In Ruth Campbell's contribution, the claim for commonalities between lip-reading and other forms of language perception is even stronger since she invokes a shared mechanism, rather than just a common language, for the coding of information from different sources. Starting from the well-established recency-effect in the serial recall of lip-read lists, she first presents evidence to show that lip-reading and silent articulation (mouthing) may not depend completely on a common coding system, as a strong theory of common coding of speech perception and production would claim. On the other hand, she then demonstrates a similar recency effect for pseudo-homophones of numbers (which is much weaker than for lip-read material, but stronger than for orthographically correct number words). On the basis of this and other evidence she argues that the underlying mechanism could be a shared coding system that translates various forms of linguistic inputs into a common abstract phonological (and still strictly pre-lexical) code. This common code can be accessed by heard, by lip-read, and, under conditions where sublexical phonological assembly is required, also by written input information.

6 Integrating multiple sources of information in listening and reading

Dominic W. Massaro

Abstract

Reading and listening are viewed as having analogous stages of language processing. In both domains, there are multiple sources of information supporting the identification and interpretation of the language input. The results from a wide variety of experiments are described within the framework of a fuzzy logical model. The assumptions central to the model are (1) each source of information is evaluated to give the degree to which that source specifies various alternatives, (2) the sources of information are evaluated independently of one another, (3) the sources are integrated to provide an overall degree of support for each alternative, and (4) perceptual identification and interpretation follows the relative degree of support among the alternatives. A formalization of these assumptions is applied to results of experiments studying (1) speech perception by ear and eye, (2) audible features in listening, (3) visible features in reading, (4) orthographic context in reading, (5) phonological context in listening, (6) lexical context, and (7) syntactic and semantic context. The adequate description of the results by the model indicates that the sources of support provide continuous rather than categorical information. The integration of the multiple sources results in the least ambiguous sources having the most impact on processing. The interpretation of these results provides major constraints to be met by potential theories of language processing.

Introduction

This paper focusses on the domains of reading and listening with regard to the question of common processes. Before beginning, it is necessary to distinguish between alternative meanings of common processes and the

LANGUAGE PERCEPTION AND PRODUCTION
ISBN 0-12-052750-2

meaning intended here. At a fundamental level common processes could imply identical processes or analogous processes. No one would make the claim that there are identical processes in speech perception and reading, certainly with regard to early stages of processing of the input. With regard to early stages then one can at most expect to uncover analogous processes in the two domains whereas identical processes may be functional at later stages of information processing. Thus, for example, we might look for analogous processes in processing segmental information in the sense that auditory segments of speech and visual segments of writing must be processed by the appropriate perceptual systems (Massaro, 1975a, 1975b). Given that speech and writing represent the same language, however, identical processes may begin to play a role once the more concrete characteristics of the language input are resolved. If it turns out that we have a single lexicon functional across these two domains, we can expect that phrase and sentence interpretation will behave similarly. Thus it is primarily an empirical question regarding first whether there are analogous processes functional in the early stages of information processing in listening and reading and secondly at what stages of processing do identical processes handle the comprehension of spoken and written messages.

Many processing domains can be viewed as providing multiple sources of information for appropriate perception and action. These domains include the visual perception of a three-dimensional world, the recognition of objects and concepts, and social judgement of people and actions. The goal of the present chapter is to consider evidence for the idea that a fundamental property of speech perception and reading is also the availability of multiple sources of information. Sources of information are defined as the features or attributes of the environment that are utilized in perception and action (Massaro, in press a d). The belief is that there are multiple sources of information available in language and a fundamental process is the evaluation and integration of these sources. Following the idea of analogous rather than common processes, we can expect that the sources of information available in the two domains will differ to a considerable degree. Some top-down sources might be similar, however, in the sense that they are common to both speech perception and reading. However, there can be no argument with the idea that the fundamental bottom-up sources of support in the two domains must be different. The question of interest with respect to the idea of analogous processes is whether the same procedures govern the evaluation and integration of the multiple sources across the two domains. This distinction might be considered in light of Fodor's (1983) concept of modularity. Even if we have different modules for reading and listening, these modules might share or have analogous processes.

