

Gass and Varonis 1984 and considerations of comprehensibility and intelligibility

BROWNE Kevin C.

Abstract

Gass and Varonis' (1984) article, *The effect of familiarity on the comprehensibility of nonnative speech* has been widely cited in the literature of applied linguistics since its publication. Their conclusions continue to be echoed by researchers, especially by those investigating both nonnative speech processing and how different types of familiarity function contribute to speech processing. Among their findings were that accent familiarity contributes to comprehensibility. This paper presents a strong theoretical stance that their conclusions that familiarity with nonnative speech in general and accent-familiarity contribute to the comprehensibility of nonnative speech was not reliably measured and that their findings revealed more about intelligibility than comprehensibility due to their choice of methodology that relied solely on transcription task data. Included also is a discussion of the problem of how terms like intelligibility, comprehensibility and interpretability have been applied, sometimes interchangeably, by researchers that likely led to Gass and Varonis' attributing their findings to comprehensibility. The findings of Gass and Varonis (1984) are important and valuable to all researchers interested in how different types of familiarity impact speech processing, and this paper does not aim to discredit it; however, their findings and conclusions warrant review and may more appropriately describe how familiarity affects intelligibility than comprehensibility.

key words: intelligibility, comprehensibility, comprehension

I. Introduction

A widely cited piece of research concerning how different types of familiarity affects the comprehensibility of L2 accented English is Gass and Varonis' (1984) article *The effect of familiarity on the comprehensibility of nonnative speech* (e.g., Anderson-Hsieh and Koehler, 1988; Munro, Derwing and Morton, 2006; Smith and Nelson, 1985). They hypothesized their construct of comprehensibility as an aspect of speech processing that can be measured through word/utterance identification alone, which was arguably (and is currently) more often a methodology applied to measuring intelligibility (Browne,

2016; Browne & Fulcher, 2016; Field, 2005; Jenkins, 2000); however, this point has not been raised for discussion by any researchers. The following is a critical review of this important piece of research that presents a strong theoretical stance that their findings revealed how different types of familiarity affect the intelligibility of nonnative speech, and that their conclusions concerning how familiarity with nonnative speech in general and with a particular accent contribute to comprehension were made without any compelling evidence from the data. Presented first is a summary of Gass and Varonis' study and their main findings followed by a discussion of how

terms like comprehensibility and intelligibility have been defined and applied inconsistently throughout the literature. These are followed by an explanation of the rationale supporting the hypotheses presented in this paper and a conclusion.

II. Overview of Gass & Varonis (1984)

The aims Gass and Varonis investigated were the effects four types of familiarity have on native speakers' ability to process nonnative speech. The four types of familiarity they investigated were:

1. Familiarity with topic of discourse-both with a specific topic and based on "real world" knowledge familiarity
2. Familiarity with nonnative speech in general
3. Familiarity with a particular nonnative accent
4. Familiarity with a particular nonnative speaker

Four speaker participants, all male nonnative speakers (NNS) of English (Arabic L1 n=2; Japanese L1 n=2), were selected from a group of fifteen advanced level English as a Second Language (ESL) students, and 142 native speaker (NS) students were recruited as listening participants from the University of Michigan in the United States. The four speaker participants were deemed to be "equally comprehensible" (p.67) by a review of ten ESL teachers, and were each recorded completing three tasks: (1) reading the North Wind story; (2) reading a set of five 'related sentences' that pertained to the story though themselves were not included in the reading; and (3) reading a set of 'unrelated sentences' that had nothing to do with the story whose contexts or topics were considered pertaining to 'real world knowledge.' The recordings

were used to produce 24 different 'tapes.' Each tape included different combinations of speakers reading the story, related sentences and unrelated sentences in various orders. For example, 'tape 3' had Japanese Speaker 2 reading the related sentences first followed by Arabic Speaker 1 reading the story and then the unrelated sentences, but 'tape 8' included Arabic Speaker 1 reading the related sentences first, followed by Japanese Speaker 1 reading the story and then Japanese Speaker 2 reading the related sentences. In this way, they were able to create sequences of speaker and sentence type relative to either pre or post story reading in order to attempt to measure the four targeted facets of familiarity, and provided four independent variables:

