variable, and one major source of speech variation is coarticulation. Coarticulation consists of universal biomechanical and language-specific phonological components: biomechanical constraints determine the direction of coarticulatory perturbations (e.g. /u/ is fronted in the context of alveolar consonants), and phonological knowledge guides the degree of coarticulation. Universal and language-specific components are also found in speech perception. Listeners‟ can generally compensate for systematic covariations of the acoustic properties of natural speech, but the degree of compensation varies systematically depending on listener's linguistic experience and listener expectation toward normative range of speech variation in certain linguistic contexts. There is a large body of experimental studies that examine the effects of context on the perception of target speech sounds.

In a commonly used methodology, an experimenter would create one or more acoustic continua, where the acoustic property of each sound is systematically altered along a relevant dimension such as VOT or formant frequency, so that the perceived phonemic category of each sound in a given continuum transforms, either gradually or abruptly, from one category into another. Each of these target sounds is embedded in a particular context, and subjects are asked to determine the phonemic identity of each target sound. The null hypothesis is that subjects‟ identification of the target sound will remain constant regardless of the context in which the sound occurs, and the alternative hypothesis is that subjects‟ identification of the target sound will vary in a way that demonstrates listener compensation for contextual effects.

Mann and Repp's study represents one of the early studies that used sound continua to examine listener compensation for coarticulation. They examined listener identification of synthetic fricative noise from an [s]-[ʃ] continuum when followed by either [a] or [u]. In natural speech production, fricative noise in /s/ is realized as a little more [ʃ]-like when followed by [u] because anticipatory lip protrusion for an upcoming round vowel lowers the center frequency of fricative noise. Thus, if the listener compensates for coarticulation, then the listener would identify an ambiguous fricative noise stimulus more often as [s] in the [u] context than in the [a] context. Mann and Repp‟s results showed this pattern. Their listeners‟ [s]-[ʃ] category boundary shifted toward the [s]-end in the [u] context relative to the boundary in the [a] context. Since then, numerous studies have shown converging results - perceptual category boundary shift—for

Copyright (c) 2022 Author (s). This is an open-access article distributed under the terms of Creative Commons Attribution License (CC BY). To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

consonant contexts and vowel targets, consonant contexts and consonant targets, and vowel contexts and vowel targets.

Compensatory perception occurs in the context of covarying features, as well. It has been repeatedly demonstrated that F0 tends to be lower for vowels immediately following voiced consonants than those following voiceless consonants, so if listeners compensate for this covariation, then they would more likely hear an ambiguous onset to be [+voice] when followed by a low-F0 vowel than a high-F0 vowel. Fujimura tested this hypothesis by using a synthesized stimulus series varying perceptually from [k] to [g]. For the ambiguous tokens from the middle of the continuum, listeners more often reported hearing [g] when the F0 of the following vowel is low. [1]

Further, compensatory effects can be triggered and the degree of the effects can be influenced by non-segmental contexts. For example, Ladefoged and Broadbent tested, among other things, listeners‟ identification of a synthesized /bVt/ word when played back after a precursor phrase, the F1 of which was shifted down from the standard precursor. The test word was identified as bit (/bɪt/) by 87% of the subjects when preceded by the standard precursor but the same word was identified as bet (/bɛt/) by 90% of the subjects when preceded by the precursor that had lower F1, presumably because listeners took the overall low- or high-F1 context into account when judging the height of the vowel in the test word. Later, Ohala and Shriberg showed that low-pass and high-pass filtering of the precursor phrase and the target vowel stimuli can alter listeners‟ perceptual judgments of the target vowels along the front-back dimension.

These findings offer two important insights. First, compensation and other contrastive context effects are closely related phenomena: compensation is achieved by a dynamic process involving both local-level adjustments of a target acoustic signal relative to the immediate context as well as larger-level adjustments of the perceptual scale. The second insight is that compensation is closely linked to listener knowledge about the various types of systematic and context-dependent surface variations found in day-to-day spoken communication. The next section reviews research on this second point—influence of linguistic knowledge on speech perception.[2]

Speech perception and word recognition involve interpreting acoustic signals in terms of phonemes and then to words. In addition, there is a rich body of evidence that higher-level knowledge such as semantics and lexical knowledge influence perceptual judgments on the lower-level linguistic unit such as phonemes and features. For example, Marslen-Wilson and Welsh have shown that listeners are able to shadow (i.e., repeat what they have just heard) faster when the sentences they were asked to repeat were both semantically and syntactically well-formed. Subjects were least successful in shadowing random meaningless sequences of words. For well-formed sentences, their subjects shadowed them with very short latencies, about 250 ms, or roughly the length of a single syllable.