Perhaps it is a truism that multiple sources of information are available to the language perceiver and yet few studies actually have addressed the important implications of this observation. These include the number and nature of the sources, how each source is processed, how the multiple sources are selected or integrated, and the nature of categorization resulting from these processes. I will first illustrate the approach I have taken to the integration issue by discussing audible and visible contributions to speech perception. The same approach will be extended to account for the range of domains involving language processing given multiple sources of information.

Although the visible mouth movements of the speaker are potentially informative and although perceivers have been shown to utilize some of this information given no or degraded auditory input, only within the last decade was it demonstrated that watching a speaker actually influences what the perceiver hears (McGurk and MacDonald, 1976). This phenomenon offers a potentially valuable paradigm for addressing the issues revolving around the conceptualization of multiple sources of information supporting language understanding. One reason is that it is easy to manipulate the audible and visible sources independently of one another. A second reason is the naturalness of the identification task given either just auditory or just visual or both sources of information. The goal of the experiment reported here is to demonstrate conclusively that the two sources are integrated in speech perception and to uncover the nature of the processes involved. Bimodal speech events were created by the combination of synthetic speech sounds along an auditory /ba/ to /da/ continuum paired with /ba/ or /da/ visual articulations. The single auditory and single visual cue conditions also were included. The analysis of the experiment provides a framework for conceptualizing this study and the description of more general problems in reading and listening. This framework will be applied to a variety of domains involving a comprehensive range of sources of information in language processing.

Integrating audible and visible speech

Method

Eight college students participated for one hour in the experiment. All test stimuli were recorded on color videotape, following the procedure of Massaro and Cohen (1983c). On each trial, a male speaker articulated /ba/, /da/, or nothing. The original audio track was replaced with synthetic speech consisting of a nine-step /ba/ to /da/ auditory continuum. Nine levels along the auditory /ba/ to /da/ continuum were factorially combined with two possible visual articulations, /ba/ and /da/. These 18 trials represent the bimodal condition. There were also auditory-alone and visual-alone conditions. In the auditory-alone condition one of the nine auditory stimuli was presented, but the speaker did not move his mouth. In the visual-alone condition the speaker's mouth articulated either /ba/ or /da/, but no auditory speech was presented. In every block of 54 trials, there were 18 bimodal conditions, 18 auditory-alone conditions, and 18 visual-alone conditions. Each subject was tested for a total 594 experimental trials.

Subjects were instructed to watch the speaker on the TV monitor and to listen to a possible speech sound coming from the TV. They were told of the three different kinds of trials: the bimodal trials, the auditory-alone trials, and the visual-alone trials. Subjects were asked to identify the speech event as one of eight alternatives by pressing one of eight buttons on the keyboard in front of them. The eight buttons corresponded to the alternatives /ba/, /da/, /bda/, /dba/, /ga/, /va/, /tha/, and 'other'. These alternatives were determined on the basis of pilot work and a previous experiment (Massaro and Cohen, 1983c). Subjects were instructed to use the identification category 'other' when none of the other seven response alternatives was appropriate. Subjects were told that although they might not be sure about what the speaker said, they should simply make the best decision that they could. Each trial was preceded by an auditory warning stimulus with a response interval of 2.75 s. Each subject was seated and viewed a 12 in color monitor, which presented both the video and audio. The subjects sat about two to three feet away from the TV with the loudness level of the speech at a comfortable listening level (70 dB-A).

Results

Figure 1 presents the observed proportion of each of the eight responses for the 29 unique speech events. Three separate analyses of variance were carried out across the three conditions. In all three analyses, the variables of interest were statistically significant. Subjects' judgements, as indicated by the proportion of responses, were influenced by the both visual and the auditory information. The interaction of the visual and auditory information with response was also significant in the bimodal condition.

