Visual Lexical Stress Information in Audiovisual Spoken-Word Recognition

Alexandra Jesse [1], *James M. McQueen* [1]

[1] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

alexandra.jesse@mpi.nl, james.mcqueen@mpi.nl

Abstract

Listeners use suprasegmental auditory lexical stress information to resolve the competition words engage in during spoken-word recognition. The present study investigated whether (a) visual speech provides lexical stress information, and, more importantly, (b) whether this visual lexical stress information is used to resolve lexical competition. Dutch word pairs that differ in the lexical stress realization of their first two syllables, but not segmentally (e.g., 'OCtopus' and 'okTOber'; capitals marking primary stress) served as auditory-only, visual-only, and audiovisual speech primes. These primes either matched (e.g., 'OCto-'), mismatched (e.g., 'okTO-'), or were unrelated to (e.g., 'maCHI-') a subsequent printed target (*octopus*), which participants had to make a lexical decision to. To the degree that visual speech contains lexical stress information, lexical decisions to printed targets should be modulated through the addition of visual speech. Results show, however, no evidence for a role of visual lexical stress information in audiovisual spoken-word recognition.

Index Terms: visual prosody, lexical stress, priming, audiovisual spoken-word recognition

1. Introduction

Recent research in spoken-word recognition using auditory-only speech input has shown that lexical competition among words is resolved not only based on segmental information but also based on suprasegmental information [1-3]. In Dutch, word pairs exist that differ in the suprasegmental lexical stress realization of their first two syllables, but not segmentally (e.g., 'OCtopus' and 'okTOber'; capitals marking primary stress). That is, the stressed and unstressed versions of the syllables differ in fundamental frequency (F0), amplitude, and duration. Hearing the first two syllables of such words as auditory fragment primes (e.g., 'OCto-') leads to faster lexical decision responses to a subsequently presented printed target when it matches ('OCtopus') the fragment in stress compared to the case where the prime was an unrelated word fragment (e.g., 'maCHI-' from 'machine') [2]. Likewise, hearing the first two syllables of words that match segmentally but mismatch suprasegmentally in their stress ('okTOber') leads to slower responses to a printed target (*octopus*) than when it is preceded by an unrelated word fragment. That is, suprasegmental auditory stress information is used to resolve lexical competition in on-line spoken-word recognition.

While segmental information strengthens both the target and the stress competitor in both the matching and the mismatching priming condition, suprasegmental information has a different effect across conditions. In the matching priming condition, suprasegmental information adds more support to the target than to the stress competitor. In the mismatching priming condition, suprasegmental information weakens the target but strengthens the stress competitor. That is, in the matching condition, target recognition benefits from both matching segmental and suprasegmental information. In the mismatching condition, the target is segmentally supported but is disfavoured suprasegmentally.

As we know, however, speech is a multimodal phenomenon. Visual speech aids word recognition by providing information about the word's segments [e.g., 4]. But visual speech also provides prosodic information [e.g., 5-10]. However, little is known about whether and how visual prosodic information influences word recognition. Only few studies have investigated the transmission of lexically-defined stress on the word level in visual speech. Minimal word pairs in English (such as '(to) obJECT' and '(an) OBject') and in Swedish can be discriminated above chance in visual-only speech presentations [5,6]. Production data for the English minimal lexical stress pairs showed no correlation with head or eyebrow movement [7]. The articulatory correlates of stress in this study were instead found to be interlip distance and chin opening.

Similarly, phrase-level accents (prominence) that are conveyed by similar acoustic variables can also be detected in visual-only speech [8-10]. Here, seeing the lower part of a face is sufficient for visual prominence detection [9], although eyebrow and head movement also influence prominence detection [10]. Generally, even though they are unreliably present, eyebrow and head movements seem to correlate with changes in F0 [11-13].

These results show that visual speech can provide information about phrase and word-level stress, but it is unclear whether this information is indeed used during word recognition to resolve lexical competition. The present study investigates therefore whether visual cues to lexical stress can also be used in lexical disambiguation.

