PATTERNS OF INDIVIDUAL DIFFERENCES IN REDUCTION: IMPLICATIONS FOR LISTENER-ORIENTED THEORIES

Rory Turnbull

Ohio State University
turnbull@ling.osu.edu

ABSTRACT

Many listener-oriented theories of phonetic reduction assume that the talker has tacit knowledge of an interlocutor's mental state, and consequently predict that talkers with poor theory of mind should exhibit inconsistent behavior between semantic contexts. This study examined effects of individual differences in theory of mind skills on extent of phonetic reduction in three acoustic domains (word duration, vowel duration, and vowel dispersion) in contexts of high and low semantic predictability. Overall, a trend was observed whereby participants with stronger theory of mind skills produced less phonetic reduction than participants with poorer theory of mind skills. These results are not consistent with the predictions of listener-oriented theories, or indeed other contemporary theories. We conclude that individual differences ought to be considered in more detail by phonetic theory.

Keywords: phonetic reduction, H&H theory, individual differences, theory of mind, predictability

1. INTRODUCTION

1.1. Effects of semantic predictability on reduction

Lieberman [15] established that semantic predictability affects the phonetic realization of words. For instance, the word *nine* in (1) is relatively unpredictable compared to the word *nine* in (2); Lieberman found that the *nine* in (1) was more intelligible in isolation than the *nine* in (2).

(1) The next number you will hear is nine.

(2) A stitch in time saves nine.

Subsequent acoustic studies have confirmed that words produced in predictable contexts tend to have shorter word duration, shorter vowels, and less dispersed vowels relative to words produced in unpredictable contexts [9, 21].

1.2. Listener-oriented theories

One family of explanations for these effects is the listener-oriented theories of phonetic reduction [2, 17]. According to these theories, speech production involves a balance between the competing demands of conservation of effort on the part of the talker, and maximization of intelligibility on the part of the listener. To achieve this balance, talkers reduce predictable words, which can be inferred from context, while unpredictable words are not reduced, so as to preserve intelligibility. A corollary of this mechanism is that talkers must be able to construct and update a mental model of their interlocutor(s), so that they reduce (or fail to reduce) in contextually appropriate positions.[1] This modeling of one's interlocutor requires a well-developed theory of mind, the ability to impute mental states to others [18].

1.3. Theory of mind

At its most basic, theory of mind (ToM) involves understanding that other individuals have thoughts and beliefs which are different from one's own. Despite ToM often being discussed in terms of an individual either 'having' or 'not having' ToM, it is best understood as a gradient of aptitude [22]. It has been demonstrated that neurotypical adults show variance in the extent and application of their ToM skills [14, 16, 19].

One assessment designed to test ToM in adults is the 'Reading the Mind in the Eyes' test (RMITE) [5], an emotional cognition task. In this test, participants are shown a picture of the eye region of the face and are asked to choose which of four words correctly expresses what the person in the picture is thinking or feeling. This test has been shown to be a useful measure of ToM ability and has good test-retest reliability over long periods of time [5, 10]. Scoring the test provides a score from 0 to 36, where a higher score indicates more adept use of ToM. Chance performance—that is, guessing—gives an expected value of 9 ± 3. Previous studies have suggested that a mean score of around 27 is expected for neurotypical populations [5, 10], meaning that this test is appropriate for quantifying variability within a non-clinical sample.

Another ToM assessment is the Autism-spectrum Quotient (AQ) [7]. The link to ToM is less direct than with the RMITE. The AQ measures the extent of cognitive and personality traits which resemble aspects of the autism phenotype. An underdevel-

oped ToM has been proposed to be one of the central deficits of the autism phenotype [4], and although other traits have been identified as being central to autism, the basic assertion that people with autism lack a ToM remains unchallenged [8, 11]. The AQ score can therefore be used as an indirect measure of ToM, although it should be borne in mind that it also reflects other traits (such as attention-switching skill). The score is derived from a self-report questionnaire, where participants respond to statements such as "I find it difficult to work out people's intentions" along a four-point scale from 'strongly disagree' to 'strongly agree'. Potential scores range from 50 (fewest autistic traits) to 200 (most autistic traits). Research into neurotypical samples suggests means of 109.4 for males ($SD = 14.0$), and 104.1 ($SD = 13.5$) for females [1]. Note here that the AQ and RMITE tests have opposite polarity—a higher score in the AQ indicates greater prevalence of autistic traits, and therefore by hypothesis a poorer ToM, while a *lower* score in the RMITE indicates a poorer ToM.

2. METHODS

2.1. Materials

Eighty-two sentences, in high- and low-predictability pairs, were selected from the SPIN test [13]. Examples of high- and low-predictability sentences are provided in (3) and (4), respectively. Note that the final word is either predictable or unpredictable from the preceding sentential context.