1. The four speakers
2. The two possible positions of the speakers (pre-story or post-story)
3. Three potential possibilities for the post-story position (1. different accent; 2. different speaker; 3. same speaker)
4. Two conditions ('related' or 'unrelated') of the sentences to the story

Each of the 142 NS participants listened to one tape each and completed two tasks: transcribing each sentence from both lists and writing a short summary of the story to determine comprehension. Mistakes in the sentence transcriptions were counted as 'errors', and the mean number of errors were calculated for each speaker. These error calculations comprised the only analyses of the data.

As stated earlier, comprehensibility was measured through successful word/utterance recognition, but it was never clearly defined as such. Rather than clearly define comprehensibility in a definition form, it was instead schematized how they believed

comprehensibility of nonnative speech to native speakers occurs:

$$C = p_{\alpha} + g_{\beta} + f1_{\gamma} + f2_{\delta} + f3_{\epsilon} \dots fl_{\zeta} + s_{\eta} \dots$$

C=comprehensibility,

p=pronunciation, g=grammar,

f1=familiarity with topic,

f2=familiarity with person,

f3=familiarity with speaker's native language, fl=fluency, s=social factors

(Gass & Varonis, 1984, p.67)

This schematization takes into account various aspects of speech, but the measures of the effects of different familiarities on comprehensibility were determined through transcription error ratios alone. This approach is puzzling, as transcription tasks were already commonly being used at that time to measure intelligibility (e.g.: Catford, 1950; Smith & Rafiqzad, 1979). Comprehensibility was more commonly measured using rating scales to score how much the listeners felt they 'understood', or locutionary force (e.g., Smith & Rafiqzad, 1979). The only aspect of Gass and Varonis' research that delved into locutionary force was having the listening participants write brief summaries of the story they heard "so that we could determine whether they had indeed understood the story" (p.69). However, there was no discussion of these summaries or any results of analyses of that data anywhere in the paper beyond that mentioning collecting them in the methodology. It can only be speculated, therefore, that all of the summaries the listening participants provided were accurate descriptions of the North Wind story, but again that is only speculation. It is equally interesting that neither terms, 'intelligible' nor 'intelligibility' are included anywhere in the paper that could have served to clarify how they differed from comprehensibility in their work. The basis of all of their findings and

conclusions derive solely from the results of transcription exercises (intelligibility tasks) that did not include measurements or other evidence to confirm the locutionary force of speakers' utterances were successfully conveyed. It is therefore arguable to suggest that the findings of Gass and Varonis (1985) are more associated with the how different types of familiarity impact intelligibility and not comprehensibility. Of course, in order to justly confirm or contest their findings it would be necessary to clearly understand how they interpreted and differentiated the terms 'comprehension', 'comprehensibility', 'intelligibility' and 'intelligible'. Unfortunately, we do not know.

Gass and Varonis' found that of the four familiarity types researched, 'familiarity of topic' had the greatest impact to intelligibility. Results of one-tailed t-tests comparing the pre- and post-text positions of the related sentences revealed significant differences of means of errors ($p < .05$) for three of the four speakers (p.72). They found more errors reported in the pre-story transcriptions of the 'related' sentences than those from the post-story position, which suggests that NSs were more capable of determining the content of NNSs' speech if they were familiar with the specific topic. Similarly, the 'unrelated' sentences composed of information reflecting 'real world knowledge' experienced significantly lower instance of errors ($F = 19.64$, $p = .0001$) when compared to the 'related' sentences when they were read in the pre-story position on the tapes. The differences in error ratios suggest that native speakers employ syntactic and contextual information to decipher segmental errors. These findings support Dauer's (2005) conclusions that native speakers utilize top-down strategies for comprehension and

intelligibility.