This means that in polysyllabic words they were able to recognize and begin repeating a word even before it was presented completely. These results show that: 1) listeners start narrowing down lexical candidates the moment the speech signal starts; and 2) assuming lexical candidates expedite subsequent perceptual processing. In another study, Warren demonstrated that lexical knowledge causes the phoneme restoration effect. When a single segment within a word (i.e. /s/ in

---

[1] Amos, J. (2007). Wadda boo'iful place: an analysis of the variables (ju) and (t) in Mersea Island English. (Unpublished M.A. thesis). University of Essex.
Barnes, J. (2006).

[2] Phonetic laws. In E. Stankiewics (Ed. & Trans.), A Baudouin de Courtenay Anthology: The Beginning of Structural Linguistics (pp. 260-277).

ISSN 2349-7793 (online), Published by INTERNATIONAL JOURNAL OF RESEARCH IN COMMERCE, IT, ENGINEERING AND SOCIAL SCIENCES.,
under Volume: 16 Issue: 01 in January-2022
https://www.gejournal.net/index.php/IJRCIESS
Copyright (c) 2022 Author (s). This is an open-access article distributed under the terms of Creative Commons Attribution License (CC BY). To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

legislature) was replaced by a cough-like sound, his subjects recognized the word without any problem, and could not even tell which segment was replaced by the sound of cough, presumably due to restoration of the missing phoneme, which is guided by lexical knowledge. [3]

Later, Elman and McClelland showed that lexically restored phonemes can cause compensation for coarticulation. Prior to their study, Mann and Repp and Repp and Mann showed that American listeners shift perceptual phonemic category boundary location on a /t/-/k/ continuum toward the /k/-end (ambiguous sounds receive more /t/-responses) in a context of preceding /ʃ/ than in a context of preceding /s/, presumably because the listeners compensate for a coarticulatory retraction of /t/ when it is heard after /ʃ/. Elman and McClelland replicated this compensation effect by using a pair of words such as progress and abolish, for which the final phoneme is /s/ and /ʃ/, respectively, as contexts but with the final consonants replaced with a synthesized sound that is intermediate between [s] and [ʃ]. Their subjects tended to perceive the ambiguous final consonant as /s/ or /ʃ/ in a way to form a real word than a non- word context (e.g. progress is a real word but progresh is a non-word) and subsequently compensate for coarticulation on a target sound from a /t/-/k/ continuum.

Compensation for coarticulation is also induced by visual stimuli. For example, Fowler and her colleagues replicated Mann's (1980) finding for /da/ bias on /da/-/ga/ continuum when preceded by /ar/ but not /al/ (due to compensation for retraction and lowered F3 of /d/ after /r/) when the context syllable was perceptually ambiguous between /al/ and /ar/, but clearly disambiguated by a simultaneous video of a speaker hyperarticulating /alda/ or /arda/. Another type of listener knowledge that influences perceptual judgments of phoneme identity is the knowledge about gender variation in speech sounds. For example, in a vowel normalization study Johnson demonstrated that listeners actively adjust perceived vowel quality depending on perceived speaker identity. He used a hood-hud ([hʊd]-[hʌd]) continuum, and the target stimuli were embedded in a carrier sentence that had either a rising or falling F0 contour, ending at constant F0, which is same as target word‟s F0. These pitch contours were designed to mimic male speakers‟ interrogative (rising contour, starting with low F0) and female speakers‟ declarative (falling contour, starting with high F0) pitch contours. Listeners made more hood responses for the ambiguous tokens in the perceived female condition than in the perceived male condition. That the observed shift in perceptual judgment was not due to a formant shift in the precursor phrase as in the case of the Ladefoged and Broadbent highlights the role of listener expectation, in this case that male talkers tend to realize /ʊ/ as slightly lower variant, somewhat more similar to /ʌ/, than females.[4]

