The Diachronic Influences of Perception: Experimental Evidence from Turkish
Author(s): Jeff Mielke

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via http://linguistics.berkeley.edu/bls/.

The Annual Proceedings of the Berkeley Linguistics Society is published online via eLanguage, the Linguistic Society of America's digital publishing platform.
The Diachronic Influence of Perception: Experimental Evidence from Turkish

JEFF MIELKE
The Ohio State University

1. Phonetics in Phonology: Conflicting Hypotheses
Two conflicting strategies have been proposed to explain the role of phonetic factors such as speech perception in phonology. According to the “Sound Change” view, phonetic explanation should be invoked solely in explaining diachronic changes which give rise to synchronic alternations (see, e.g., Ohala (1981, 1993), Anderson (1981), Dolby and Hansson (1999), Buckley (2000), Hale and Reiss (2000), Hume and Johnson (2001), Hyman (2001), Janda and Joseph (2002), Myers (2002)). According to the “Direct Influence” view, phonetic explanation should also be incorporated into synchronic phonology (see, e.g., Archangeli and Pulleyblank (1994), Pater (1996), Hayes (1997), Steriade (1997, 2001, this volume), Flemming (1995), Myers (1997), Mielke (2000)). The purpose of this study is to test some of the predictions made by these two approaches.

An example of Direct Influence is the P-map hypothesis (Steriade 2001, this volume), which states that speakers’ knowledge of the perceptibility of contrasts influences phonology directly: “The primary function of the P-map is to guide the speaker in search of the minimal input deformation that solves a phonotactic problem. The grammatical reflex of the P-map is the projection of and ranking among correspondence constraints” (Steriade, In press:4). An alternative explanation, argued by e.g., Ohala (1981, 1993 inter alia), Hume and Johnson (2001), and Hyman (2001), is that perception is relevant at the inception of a sound change, but it is not part of the grammar.

1 Thanks to Mike Armstrong, Savas Arslan, Mary Beckman, Robin Dodsworth, Shelome Gooden, Tsan Huang, Beth Hume, Brian Joseph, Keith Johnson, Alexei Kochetov, Kristie McCrary, Grant McGuire, Scott Myers, Dave Odden, Phonies, Janet Pierrehumbert, Charles Reiss, Anton Rytting, Misun Seo, Andrea Sims, Donca Steriade, the creators of TELL, Giorgos Tserdanelis, Pauline Welby, Steve Winters, Alan Yu, audiences at OSU, at the 2002 MOT Phonology Workshop in Montréal, and at the GDR Phonologies 4èmes Journées Internationales in Grenoble, and people at poster sessions at OSU’s CogFest and LabPhon 8, for valuable assistance with this paper, including many helpful comments and suggestions.
Both approaches seek to explain why many formally statable phonological processes fail to occur (The Too-Many-Solutions problem). Steriade’s solution to the Too-Many-Solutions problem is that a speaker’s knowledge of perceptual similarity projects a ranking of correspondence constraints, and this constraint ranking predicts attested patterns. An alternative solution to the Too-Many-Solutions problem is that perceptibility constrains the set of likely grammars diachronically. Perception and other phonetic factors may motivate a sound change at its inception (see Janda and Joseph 2002), but do not play an active role in the synchronic grammar. Because perception contributes to the likelihood of a sound change, its influence is seen in synchronic phonological alternations.

Some observations are consistent with the P-map hypothesis. For example, there is a crosslinguistic preference for less noticeable repair strategies (ones which involve maximally similar UR and surface forms), and speakers are aware of similarity rates, based on experimental evidence and on studies of rhyming patterns (Steriade, this volume, and references cited). Steriade concludes that the preference for less noticeable repair strategies is teleological, that speakers choose less noticeable repairs because they know they are less noticeable. Even if speakers possess similarity knowledge, this is not direct evidence that they apply it, so empirical evidence is needed.

Proving the P-map hypothesis requires more than just evidence that perception influences phonology (because the influence could be diachronic) and that speakers have knowledge of similarity (because this knowledge may not be applied to phonology). It requires testing whether perceptibility knowledge is synchronically accessible to and used by speakers. Because it assumes that perception influences phonology via speakers’ adaptive strategies (following, e.g., Lindblom 1990), the P-map hypothesis predicts that perception could manifest itself in speech style variation as speakers adapt to different situations (Steriade, p.c.). The sound change hypothesis does not make this prediction, so a close examination of a perceptually-conditioned process may be informative. Turkish /h/ deletion is one such process.