The experiment was modelled closely on previous experiments using the cross-modal fragment priming paradigm [e.g., 2]). Minimal Dutch stress pairs such as 'OCtopus'-'okTOber' were recorded on video, spoken by a native female speaker of Dutch. The first two syllables of these words or of an unrelated word (e.g., 'maCHI-') were presented as auditory-only, visual-only, or audiovisual speech fragment primes. After each prime, a printed target item was displayed. The participants had to indicate by button press whether or not the target is a word in Dutch. These lexical decisions (e.g., to *octopus*) should be faster when preceded by a matching fragment prime ('OCto-') and slower when preceded by a mismatching fragment prime ('okTO-') compared to the case where the prime is segmentally unrelated ('maCHI-'). To the degree that visual speech contains lexical stress information, this auditory effect should be replicated for visual-only speech and should be strengthened through the addition of visual speech to auditory speech.

2. Experiment

2.1. Methods

2.1.1. Participants

Fifty-five native Dutch speakers from the MPI participant pool were paid for their participation. Data from seven participants were excluded from all analyses due to equipment failure during the experiment.

2.1.2. Stimuli

Twenty-four Dutch word pairs that shared the same segments in their two initial syllables and only differed in their stress pattern were selected. Two pairs had primary stress either on their first or second syllable (e.g., 'SYllabus' and 'syLLAbe'; syllabus, syllable). Eight pairs either had primary stress on the first or the third syllable (e.g., 'CAvia' and 'kaviAAR'; guinea pig, caviar). Fourteen pairs had either primary stress on the second or third syllable (e.g., 'saLAmi' and 'salaMANder'; salami, salamander). Note that all items with primary stress on the third syllable had also secondary stress on the first syllable. All items in a pair were semantically and morphologically unrelated. All items shared also the first phoneme of the third syllable. One exception is octopus-oktober, where the phonemes of their third syllable only share their manner and place of articulation, but are therefore visually highly similar. Twenty-four control items were selected that had two initial syllables that were auditorily and visually unrelated to those of the respective targets. Visual unrelatedness was based on Dutch viseme classes [14]. All but eight target words pairs were equated on overall syllable length, that is, they consisted each of three syllables. Each of the eight exemption pairs consisted of one item with three syllables and one item with four syllables. Control items showed a similar distribution of stress as the targets and were equated on their mean word frequency to the targets [15].

Twenty-four additional stress pairs and matching control items were selected as fillers. These items were closely equated to the target stress pairs and did not qualify as targets mainly because the item pairs differed in the first phoneme of their third syllable from another. Ninety-six words were selected as primes for nonword trials. All nonwords were phonotactically legal in Dutch. Forty-eight of these nonwords were preceded by segmentally related word fragment primes (i.e., nonwords began with the word fragment prime); the other half by unrelated word fragment primes. Overall, an equal number of word and nonword trials were presented to each participant, with half of each type followed by segmentally related, half by segmentally unrelated fragment primes.

All stimuli were video recorded by a female native Dutch speaker. The speaker was naïve regarding the purpose of the study and had not received any further instructions about visual stress realization (e.g., was not instructed to specifically move or not move the head). All word primes were recorded embedded sentence-finally in different neutral sentences. Stress pairs and controls were recorded in two sentences each. For example, 'Zonder enige aanleiding zei de oude vrouw' ('Without any instructions said the old woman') and 'Voor galgje koos ik het woord' ('For hangman I chose the word') were both recorded ending with 'cavia' (guinea pig), 'kaviaar' (caviar), and 'tolerant' (tolerant). Two sentences were needed so that experimental and control primes of a given item pair could be presented in two different sentences to a participant. The boundary between second and third syllable was determined based on the waveform in Praat. The video was then cut accordingly at the next upcoming frame boundary. The part between auditorily-determined boundary and frame boundary was then set to silence in Adobe Audition. A 5ms linear ramp was applied to the audio track before the silence. Note that video and audio were never separated during stimulus editing and also that they were presented only as a complete video in the experiment. For visual-only presentations, the audio track was simply muted; for auditory-only presentations, the video track was hidden.