Familiarity of speaker, familiarity of accent and familiarity of nonnative speech in general were found to contribute to the intelligibility of nonnative speech though these findings were not based on any significant differences in the data. Familiarity of speaker and familiarity of accent were determined contributors to intelligibility by observing speaker error instances in the pre and post story positions. Listeners did tend to find the speakers of the same accent more intelligible in the post story sentences when they had encountered the other speaker of the same accent reading either the pre-story sentences or the story, as well as for individual speakers but not in all cases. 'Japanese speaker two' was not found to be more intelligible reading the related sentences in the post position when he had provided some example of his speech prior when compared to his error reports having heard 'Japanese speaker one' prior. 'Arabic speaker two' had the same lack of improved error reports for the unrelated sentences. In short, no significant results were reported to substantiate the claim that familiarity of speaker contributes to intelligibility. Additionally, the data only revealed a 'tendency' for familiarity of nonnative speech in general to contribute to intelligibility. The evidence they provided for this claim was the, "small difference between pre-text and post-text unrelated sentence" (Gass & Varonis, 1984, p.77). Additionally, no data was collected or included concerning the listener participants' overall familiarity with nonnative speech to demonstrate comparative differences, so it can only be speculated what their actual familiarity with particular accents, the included accents or nonnative speech in general were.

III. Comprehensibility and intelligibility

In order to claim that the findings of Gass and Varonis reveal more about intelligibility than comprehensibility, it is necessary to examine how the terms have been interpreted in the literature. Intelligibility and comprehensibility have been defined throughout the literature in many ways and at times interchangeably. Jenkins (2000) commented, "there is yet no broad agreement on a definition of the term 'intelligibility': it can mean different things to different people" (p.70). Comprehensibility likewise shares a similar inconsistency of its meaning among researchers.

Munro, Derwing and Morton (2006), for example, suggest intelligibility is "the extent to which a speaker's utterance is actually understood" and comprehensibility, "refers to the listener's estimation of difficulty in understanding an utterance" (p.112). They measured intelligibility through transcription task and comprehensibility with a rating scale. This interpretation was also applied by Field (2005) and Kennedy and Trofomovich (2008) with all three papers including locutionary force as an aspect of comprehensibility. Interestingly, Derwing and Munro (2009) maintained that comprehensibility is a measure of how easy or difficult it is for a listener to understand a speaker but removed locutionary force from the interpretation stating that, "this dimension is a judgment of difficulty and not a measure of how much actually gets understood" (p.478). All of these studies employed rating scales to measure what the listeners perceived was their levels of understanding with or without locutionary force as an included component.

Part of the problem concerning how these terms have been interpreted and continue to

be is that if research demands intelligibility or comprehensibility include locutionary force of utterances, intentions of speakers or not, there are enough precedents in the literature to allow it or for researchers to invent new ones. The problem of so many interpretations of the terms has long been discussed. Smith and Nelson (1985) were among the first researchers to argue that intelligibility must be clearly distinguished from comprehensibility and interpretability in the literature and appealed to researchers to finally settle on how the terms should be applied. They argued that by separating 'intelligibility', 'comprehensibility' and 'interpretability' to describe word/utterance identification, locutionary force and illocutionary force respectively would provide greater clarity, and that it was not prudent to warehouse so many concepts and measurements into one term. Sadly, their plea has not been heeded. This paper echoes their plea.