2. Turkish /h/ Deletion: Divergent Predictions

/h/ is optionally deleted in Turkish casual speech, but only in certain segmental contexts (Lewis 1967, Sezer 1986). Sezer reports that /h/ is optionally deleted before sonorant consonants (1a), but not after them (1b). /h/ is optionally deleted after voiceless stops (2b) and affricates (3b), but not before them (2a & 3a). /h/ is optionally deleted before and after voiceless fricatives (4a & 4b), and it is optionally deleted intervocically (5a), but not word-initially (5b). Sezer reports that /h/ does not delete word-finally (5c), but informal native speaker judgments indicate that it deletes in this environment as well, and a previous production experiment shows that word-final /h/ deletion is conditioned by the initial segment of the following word, occurring under conditions similar to those under which word-internal /h/ is deleted (Mielke 2002a).
The Diachronic Influence of Perception

(1) a. fihrist ~ fi:rist² ‘index’
teilhike ~ te:like ‘danger’
mehmet ~ me:met proper name
köhne ~ kö:ne ‘old’

b. merhum ~ *merum ‘the late’
ilham ~ *ilam ‘inspiration’
imha ~ *ima ‘destruction’
tenha ~ *tena ‘deserted’

(2) a. kahpe ~ *ka:pe ‘harlot’
sahite ~ *sa:te ‘counterfeit’
mahkum ~ *ma:kum ‘inmate’

b. jüþhe ~ jüþe ‘suspicion’
ethem ~ etem proper name

(3) a. ahtfi ~ *a:tfi ‘cook’

b. metþhul ~ metþul ‘unknown’

(4) a. mahsus ~ ma:sus ‘special to’
tahsil ~ ta:sil ‘education’
aufab ~ a:fab ‘made of brick’

b. ishal ~ isal ‘diarrhea’
safha ~ safa ‘step’
meþhur ~ meþur ‘celebrity’

(5) a. tofium ~ toum ‘seed’
müfiendis ~ müendis ‘engineer’
saian ~ saan ‘copper food dish’
müafaza ~ muafaza ‘protection’

b. hava ~ *ava ‘air’
c. timsah ~ ?timsa: ‘crocodile’

A perception experiment (Mielke 2002b, To appear) indicates that /h/ is generally less perceptible in environments where it deletes than in environments where it does not delete. Crucially, for present purposes, perceptibility varies substantially between different deletion environments, as seen below in (6). Because /h/ deletion is optional and perceptually conditioned, the different approaches make different predictions about how the deletion rates in relatively high and relatively low perceptibility environments will be affected by speech style variation.

² When it is deleted from preconsonantal or final position, compensatory lengthening of the preceding vowel occurs.
(6) Ability of Turkish listeners to perceive /h/ in its deletion environments (Mielke 2002b, To appear)

<table>
<thead>
<tr>
<th>Context</th>
<th>Sensitivity (d')</th>
</tr>
</thead>
<tbody>
<tr>
<td>V __ glide</td>
<td>3.462</td>
</tr>
<tr>
<td>V __ liquid</td>
<td>3.340</td>
</tr>
<tr>
<td>V __ nasal</td>
<td>3.303</td>
</tr>
<tr>
<td>intervocalic</td>
<td>2.897</td>
</tr>
<tr>
<td>V __ voiceless fricative</td>
<td>2.663</td>
</tr>
<tr>
<td>voiceless affricate __ V</td>
<td>2.633</td>
</tr>
<tr>
<td>voiceless stop __ V</td>
<td>2.408</td>
</tr>
<tr>
<td>voiceless fricative __ V</td>
<td>2.300</td>
</tr>
<tr>
<td>word-final</td>
<td>1.123</td>
</tr>
</tbody>
</table>

The P-map hypothesis assumes that Turkish speakers are aware of differences in perceptibility, and that they delete /h/ in environments where a change is less likely to be noticed. A further prediction is that in more careful speaking styles, /h/ will delete more frequently in environments where deletion is less noticeable. In a production experiment with /h/ deletion rate as a dependent variable, this should appear as a significant interaction between speaking style and segmental context.