2.1.3. Procedure and Design

Participants were instructed that on each trial they would either first see, or hear, or see and hear a person speak. Immediately afterwards, a printed word or nonword would appear on the screen. The task was to indicate with a button press as fast and as accurately as possible whether this item was a word or not in Dutch. A practice phase consisting of twelve trials preceded the main experiment.

Priming condition was implemented as a between-subject variable. All participants would see all critical stress pair items once as a target on the screen. One item of each pair was preceded by its control prime (e.g., 'tole-' - *kaviaar)*, the other item by an experimental prime that was either matching (e.g., 'CAvi-' - *cavia*) or mismatching (e.g., 'kavi-' - 'cavia') in its stress pattern. One group of participants *('matching' group*) received only matching and control primes for the critical trials and therefore only mismatching and control primes for stress filler trials. The other group *('mismatching' group*) were exposed only to mismatching and control primes for the critical trials and matching and control primes for stress filler trials. A third of each type of trials were presented under each modality condition. Note that, for a given participant, the control and experimental primes of an item pair were always presented in the same modality condition. Modality condition, sentence, and priming condition were counterbalanced in lists across participants.

2.2. Results and discussion

Table 1 shows the raw mean lexical decision latencies (RT) for correct responses in control and experimental conditions for auditory-only, visual-only, and audiovisual presentation conditions. Reaction times were measured from the onset of the printed target presentation. Table 2 shows the error rates for these conditions. Figure 1 shows the priming effect for the matching and mismatching priming conditions for each modality condition for mean reaction times. The priming effects for reaction times are the difference between mean reaction time for correct responses in control and experimental conditions for each modality. Figure 2 shows the priming effects for each condition for percentage of errors. Priming effects here are the differences between percentage of errors in control and experimental conditions for each modality.

For mean RTs, two ANOVAs with experimental condition (experimental prime vs. control prime), priming type (matching vs. mismatching), and presentation modality (auditory-only, visual-only, or audiovisual speech prime) as fixed factors and subjects and items as random factors, respectively, showed no significant effect of experimental condition, priming type, or presentation modality on reaction times (for all effects $p>.05$). Only the interaction between

modality and priming type was marginally significant ($F_s(2,92)=2.35$, p=.10; $F_i(2,45)=1.22$, p=.31).

Planned comparisons between the six experimental prime conditions and their respective controls were conducted. These comparisons were all directional tests, in that, for response latencies, it was tested for each modality condition whether responses in the matching prime condition were faster than in their control condition and whether responses in the mismatching prime condition were slower than in their control condition. Likewise, for error rates, it was tested whether fewer errors were made in the matching prime conditions than in their control conditions, and whether more errors were made in the mismatching prime conditions than in the control conditions.

Table 1. *Mean lexical decision latencies (mean reaction times, in ms, from target onset) based on target pairs for each priming type under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

Mo--dality	Matching		Mismatching	
	Control	Exp.	Control	Exp.
A	638	624	641	662
V	634	633	632	624
AV	644	616	626	614

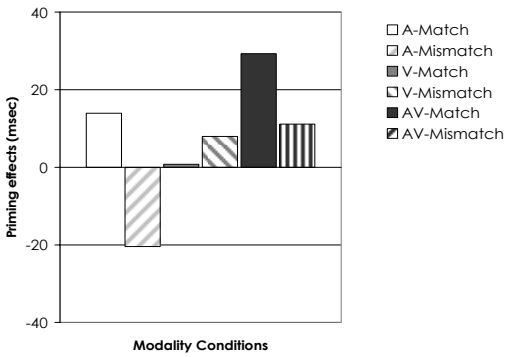

Figure 1: *Priming effects for RT on target stress pairs for each priming type (matching, mismatching) under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