(3) For your birthday I baked a cake.

(4) Tom wants to know about the cake.

2.2. Procedure

Sentences were presented on a computer screen, blocked by predictability condition, and participants were instructed to read the sentences aloud "as if talking to a friend". After reading the sentences, participants completed the AQ questionnaire and the RMITE test.

2.3. Measurements

Measures of word duration, vowel duration, and vowel midpoint F1 and F2 were taken of the target (final) word of each sentence. The duration measures were taken in milliseconds, and the formant measures in ERB.[2] The formant values were used to calculate vowel dispersion which is defined for each token as the Euclidean distance from that token to the talker-specific vowel center. The vowel center is defined as the mean of by-vowel means for that talker.

2.4. Participants

Twenty-one undergraduate participants completed the task for partial course credit. Eligibility was restricted to native monolingual speakers of American English with no history of speech, language, or hearing disorders. None of the participants reported any history of autism-spectrum conditions.

2.5. Analysis

A linear mixed effects regression model was constructed for each of the acoustic variables—word duration, vowel duration, and vowel dispersion. Fixed effects were predictability condition (high versus low, reference: high), AQ score, RMITE score, and 2-way interactions between predictability condition and AQ, and predictability condition and RMITE. All continuous variables were centered around the mean before being entered into the model. Random intercepts for talker and word identity were included, with random slopes for predictability condition for both talker and word. In all models, p-values were calculated by treating the absolute t-statistic as if it came from a t-distribution with degrees of freedom equal to the number of data points minus the number of model parameters [3].

From a total of 1,722 target tokens (21 participants $\times$ 82 sentences), 72 tokens were excluded due to disfluencies, restarts, or speech errors. Thus, 1,650 tokens were included in the analysis.

3. RESULTS

3.1. Individual difference scores

AQ scores ranged from 90 to 128, with a mean of 111.48 ($SD = 11.02$), and RMITE scores ranged from 15 to 33, with a mean of 26.92 ($SD = 4.77$). These figures are not far from the expected population means established in prior work [1, 5]. AQ and RMITE scores were not correlated with each other (Pearson's $r = -0.065, p = 0.757$), suggesting that the two scores measure somewhat different aspects of ToM. Due to the lack of covariance between these two measures, they can be safely entered into a regression model together.

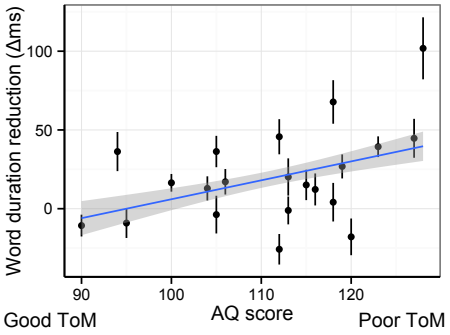
3.2. Word duration

The output of the regression analysis predicting word duration is summarized in Table 1. A simple effect of predictability condition was observed, such that unpredictable words ($M = 429$ms) were longer than predictable words ($M = 410$ms). This result is consistent with prior research, and demonstrates that phonetic reduction occurred. No simple effects of the individual difference scores (AQ and RMITE) were observed, but significant interactions were ob-

Table 1: Model output for word duration model. Significant p-values in boldface.

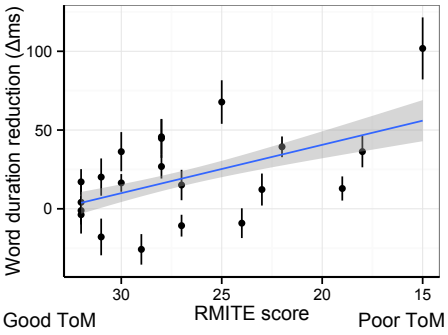
Effect	β	t	p
Intercept	−8.207	−0.447	0.655
Pred.: Low	19.874	3.885	**< 0.001**
AQ score	0.723	0.301	0.763
RMITE score	1.976	0.906	0.365
Pred. × AQ	2.759	2.425	**0.015**
Pred. × RMITE	−3.052	−2.943	**0.003**

Figure 1: Reduction in word duration from low to high predictability conditions as a function of talker AQ score. This and all subsequent graphs depict subject means and standard errors with linear trend and confidence interval overlaid.



served. AQ score interacted with predictability, such that participants with higher AQ scores had a larger difference in word duration between the predictable and unpredictable contexts than participants with lower AQ scores. That is, these participants had a larger extent of phonetic reduction than low-AQ participants. A similar interaction was observed with the RMITE scores, whereby participants with higher RMITE scores had shorter words in the unpredictable condition than participants with lower RMITE scores. After recalling that higher scores in the AQ are interpreted the same as lower scores on the RMITE, a clear pattern emerges: participants with less ToM ability have larger or more extensive reduction in word duration between the high and low predictability conditions. The interaction between predictability and AQ is visualized in Figure 1, and that of predictability and RMITE is visualized in Figure 2. Note that the x-axis of Figure 2 is inverted so that, in all the graphs, participants with poorer ToM skills cluster toward the right, while those with better ToM skills cluster toward the left.