According to the sound change hypothesis, differences in perceptibility are synchronically irrelevant, because only two arbitrary categories are relevant: the environments where deletion occurs and the environments where deletion does not occur. This predicts that speech style variation will affect deletion equally in all environments, and in careful styles, /h/ deletion will not be more skewed toward low-perceptibility environments than it is in reduced styles. Thus, there should be no significant interaction between speaking style and context.

3. Production Experiment

3.1. Methods

A production experiment was designed to test the predictions in the previous section. In order to induce speech style variation and /h/ deletion, 100 target words containing /h/ in various segmental contexts were elicited from native Turkish speakers in a word list, in a story, and in a story with a cognitive load. Subjects were nine female and one male native speakers of Turkish in Columbus, Ohio, aged 21-34. 100 target words with /h/ in the segmental contexts in (7) were selected from the TELL³ database. The target words were elicited from subjects three times and recorded in mono with a Shure SM10A head-mounted

³ The Turkish Electronic Living Lexicon (http://socrates.berkeley.edu:7037/TELLhome.html) was developed by Sharon Inkelas with John Lowe, Aylin Kuntay, Ronald Sprouse, Yelda Mesbah and Orhan Orgun.
microphone, a Symetrix SX202 dual mic preamp, and a Sony DTC-790 DAT recorder. The recordings were then digitized at 44100 Hz using Syntrillium’s CoolEdit audio editing software.

(7) Target words

<table>
<thead>
<tr>
<th>Deletion Environment</th>
<th>#</th>
<th>Non-deletion Environment</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop __ V</td>
<td>10</td>
<td>V __ voiceless stop</td>
<td>10</td>
</tr>
<tr>
<td>V __ nasal</td>
<td>10</td>
<td>nasal __ V</td>
<td>10</td>
</tr>
<tr>
<td>V __ liquid</td>
<td>10</td>
<td>liquid __ V</td>
<td>10</td>
</tr>
<tr>
<td>V __ voiceless fricative</td>
<td>10</td>
<td>word-initial</td>
<td>10</td>
</tr>
<tr>
<td>voiceless fricative __ V</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intervocal</td>
<td>10</td>
<td>TOTAL # OF TOKENS</td>
<td>100</td>
</tr>
</tbody>
</table>

The instructions for each successive production of the target words were designed to produce different speaking styles. Subjects read the 100 target words in a word list with 250 filler words in order to produce careful speech. Subjects also read the same target words in a story, which was intended to induce more reduced speech. Subjects read the story a second time with a cognitive load, to induce further speech reduction by distracting subjects from the speaking task (Harnsberger & Pisoni 1999). For the cognitive load, subjects memorized an 8-digit number before the first, third, and fifth pages of the six-page story and attempted to repeat each number after reading the next two pages. To avoid lexical priming by the word list, the story-reading tasks were performed before the word list task.

The 3000 target words were spliced out of context and separated into two groups according to whether /h/ was retained or deleted, according to the following criteria. If the word contained an audible /h/, deletion did not occur. If /h/ was inaudible or ambiguous, spectrograms were created, and if the spectrogram indicated any sign of an /h/ gesture, i.e., voicelessness or frication not attributable to another consonant, deletion did not occur. If the spectrogram indicated no sign of an /h/ gesture, deletion did occur.

3.2. Results and Analysis

For the 60 target words with /h/ in environments where Sezer (1986) reports deletion (8), /h/ deletion rates were 15.8% (95/600) for the word list task, 39.5% (237/600) for the story task, and 46.5% (279/600) for the story with a cognitive load. For the 40 target words with /h/ in environments where deletion is not reported, /h/ deletion rates were 3.5% (14/400) for the word list task, 19.3% (77/400) for the story task, and 26.0% (104/400) for the story with a cognitive load. Pairwise t-tests indicated no significant differences between deletion rates in any of the environments where deletion is not reported.
A univariate analysis of variance (ANOVA) was performed on the arcsine transforms of the deletion rates in the six deletion environments, and found main effects for **task** \( F(2,162) = 49.043; \ p < 0.001 \) and **environment** \( F(5,162) = 24.698; \ p < 0.001 \), but no significant interaction between **task** and **environment** (which is predicted by the P-map hypothesis) \( F(10,162) = 0.532 \). A second ANOVA was performed for just the 'story' and 'story with cognitive load' tasks, and found a main effect for **environment** \( F(5,108) = 16.722; \ p < 0.001 \), a near-effect for **task** \( F(1,108) = 3.562; \ p = 0.062 \), and again, no significant interaction between **task** and **environment** \( F(5,108) = 0.094 \).