Planned comparisons for response latencies showed an only marginally significant facilitatory priming effect for the audiovisual matching manipulation ($t_s(23)=1.70$ p=.052; $t_i(47)=1.61$ p=.057). Responses were faster when preceded by an audiovisual prime that matched with the target (M=616msec) than when compared to a control prime condition (M=644msec). There was no inhibition effect for the audiovisual mismatching condition, and no priming effect for any of the visual-only conditions (all p>.05). For the auditory-only condition, there was no effect for the auditory matching condition ($t_s(23)=.92$ p=.18; $t_i(47)=.66$ p=.26), but a marginally significant inhibitory trend for the auditory mismatching condition ($t_s(23)=1.22$ p=.12; $t_i(47)=$ 1.08 p=.14).

A further planned comparison between the priming effects of auditory-only and audiovisual speech showed no significant increase of priming from auditory to audiovisual presentation condition for matching ($t_s(23)=.39$ p=.35; $t_i(47)=.72$ p=.24) and mismatching priming ($t_s(23)=1.25$ p=.11; $t_i(46)=$ 1.54 p=.07).

For ANOVAs on mean percentage of errors, there was also no effect of experimental condition or modality (all p>.05), but a significant effect of priming type ($F_s(1,46)=3.21$, p=.08; $F_i(1,47)=6.08$, p<.05), with more errors in the group where targets were preceded by mismatching and control primes (M=7.55%) than in the group where targets were preceded by matching and control primes (M=5.04%). This priming type effect interacted with presentation modality, however ($F_s(2,92)=3.48$, p<.05; $F_i(2,45)=3.16$, p=.052). There were higher error rates in the mismatching group than in the matching group for auditory-only (7.29% vs. 5.99%) and audiovisual presentation conditions (9.38% vs. 3.13%), but not for visual-only conditions (5.99% for both groups).

Table 2. *Mean error rates based on target pairs for each priming type under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

Mo--dality	Matching		Mismatching	
	Control	Exp.	Control	Exp.
A	4.17	7.81	6.77	7.81
V	5.21	6.77	5.21	6.77
AV	4.69	1.56	7.81	10.94

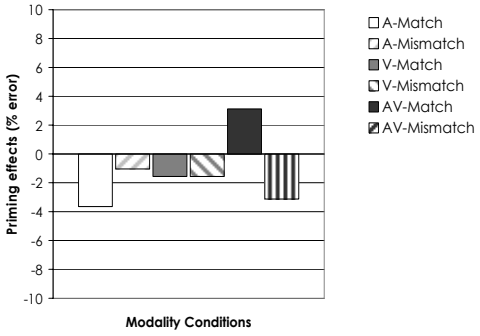

Figure 2: *Priming effects for percentage of errors on target stress pairs for each priming type (matching, mismatching) under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

Planned directional comparisons showed again a marginally significant facilitatory priming effect for the audiovisual presentation in the matching condition ($t_s(23)=1.45$ p= .08; $t_i(47)=-1.63$ p=.055). More errors were made in the audiovisual control condition (M=4.69%) than in the audiovisual matching condition (M=1.56%). For error rate, there was also a marginally significant inhibitory priming effect for the audiovisual mismatching condition ($t_s(23)=1.06$ p= .15; $t_i(47)=1.35$ p=.09). Fewer errors were made in the audiovisual control condition (M=7.81%) than in the audiovisual mismatching condition (M=10.94%).

Note that for the auditory matching condition, the difference was opposite to the direction that was predicted. More errors were made in the auditory-only matching (M=7.81%) than in the control condition (M=4.17%). An additional non-directional post-hoc test showed that this difference was marginally significant ($t_s(23)=1.77$ p=.09; $t_i(47)=1.63$ p=.11). This reversed effect for the auditory-only condition was also partially the reason why the size of priming increased between auditory-only and audiovisual presentation for the matching group ($t_s(23)=2.40$ p<.05; $t_i(47)=2.22$ p<.05). There was no increase in priming between auditory-only and audiovisual presentation for the

mismatching priming condition (t_s(23)=.49 p=.31; t_i(47)=1.73 p=.29).