Figure 2: Reduction in word duration as a function of talker RMITE score.



3.3. Vowel duration

The output of the regression analysis predicting vowel duration is summarized in Table 2. A simple effect of predictability condition was observed, such that vowels in unpredictable words (M = 188ms) were longer than vowels in predictable words (M = 181ms). No other significant simple effects or interactions were observed. However, note that the interactions between AQ and predictability and between RMITE and predictability are in the same direction as those observed in the word duration model. That is, higher AQ and lower RMITE scores both correlated with a greater extent of phonetic reduction (duration differences between high and low predictability conditions). However, given the lack of a statistically significant result, this trend cannot be reliably interpreted.

Table 2: Model output for vowel duration model. Significant p-values in boldface.

Effect	β	t	p
Intercept	−4.276	−0.425	0.671
Pred.: Low	8.778	2.023	**0.043**
AQ score	−1.045	−0.524	0.600
RMITE score	−0.810	−0.447	0.655
Pred. × AQ	1.085	1.104	0.270
Pred. × RMITE	−0.355	−0.396	0.692

3.4. Vowel dispersion

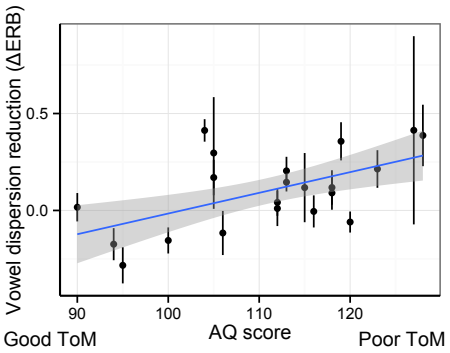
The output of the regression analysis predicting vowel dispersion is summarized in Table 3. A simple effect of predictability condition was observed, such that vowels in unpredictable words (M = 10.471 ERB) were more disperse than vowels in predictable words (M = 10.420 ERB). Additionally, a significant interaction between AQ and predictability condition was observed, in the same di-

rection as that observed for the word duration model: participants with higher AQ scores had more difference in vowel dispersion between words in the different contexts than participants with lower AQ scores, suggesting greater phonetic reduction for those participants. This interaction is depicted in Figure 3. No significant interaction of RMITE score was observed, but the sign of the coefficient is consistent with the effect established for word duration.

Table 3: Model output for vowel dispersion model. Significant p-values in boldface.

Effect	β	t	p
Intercept	−0.03597	−0.114	0.910
Pred.: Low	0.09163	2.409	**0.016**
AQ score	−0.01679	−0.707	0.479
RMITE score	0.03650	1.694	0.090
Pred. × AQ	0.02847	3.310	**0.001**
Pred. × RMITE	−0.00843	−1.067	0.286

Figure 3: Reduction in vowel dispersion as a function of talker AQ score.



4. DISCUSSION AND CONCLUSION

For all three acoustic variables (word duration, vowel duration, and vowel dispersion), a predictability effect was observed such that productions were more phonetically reduced (shorter duration and less disperse vowels) for words in predictable contexts. This result is consistent with prior research on semantic predictability-based reduction [9, 15, 21]. Additionally, interactions with individual differences in ToM measures were observed, with the overall trend being that participants with poorer ToM skills produced larger acoustic differences between the two predictability conditions than participants with better ToM skills. That is, these participants were in some sense more 'sensitive' to the semantic predictability manipulation.

A possible interpretation of these results could be that high-AQ or low-RMITE participants trend toward hyperspeech in general, giving them more 'room' for reduction. However, this interpretation is countered by the lack of any simple effects of AQ or RMITE on the acoustic measures.

At face value, this result, whereby individual differences in ToM are implicated in individual differences in extent of phonetic reduction, would seem to lend support to the listener-oriented view. However, the listener-oriented account predicts that being more adept at modeling one's interlocutor should lead to more adept use of phonetic reduction to enhance or reduce words in the appropriate contexts. The patterns revealed in the current study are in fact the opposite—participants who are less adept at ToM tasks appear to be very effective at reducing based on predictability, while close examination of Figures 1, 2, and 3 suggests that some participants with strong ToM ability are in fact not producing any phonetic reduction. In spite of their ability to model their interlocutor, they are doing nothing to assist them in comprehension.