Deletion rates (V = vowel, L = liquid, N = nasal, S = voiceless fricative, T = voiceless stop, # = word boundary)

Deletion rates are lower in the three most salient environments (V_L, V_N, and V_V), but this is not limited to any particular speaking style, as predicted by the P-map hypothesis. The reason for this becomes clear upon examining the results for the four environments where deletion is not claimed to occur. Some deletion occurs in all environments, including the four environments where deletion is not reported by Sezer (1986), and the deletion rates in three of the environments where deletion is reported are not significantly higher. Significantly higher rates of deletion exist only for three environments (V_S, T_V, and S_V).

Comparing the six deletion environments with the environments where deletion is not reported to occur reveals that deletion rates are particularly high only for three of the reported deletion environments (9).\(^4\) Pairwise t-tests show

\(^4\) Another sensible place to look for explanation is the frequency of words with /h/ in these contexts. A search of a newspaper corpus of about 350,000 words gathered by Kemal Oflazer in the late 1990s shows no correlation between deletion environments and frequency of occurrence of words which contain /h/ in these environments. Of the environments in (9), by far the most frequent in both token and type frequency are word-initial /h/ and intervocalic /h/. Among the consonant environments, two of the environments with high deletion rates are very low in
that the deletion rates before and after voiceless fricatives and after voiceless stops are significantly higher than in each of the other environments \([p < .002]\). The clustering of environments into two categories suggests the existence of more than one deletion process in Turkish, something which is addressed below.

(9) Deletion rates in all environments (averaged across three tasks)

![](image)

4. Discussion
The results support the sound change hypothesis. The lack of a significant interaction between task and environment means that the distribution of deletions does not change according to speaking style. The overall rate of deletion does indeed increase as speech style becomes more casual, but the slope does not change significantly. This means that perceptibility does not appear to be relevant for determining where deletions will occur in these speaking styles. The results of the experiment provide no evidence that perception is part of the grammar.

The results do provide further support for the role of phonetic explanation in phonology, because the deletion pattern found in this study is very much in line with the predictions of a perceptual account. In half of the environments where deletion is reported by Sezer (1986), the deletion rates are virtually the same as in environments where deletion is not reported. Deletion rates are significantly higher only before and after voiceless fricatives and after voiceless stops. The deletion rates in these environments form a homogenous subset apart from the other seven, and they also share two properties which make them likely candidates for conditioning /h/ deletion. First, /h/ is less salient to Turkish listeners in these environments than in any other environment included in the study, due ostensibly to the presence of masking frication and aspiration noise frequency, but /h/ after voiceless stops is more common than all but two of the other consonant environments.
which is adjacent to /h/ in these environments but either nonexistent or not adjacent to /h/ in the other seven (Mielke 2002b, To appear). Second, /h/ is adjacent to a segment with laryngeal opening and closing gestures in these three environments (and also before voiceless stops). These environments are more prone to articulatory simplification via aggregation of laryngeal gestures, as seen in the results of Munhall and Löfqvist’s (1992) “Kiss Ted” experiment. This is further evidence for the importance of perception and production factors as external forces which shape phonological systems.

Based on deletion rates found in this study, one possible conclusion is that the deletion process described here is not the same as the one described by Sezer. This study finds a nominal rate of deletion in all environments, and a smaller set of environments where a distinct deletion process apparently applies. The differences between the two sets of findings may be attributable to regional variation, generational differences, change in the nature of the /h/ deletion process over the course of 25 years, or simply the fact that Sezer’s study is based on native speaker intuitions and this one is based on experimental data.\footnote{Repetition of the above ANOVAs with the seven low-deletion environments and the three high-deletion environments analyzed separately also found no significant interaction, although in this case the set of high-deletion environments is too small to find a main effect for environment.}

The results of this study suggest that the Too-Many-Solutions problem has already been solved in language change. If a process is already active in a language, the Too-Many-Solutions problem is never even posed to the language learner or language innovator. This is consistent with Myers’ (2002) proposed solution to the problem of gaps in the factorial typology of Optimality Theory. Myers (2002:28) argues that gaps in factorial typology are “pervasive and natural”, because the alternations which do not occur are those that cannot arise through simple sound changes from phonetic patterns.