The single modality conditions establish the contribution of both audible and visible sources, but it is more involved to determine that both sources simultaneously contribute to bimodal speech perception. Significant effects of both variables would occur if some subjects used one modality and other subjects used the other modality, or if a given subject used one modality on some trials and the other modality on other trials. On the other hand, the

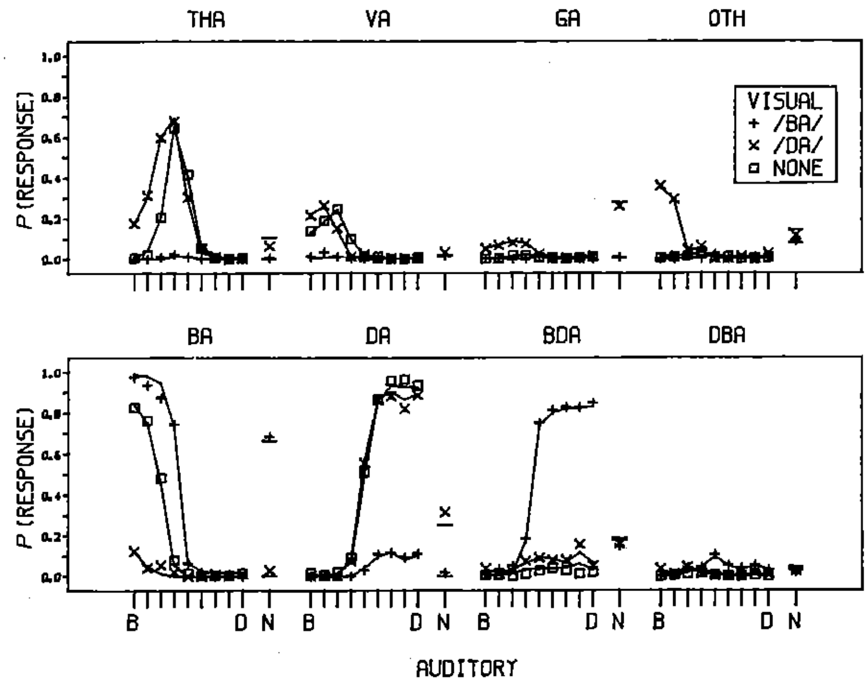

Figure 1. Observed (points) and predicted (lines) proportion of identifications for the eight response alternatives as a function of the auditory and visual dimensions for the auditory, visual, and bimodal syllables. The *x* abscissa gives the 9 levels along the auditory /ba/ to /da/ continuum and the N condition in which no auditory syllable was presented. The parameter of the graph gives whether the visual syllable was /ba/, /da/, or was not presented. The predictions are for the fuzzy-logical model of perception.

simultaneous contribution (integration) of both modalities can be established if the bimodal judgements can not be explained by a given subject using just a single dimension on each trial. Consider the three circled points in the right panel of Figure 2 for a typical subject. A visual /ba/ presented alone is identified as /ba/ about 90% of the time. The speech sound at the /ba/ end of the auditory continuum is heard as /ba/ about 82% of the time when presented alone. However, the same speech sound at the /ba/ end of the auditory continuum paired with a visual /ba/ is identified as /ba/ 100% of the time. Similar results exist for the other subjects as can be seen from the group results in the left panel of Figure 2. Thus, the visual