IV. Theoretical argument and rationale

Gass and Varonis claim that familiarity with nonnative speech in general and accent-familiarity are contributing factors to word/utterance identification (or 'intelligibility' as this paper defines it), but they did so without significant or compelling evidence. As stated earlier, it is arguable because there was no discussion of the data that could have revealed if the locutionary force of the utterances they listened to were perceived. In fact, there is very little discussion in their paper dedicated to these two claims. It appears that perhaps the inclusion of accent-familiarity in the study was a post hoc theory that occurred during the analyses. The conclusion that familiarity with a particular nonnative accent may have been an afterthought in their research is based

on the following passage:

A final point to mention is that there is a tendency for the Arabic speakers to elicit more comprehension errors than the Japanese speakers. There are a number of possible explanations for this: (1) It may be that many of our subjects were more familiar with a Japanese accent than an Arabic one, as opposed to the experienced teachers who were very familiar with both (in fact, a number of native-speaking judges mentioned having Oriental TAs)

(Gass & Varonis, 1984, p.74)

It is clear that accent familiarity data was not collected from the listening participants for either of the accents represented in the samples. Why Gass and Varonis made a point to suggest that perhaps why some judges may have been familiar with Japanese accented English was because some participants had "Oriental TAs". This would imply that all 'Orientals', or Asians, speak with the same or similar accents, which is simply not true. Their claim, therefore, that familiarity with a speaker's nonnative accent contributes to NSs' processing of NNSs' speech was not substantiated in their study. Equally, it was not substantiated in their analyses that familiarity with nonnative speech in general is a contributing factor to comprehensibility. Though these are logical hypotheses, they were not reliably measured in their research.

It would not be logical to presume that Gass and Varonis' findings imply that accent-familiarity includes either familiarity with all topics and knowledge or the entire lexical inventory all speakers of an accent may employ. Unfortunately, this limitation of their schematic of 'comprehension' was not included concerning how accent-familiarity

functions to benefit speech processing. This paper argues that any definition of intelligibility that includes both locutionary and illocutionary force cannot be applicable to accent-familiarity benefits.

Accent-familiarity primarily facilitates, or increases word or utterance identification success-rates, or 'intelligibility', and must be limited to only advantages determining the phonological content of speech. Context familiarity facilitates understanding or implied meaning-aspects more associated with locutionary force, or 'comprehensibility' and illocutionary force, or 'interpretability' (as they are defined by Smith and Nelson, 1985). It is unfortunate that comprehension was never clearly defined by Gass and Varonis, or how content familiarity and accent-familiarity were differentiated. Nevertheless, accent-familiarity was included as a contributing factor of facilitating comprehension was grouped alongside context-familiarity in their conclusions, and that these important differences concerning how different types of familiarity affect speech processing (intelligibility, comprehension and interpretability) were not explained more clearly.

Gass and Varonis (1984) have been cited in several papers related to, but not limited to, familiarity (Anderson-Hsieh, Johnson & Koehler 1992; Isaacs 2008, Winke, Gass & Myford, 2011, 2013) and comprehension of nonnative speakers (e.g. Pica, Holliday, Lewis, and Morgenthaler, 1989; Munro & Derwing, 1995, 1999; Derwing & Munro, 1997). It is surprising that the lack of clarity in their defining of comprehensibility and the lack of empirical evidence to support all of the claims they made has not been challenged or questioned in the existing literature. Though

the conclusions they made are perfectly logical, if not obvious, it does not excuse or render them unchallengeable. It is not the objective this paper to argue that their conclusions are false, but that what they measured and what the empirical evidence revealed is not sufficient to support the claim that familiarity with a particular nonnative accent or familiarity with nonnative speech in general contribute to comprehensibility.

Gass and Varonis may not have fully substantiated the claim that accent-familiarity contributes to non-native speech processing with significant statistical analyses, but other studies have determined that they were correct nonetheless (see Browne, 2016; Browne and Fulcher, 2016). Gass and Varonis did provide reliable evidence to suggest that familiarity of topic (or context) significantly increases comprehension, or intelligibility (depending on how the terms are defined) of NNSs' speech by NSs. Bent and Bradlow (2003) confirmed contextual clues benefit the intelligibility of NNS's by NS's.