This explanation is similar to Steriade’s in that both are explanatory on functional grounds. The difference lies in whether or not the functional grounding which transparently interacts with language change should be included in the synchronic phonology as well. Steriade has demonstrated that inclusion of perceptual knowledge in the phonological grammar is possible, because speakers have perceptual similarity judgments, and because languages frequently employ relatively non-salient repair strategies. However, it has not been shown that speakers employ similarity judgments in their phonological processing. This study of the interaction of phonology with speech style variation finds no evidence for such a synchronic relationship between functional factors and phonology. That said, there are a number of ways in which this argument against the P-map hypothesis can be shown to be wrong.

First, one could argue that deletion is not the best example of a listener-oriented phonological process, and that the lack of evidence for listener-oriented adaptive strategies in Turkish /h/ deletion does not rule out the possibility of finding such strategies in fortition processes or in hyperarticulated speech.
The Diachronic Influence of Perception

Whether the latter finding would be significant for phonological theory is a matter for debate. Second, if such perceptually-based adaptive strategies exist, they might manifest themselves more readily in types of speech that are more obviously listener-oriented, such as a conversation in a noisy room or infant-directed speech. Third, if one chose to take the position that deletion is more prevalent in fast speech exclusively because fast speech requires more effort, and not because deletion in fast speech is less noticeable (following Kirchner’s (2001) effort-based approach to consonant lenition), then it can be argued that the P-map hypothesis does not predict that deletion should be more frequent in perceptually weak environments in some speech styles and not in others.

While these caveats are duly noted, this study provides experimental evidence that is convergent with a conclusion that has been reached by many phonologists, i.e., that sound change is not part of the grammar (Ohala 1995), that diachronic phonology is not the same as synchronic phonology (Hyman 2001), that phonology does not need to explain all historical facts (Buckley 2000), that phonetic grounding explains facts which can be derived without reference to phonology (Hale and Reiss 2000), and that functional factors such as perception, production, generalization, and conformity can be construed as diachronic filters (Hume and Johnson 2001).

Functional explanation can exist without being in the grammar, so harmonic constraint rankings (which are used in Optimality Theory in an effort to make similar typological predictions) may be epiphenomenal rather than explanatory. Given what is known about the requirements of production, perception, communication, learnability, cognition, social identity, language change, etc., how much crosslinguistic regularity of phonological patterning (phonetically-motivated or otherwise) has a grammar-external explanation? To what extent are other phonological constructs such as distinctive features epiphenomenal rather than explanatory? (Mielke 2003, In prep.)

5. Conclusion
The explanation for why /h/ deletion is more prevalent before and after voiceless fricatives and after voiceless stops involves a conspiracy of perceptual and articulatory factors. Although this phonological process can be explained functionally, no listener-oriented adaptive strategies are apparent in the speech of the talkers in the experiment, indicating that this functional explanation lies in the domain of sound change rather than in synchronic phonology. Therefore it is shown that speakers can possess functionally-motivated phonological rules without any synchronic access to the functional motivation. This conclusion supports removing the burden of explanation from formal synchronic accounts of phonological alternations when explanation can be found in language change and language use. Functional factors certainly explain much about observable patterns in phonological systems, and functional explanation is crucial to explaining how individual phonological systems came to be the way they are, but it does not
necessarily explain how phonology works in real time inside the mind of each speaker.

References


The Diachronic Influence of Perception

Mielke, J. 2002b. Turkish /h/ deletion: evidence for the interplay of speech perception and phonology. In M. Hirotani, ed. NELS 32, 383-402. Amherst, Mass.: GLSA.

Jeff Mielke
222 Oxley Hall
1712 Neil Ave
Columbus, OH 43210
mielke@ling.ohio-state.edu