To increase power, the stress fillers were added to the analysis. Note that these twenty-four stress filler pairs were similar to the experimental items and only differed in that the first phoneme of the third syllable of each pair was not always identical. The *'matching' group* had received experimental stress pairs under the matching priming condition and stress fillers under the mismatching priming condition. For the *'mismatching' group*, the opposite was the case. This means, with stress fillers added to the data set, prime type became a within-subject rather than a between-subject variable (see Table 3 for mean RTs and Table 4 for mean percentage of errors; see Figure 3 for RT priming effects and Figure 4 for error priming effects).

Prime type had a significant effect on reaction time (F_s(1,47)=4.18, p<.05; F_i(1,90)=8.15, p<.01). Responses for words preceded by matching primes or their controls (M=646msec) were faster than when preceded by mismatching primes or their controls (M=658msec). The effect of modality was marginally significant over items but not over subjects (F_s(2,94)=1.59, p=.21; F_i(2,180)=2.81, p=.06). The interaction between prime type and modality was also significant (F_s(2,94)=13.54, p<.001; F_i(2,89)=8.53, p<.001). The interaction between prime type and experimental condition was significant over items, but only marginally significant over subjects (F_s(1,47)=3.13, p=.08; F_i(1,90)=4.45, p<.05). The triple interaction between prime type, experimental condition, and modality was only marginally significant (F_s(2,94)=2.66, p=.08; F_i(2,180)=2.24, p=.11).

Table 3. *Mean lexical decision latencies (mean reaction times, in ms, from target onset) based on target and filler stress pairs for each priming type under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

Mo-	Matching		Mismatching	
-dality	Control	Exp.	Control	Exp.
A	657	633	669	696
V	659	670	646	645
AV	658	623	658	664

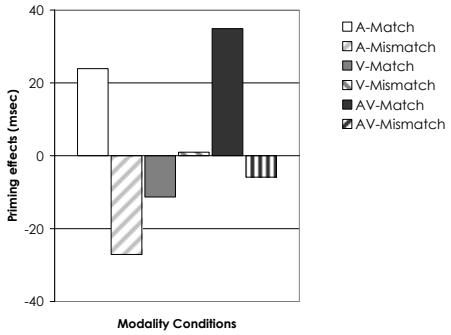

Figure 3: *Priming effects for RT on target and filler stress pairs for each priming type (matching, mismatching) under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

Planned comparisons for mean RTs between experimental priming conditions and their respective controls showed a marginally significant facilitatory effect of matching primes in the auditory-only presentation condition (t_s(47)=1.11, p=.14; t_i(92)=1.59, p=.06). Target responses were faster when preceded by a matching auditory prime (M=633msec) than when preceded by an unrelated auditory prime (M=657msec). There was a significant inhibitory effect of mismatching primes in the auditory-only condition (t_s(47)=1.97, p<.05; t_i(93)=1.61, p<.05). Target responses were slower when preceded by a mismatching prime (M=696msec) than by an unrelated auditory prime (M=669msec). None of the priming effects for the visual-only condition was significant (p>.05). Responses after matching audiovisual primes (M=623msec) were faster than after unrelated audiovisual primes (M=658msec; t_s(47)=2.67, p<.01; t_i(94)=2.37, p<.01). However, there was no inhibitory priming effect in the audiovisual mismatching condition (t_s(47)=.14, p=.44; t_i(95)=.34, p=.18). The size of the priming effects did not increase from auditory-only to audiovisual presentations for the matching condition (t_s(47)=1.00, p=.16; t_i(91)=.54, p=.30). For the mismatching condition, the priming effect decreased in the audiovisual compared to the auditory-only condition (t_s(47)=1.70, p<.05; t_i(93)=.72, p=.24). That is, seeing and hearing a speaker saying a mismatching prime did not lead to more inhibition than when only hearing the speaker. However, this could be due to the fact that visual speech also provides segmental information. Seeing in addition to hearing a speaker say 'okTO' should also support both *octopus* and *oktober* more than unrelated competitors by providing visual segmental information. If the addition of visual speech would not provide any suprasegmental information then the audiovisual benefit in support compared to the auditory-only case should be equal for the matching and the mismatching priming condition. This means that the difference in priming effects for the audiovisual and the auditory-only conditions should be the same for matching and mismatching conditions. If, however, visual speech indeed also provides suprasegmental stress information, then the difference between the priming effects from auditory to audiovisual speech should be larger for the matching condition than for the mismatching condition. However, this was not the case here (t_s (47)=.58, p=.28; t_i(90)=.10, p=.46). Therefore, there is no evidence for visual speech providing suprasegmental stress information.