V. Conclusion

It is highly likely that Gass and Varonis (1984) paper will continue to be cited in future research because the findings make perfect sense, regardless that not all were substantiated through the data. Again, this critical review was not written with the intention to disprove their claims but to shed light on the lack of evidence that supported some of their claims. If we as researchers do not scrutinize the measures of each other's claims or employ adequate hedging to our own conclusions the result will be the literature littered with unreliable precedents.

It would be inappropriate to infer that Gass and Varonis' application, or schematization, of

comprehensibility is an example of catachresis, or a willful example of misusing the term. Rather, it simply represents another example from the literature of a term whose meaning was in no way fixed, and they simply chose what they determined was most appropriate at its conception. They likely could not predict how comprehensibility, intelligibility and interpretability would be applied so differently and interchangeably in the future.

We researchers must consider settling on how terms like intelligibility, comprehensibility and interpretability are to be applied. Though the literature continues to recycle and create new interpretations, it seems prudent to suggest that any definition of comprehensibility should include aspects of understanding the content of a speaker's utterance. 'Understanding' in this way suggests notions of locutionary force and possibly illocutionary force that cannot be reliably measured through transcription data alone. This paper suggests word/utterance identification, or 'intelligibility' as Smith and Nelson (1985) suggested, should be considered the first, or lowest stage, in the processing of speech with 'comprehensibility' describing word/utterance meaning (locutionary force) and intended meaning of word/utterance defining 'interpretability'. If researchers can reach an accord concerning the meaning of these terms it can greatly aid future researchers by reducing the amount of cross-checking the literature that is currently required when seeking examples of related research.

References

- Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language learning*, 38 (4), 561-613.
- Bent, T. & Bradlow, A.R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, 114 (3), p.1600-1610.
- Browne, K. C. (2016). Raters' accent-familiarity levels and their effects on pronunciation scores and intelligibility on high-stakes English tests (Doctoral dissertation, School of Education).
- Browne, K., & Fulcher, G. (2016). Pronunciation and intelligibility in assessing spoken fluency. *Second Language Pronunciation Assessment*, 37.
- Catford, J.C. (1950). Intelligibility. *ELT Journal*, (1), p.7-15.
- Dauer, R.M. (2005). The Lingua Franca Core: A new model for pronunciation instruction?. *Tesol Quarterly*, 39 (3), p.543-550.
- Derwing, T.M. & Munro, M.J. (1997). Accent, intelligibility, and comprehensibility. *Studies in second language acquisition*, 19 (01), p.1-16.
- Derwing, T.M. and Munro, M.J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*., 42 (4), p.476-490.
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL quarterly*, 39 (3), p.399-423.
- Gass, S. & Varonis, E.M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language learning*, 34 (1), p.65-87.
- Isaacs, T. (2008). *Assessing second language pronunciation: A mixed methods study*. Germany: VDM Publishing.
- Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford: Oxford University Press.
- Kennedy, S. and Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review* 64, p.459-490.
- Munro, M.J. & Derwing, T.M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language*

- learning*, 45 (1), p.73-97.
- Munro, M.J., & Derwing, T.M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49 (1), p.285-310.
- Munro, M. J., Derwing, T. M., & Morton, S.L. (2006). The mutual intelligibility of L2 speech. *Studies in second language acquisition*, 28 (1), 111-131.
- Pica, T., Holliday, L., Lewis, N. & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in second language acquisition*, 11 (1), p.63-90.
- Smith, L. E., & Nelson, C. L. (1985). International intelligibility of English: Directions and resources. *World Englishes*, 4 (3), 333-342.
- Smith, L.E., & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *Tesol Quarterly*, 13 (3), p.371-380.
- Winke, P., Gass, S. & Myford, C. (2011). The relationship between raters' prior language study and the evaluation of foreign language speech samples. *ETS Research Report Series*, 2011 (2), p.i-67.
- Winke, P., Gass, S., & Myford, C. (2013). Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing*, 30 (2), p.231-25.