Evidence for the link between speech perception and phonological knowledge also comes from cross-linguistic studies on speech perception variation, which correlates with language- specific sound patterns. For example, velum lowering in Thai and American English vowel- nasal coda (VN) sequences starts during the vowel, but Thai exhibits less overlap than English. Consistent with this shorter duration of the nasal portion of the vowel, Thai listeners exhibit less compensation for nasalization in nasal contexts than English listeners; that is, Thai listeners perceive greater nasality from the nasalized vowels in [NVN] context than English listeners do. In addition, speakers of languages that differ in the degree of nasal overlap prefer different amounts of nasalization and

---

[3] Amos, J. (2007). Wadda boo'iful place: an analysis of the variables (ju) and (t) in Mersea Island English. (Unpublished M.A. thesis). University of Essex.
Barnes, J. (2006).
[4] Phonetic laws. In E. Stankiewics (Ed. & Trans.), A Baudouin de Courtenay Anthology: The Beginning of Structural Linguistics (pp. 260-277).

Copyright (c) 2022 Author (s). This is an open-access article distributed under the terms of Creative Commons Attribution License (CC BY). To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

temporal
patterns of overlap when judging stimulus naturalness. These studies show a language-specific relationship between patterns of vowel nasalization and the perceptual judgments on nasalized vowels.[5]

One of the significant implications of these studies is that the knowledge about the language-specific degree of coarticulation also influences perceived degree of coarticulatory perturbation on the segments. Another aspect of speech perception where a cross-linguistic difference has been observed is weighting of acoustic cues. In a study on the acoustic cues for place of articulation of stops in Japanese and American English, Fujimura, Macchi, and Streeter showed, firstly, that CV release cues dominate over VC closure cues when these cues conflict. Thus, for example, a stimulus made up by splicing /ab/ (except for the release burst) onto /da/ (starting from the burst) was heard as /ada/, instead of /abda/. Secondly, and more importantly for the purpose of the present review, their study showed different response patterns that were influenced by the stress/accent patterns of the subjects‟ native languages.

Only American subjects showed an attenuation of the dominance of the CV release cue when the [VCCV] stimuli had a high pitch V1 and low pitch V2 pattern compared with the opposite pitch pattern. American subjects responded to the release cue more strongly when it was high-pitched than low-pitched, presumably because the American subjects interpreted high-pitched syllables as stresses syllables. This study suggested that in addition to any physical differences between VC and CV cues, listeners‟ linguistic experience dictates which cues they pay most attention to. Collectively, findings from these studies suggest that memorized sound patterns and articulatory configurations for speech sounds and sequences of these sounds that make up words influences what listeners think they hear as well as how the perceptual system processes incoming acoustic signals.

References:

1. Alivuotila, L., Hakokari, J., Savela, J., Happonen, R-P., & Aaltonen, O. (2007). Perception and imitation of Finnish open vowels among children, naïve adults, and trained phoneticians. Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, pp. 361- 364.

2. Amos, J. (2007). Wadda boo‗iful place: an analysis of the variables (ju) and (t) in Mersea Island English. (Unpublished M.A. thesis). University of Essex.
Barnes, J. (2006).

3. Strength and weakness at the interface: Positional neu-tralization in phonetics and phonology. Berlin: Mouton de Gruyter. Baudouin de Courtenay, J. (1970).

4. Phonetic laws. In E. Stankiewics (Ed. & Trans.), A Baudouin de Courtenay Anthology: The Beginning of Structural Linguistics (pp. 260-277).

5. Indiana University Press. (Translated from French summary ―Les Lois phonétiques‖ (pp. 57-82) of ―O prawach głosowych,‗‖ Rocznik slawistyczny, 3 (pp. 1-57). Original published in 1910.)

[5] Amos, J. (2007). Wadda boo‗iful place: an analysis of the variables (ju) and (t) in Mersea Island English. (Unpublished M.A. thesis). University of Essex.
Barnes, J. (2006).

ISSN 2349-7793 (online), Published by INTERNATIONAL JOURNAL OF RESEARCH IN COMMERCE, IT, ENGINEERING AND SOCIAL SCIENCES., under Volume: 16 Issue: 01 in January-2022
https://www.gejournal.net/index.php/IJRCIESS
Copyright (c) 2022 Author (s). This is an open-access article distributed under the terms of Creative Commons Attribution License (CC BY). To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/