Table 4. *Mean error rates based on target and filler stress pairs for each priming type under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

Mo-	Matching		Mismatching	
-dality	Control	Exp.	Control	Exp.
A	11.72	9.64	10.68	10.42
V	8.33	10.16	6.51	9.38
AV	8.07	7.29	8.85	11.72

In the ANOVAS on error rate, there was a marginally significant effect of modality (F_s(2,94)=2.56, p=.08; F_i(2,190)=2.96, p=.054), and a marginally significant interaction of modality with priming type (F_s(2,94)=2.01, p=.14; F_i(2,94)=2.53, p=.09). Planned comparisons showed more errors were made after a mismatching audiovisual prime (M=11.72%) than after an unrelated prime (M=8.85%; t_s(47)=1.36 p= .09; t_i(95)=1.37 p=.09). Also there was a marginally significant tendency for more errors to occur in the visual-only condition after mismatching primes (M=6.51%) than after control primes (M=9.38%; t_s(47)=1.40 p=.08; t_i(47)=1.55 p=.06). No comparison of priming effects between auditory-only and audiovisual was significant (all

p>.05). Furthermore, the difference between matching and mismatching conditions did not vary between auditory-only and audiovisual conditions (all p>.05).

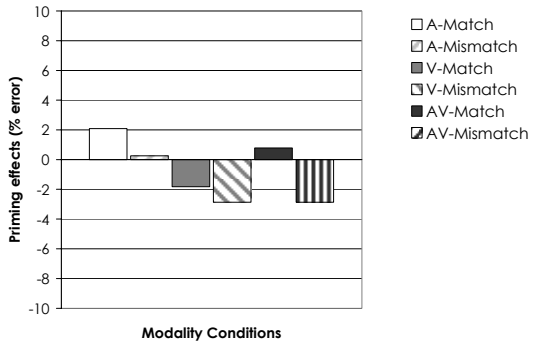

Figure 4: *Priming effects for percentage of errors on target and filler stress pairs for each priming type (matching, mismatching) under each modality condition (A= auditory-only, V= visual-only, AV= audiovisual).*

3. General Discussion

This study investigated whether visual speech provides information about lexical stress of words and whether this information is used during word recognition to resolve lexical competition. Results failed to show any evidence for the use of visual lexical stress information in spoken word recognition.

With forty-eight stress targets analyzed, only evidence for facilitatory priming for matching audiovisual speech primes was found in the response latencies. No inhibition was found for audiovisual mismatching primes (although both effects were found for error rate in the audiovisual condition). No priming effects were found for visual-only conditions and only trends for the auditory-only condition. Therefore, the experiment not only failed to provide any evidence for visual speech cues to lexical stress but also failed to replicate previous auditory-only studies [e.g., 2]). However, this failure could be due to the differences in design. In the present study, each participant only contributes eight data points to each condition. In addition, priming had here to be manipulated as a between-subject rather than a within-subject variable as in previous studies to increase the number of data points provided by each subject from four to eight.