A potential, *post hoc* explanation for these patterns relies on a perspective from the broader autism phenotype. Here, we note that people with autism tend to have great skill in 'systematizing', or noticing patterns in the whole, and understanding and predicting the behavior of non-agentive systems [6, 23]. To the extent that neurotypical adults with high AQ scores approximate people with autism, and to the extent that semantic predictability can be conceived of as a system with patterns, then it is expected that people with a more autistic cognitive style should be more able to notice and use systematic trends in language to aid interlocutor comprehension. However, this explanation is *post hoc*, ignores other aspects of the autism phenotype which suggest that the local is prioritized over the global in autistic perception [12], and therefore requires further empirical support.

This study examined effects of semantic predictability on phonetic reduction, and, in line with previous studies, established that words in predictable contexts are phonetically reduced relative to words in unpredictable contexts. Additionally, individual differences of reduction were observed, and these differences were found to correlate with scores on tasks assessing theory of mind ability. Participants with more adept theory of mind produced less phonetic reduction than participants with poorer theory of mind. These somewhat unexpected findings suggest that phonetic theory needs to take individual differences seriously to develop cognitively plausible mechanisms for speech communication.

5. ACKNOWLEDGEMENTS

I wish to thank Cynthia Clopper for comments and suggestions on an earlier draft. This research was partially supported by a Presidential Fellowship from the Ohio State University Graduate School.

6. REFERENCES

[1] Austin, E. J. 2005. Personality correlates of the broader autism phenotype as assessed by the autism spectrum quotient (AQ). *Personality and Individual Differences* 38(2), 451–460.

[2] Aylett, M., Turk, A. E. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47(1), 31–56.

[3] Baayen, R. H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

[4] Baron-Cohen, S., Leslie, A. M., Frith, U. 1985. Does the autistic child have a 'theory of mind'? *Cognition* 21(1), 37–46.

[5] Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I. 2001. The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry* 42(2), 241–251.

[6] Baron-Cohen, S., Wheelwright, S., Lawson, J., Griffin, R., Hill, J. 2002. The exact mind: Empathising and systemising in autism spectrum conditions. In: Goswami, U., (ed), *Handbook of Cognitive Development*. Oxford: Blackwell.

[7] Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E. 2001. The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders* 31(1), 5–17.

[8] Bennett, T. A., Szatmari, P., Bryson, S., Duku, E., Vaccarella, L., Tuff, L. 2013. Theory of mind, language and adaptive functioning in ASD: A neuroconstructivist perspective. *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 22(1), 13–19.

[9] Clopper, C. G., Pierrehumbert, J. 2008. Effects of semantic predictability and regional dialect on vowel space reduction. *Journal of the Acoustical Society of America* 124, 1682–1688.

[10] Fernández-Abascal, E. G., Cabello, R., Fernández-Berrocal, P., Baron-Cohen, S. 2013. Test-retest reliability of the 'reading the mind in the eyes' test: a one-year follow-up study. *Molecular Autism* 4(1), 33.

[11] Frith, U. 2001. Mind blindness and the brain in autism. *Neuron* 32, 969–979.

[12] Frith, U., Happé, F. G. E. 1994. Autism: beyond "theory of mind". *Cognition* 50, 115–132.

[13] Kalikow, D. N., Stevens, K. N., Elliott, L. L. 1977. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America* 61(5), 1337–1351.

[14] Keysar, B., Lin, S., Barr, D. J. 2003. Limits on theory of mind use in adults. *Cognition* 89, 25–41.

[15] Lieberman, P. 1963. Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech* 6, 172—187.

[16] Lin, S., Keysar, B., Epley, N. 2010. Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology* 46, 551–556.

[17] Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, W. J., Marchal, A., (eds), *Speech production and speech modelling*. Dordrecht: Kluwer.

[18] Premack, D., Woodruff, G. 1978. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1(4), 515–526.

[19] Schneider, D., Lam, R., Bayliss, A. P., Dux, P. E. 2012. Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science* 23(8), 842–847.

[20] Turk, A. 2010. Does prosodic constituency signal relative predictability? A smooth signal redundancy hypothesis. *Laboratory Phonology* 1(2), 227–262.

[21] Turnbull, R., Clopper, C. G. 2013. Effects of semantic predictability and dialect variation on vowel production in clear and plain lab speech. *Proceedings of Meetings on Acoustics* 19(060116).

[22] Yirmiya, N., Erel, O., Shaked, M., Solomonic-Levi, D. 1998. Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin* 124(3), 283–307.

[23] Yu, A. 2010. Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change. *PLoS ONE* 5(8), e11950.

[1] Some formulations of these accounts allow for a 'generic listener', rather than a model of the specific interlocutor [20].

[2] ERB, or equivalent rectangular bandwidth, is a psychoacoustically-motivated scale, similar to the Bark scale.