Adding the stress fillers as additional targets shows that with increased power, not only a significant facilitatory priming effect was found for audiovisual matching primes on response latencies, but also an inhibitory priming effect for auditory-only mismatching primes. The facilitation for auditory-only matching primes is only marginally significant. However, there was no evidence for priming in the visual-only condition and furthermore the size of the priming effects was not modulated through the addition of visual speech to auditory speech in a way that could clearly be interpreted as evidence for visual lexical stress information.

Follow-up investigations will examine whether the minimal fragment pairs used in this study differ sufficiently in their acoustic stress characteristics, and whether visual stress can be discriminated in an off-line two-alternative forced-choice task.

4. Acknowledgements

This work was supported by a grant from the German Science Foundation (DFG) to the first author.

5. References

[1] Cooper, N., Cutler, A., and Wales, R. “Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners”, Language and Speech, Vol. 45, 2002, pp 207-228.

[2] Donselaar, W. van, Koster, M., and Cutler, A. “Exploring the role of lexical stress in lexical recognition”, Quart. J. Exp. Psych., Vol. 58A, 2005, pp 251-273.

[3] Soto-Faraco, S., Sebastián-Gallés, N., and Cutler, A. “Segmental and suprasegmental mismatch in lexical access”, J. Memory and Lang., Vol. 45, 2001, pp 412-432.

[4] Sumby, W.H., and Pollack, I. "Visual contribution to speech intelligibility in noise", JASA, Vol. 26, 1954, pp 212–215.

[5] Keating, P., Baroni, M., Scarborough, R., Alwan, A., Auer, E.T., and Bernstein, L.E. "Optical phonetics and visual perception of lexical and phrasal stress in English", Proc. 15th ICPhS, 2071- 2074, 2003.

[6] Risberg, A., and Lubker, J. "Prosody and speech-reading", Speech Transmission Lab. Quarterly Progress Status Rep., Vol. 4, 1978, pp 1-16.

[7] Scarborough, R., Keating, P., Baroni, M., Cho, T., Mattys, S., Alwan, A., Auer Jr., E., and Bernstein, L.E. "Optical Cues to the Visual Perception of Lexical and Phrasal Stress in English", Speech Prosody, 217-220, 2006.

[8] Bernstein, L.E., Eberhardt, S.P., and Demorest, M.E. "Single-channel vibrotactile supplements to visual perception of intonation and stress", JASA, Vol. 85, 1989, pp 397-405.

[9] Dohen, M., Loevenbruck, H., Cathiard, M.-A., and Schwartz, J.-L. "Visual perception of contrastive focus in reiterant French speech", Speech Comm., 44, 2004, pp 155-172.

[10] Swerts, M., and Krahmer, E. "Cognitive Processing of Audiovisual Cues to Prominence", *Proc. AVSP,* 29-30, 2005.

[11] Cavé, C., Guaïtella, I., Bertrand, R., Santi, S., Harlay, F., and Espesser, R. "About the relationship between eyebrow movements and F0 variations", Proc. Interna. Conf. Speech Lang. Processing, 2175-2178, 1996.

[12] Munhall, K.G., Jones, J.A., Callan, D.E., Kuratate, T., and Vatikiotis-Bateson, E. "Visual prosody and speech intelligibility: Head movement improves auditory speech perception", Psych. Sci., Vol. 15, 2004, pp 133-137.

[13] Yehia, H. C., Kuratate, T., and Vatikiotis-Bateson, E. "Linking facial animation, head motion and speech acoustics", J. Phon., Vol. 30, 2002, pp 555-568.

[14] Son, N. van, Huiskamp, T.M.I., Bosman, A.J., and Smoorenburg, G.F. "Viseme classifications of Dutch consonants and vowels", JASA, Vol. 96, 1994, pp 1341-1355.

[15] Baayen, R.H., Piepenbrock, R., and Gulikers, L "The CELEX Lexical Database (Release 2)" [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor], 1995.