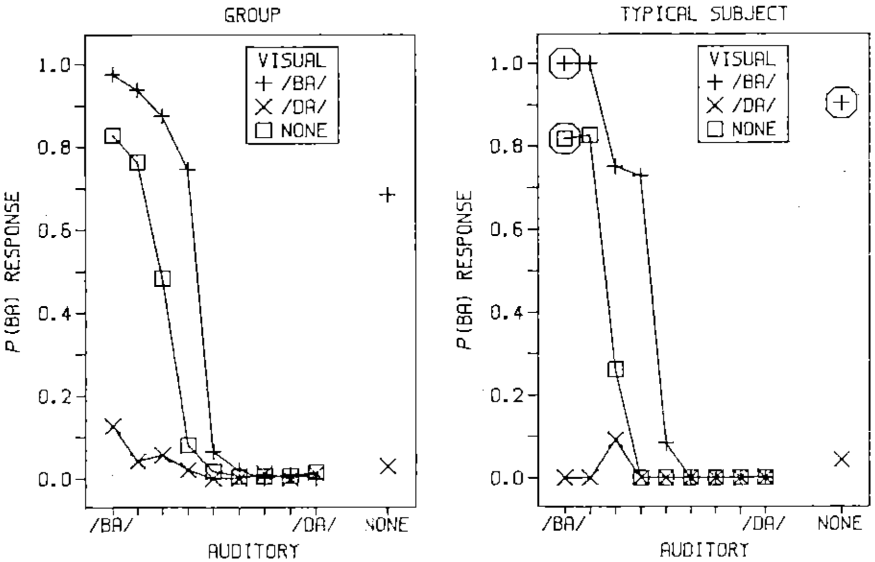

Figure 2. Proportion of /ba/ judgements for the group (left panel) and a typical subject (right panel) to the auditory, visual and bimodal speech events.

/ba/ contributes to the /ba/-ness of a syllable with an auditory /ba/ more than can be predicted by nonintegration.

The results of the experiment with open-ended response alternatives provide strong evidence for a true integration of the auditory and visual sources. In addition, the judgements given conflicting sources of information reveal a smart perceptual process. Given a visual /ba/ and an auditory /da/, subjects often reported /bda/. In contrast, a visual /da/ paired with an auditory /ba/ seldom produced /dba/ judgements. To explain the /bda/ judgements, the auditory and visual components must have been processed to a fairly deep level relatively independently of one another (i.e., without crosstalk). The process must also involve some assessment of the auditory and visual sources in terms of their compatibility with candidate categories. The alternative /bda/ given visual /ba/ and auditory /da/ is reasonable whereas the alternative /dba/ given visual /da/ and auditory /ba/ is not. A visual /ba/ is similar to a visual /bda/ whereas a visual /da/ is not at all similar to a visual /dba/. Thus, what seems as an unintuitive result makes good sense when we consider how well the two dimensions fit potential identification categories.

What model best describes the complex pattern of results? Two general classes of models can be contrasted on the basis of how the auditory and visual sources of information are evaluated and integrated in perceptual

recognition. Before addressing this question, it is important to claim that the identification judgements represent direct indices of the perceptual experience of the speech event, not simply some *post hoc* judgement. Supporting evidence for this claim has been found using continuous rating judgements and reaction times (Massaro, in press b c d). One class of models assumes that the auditory and visual sources of information are categorized before integration, whereas the contrasting class assumes that categorization does not occur until integration is complete. These classes have been called categorical and continuous, respectively (Massaro, in press a b c; Massaro and Cohen, 1983 a b c). To test the alternatives against the present results, the categorical model formalized by Massaro and Cohen (1983c) was contrasted with a continuous model called a fuzzy logical model of perception (Oden and Massaro, 1978). It is reasonable to contrast these two models since there is no other testable categorical model in the literature and the fuzzy logical model has not yet been tested against an expanded factorial design with open-ended response alternatives. Furthermore, the two models require a similar number of free parameters to describe the results. Because of space limitations, the reader is referred to the other published descriptions of the models for their details (Massaro, in press a b c d; Massaro and Cohen, 1983 a b c).

The two models were fit to the results of each of the individual subjects using the program STEPIT (Chandler, 1969). The predictions of a model are determined by finding those parameter values that optimize the description of the results. Figure 1 gives the average predicted results for the fuzzy logical model. The root mean squared deviation (RMSD) between the predicted and observed values varied between 0.022 and 0.041 across the eight subjects with an average value of 0.030. This good description is very impressive since it is predicting an essentially open ended set of response alternatives with identical information for the unimodal and bimodal speech stimuli. Figure 3 gives the average predictions for the categorical model. The RMSD varied between 0.124 and 0.173 with an average value of 0.148. The results of each of the eight subjects were better described by the fuzzy logical model than the categorical model.

The evaluation and integration of multiple sources of information is proposed to be central in language processing, as it was in speech perception by ear and eye. The second section of the paper reviews a variety of domains in listening and reading supporting this proposal. In all cases, the fuzzy logical model provides a good description of the results whereas the categorical model would fail miserably. Given the constraints of the present chapter and the large number of different domains, the details of the experiments and the model fits can not be described fully. In many cases, however, this information is available in the literature. The goal of the

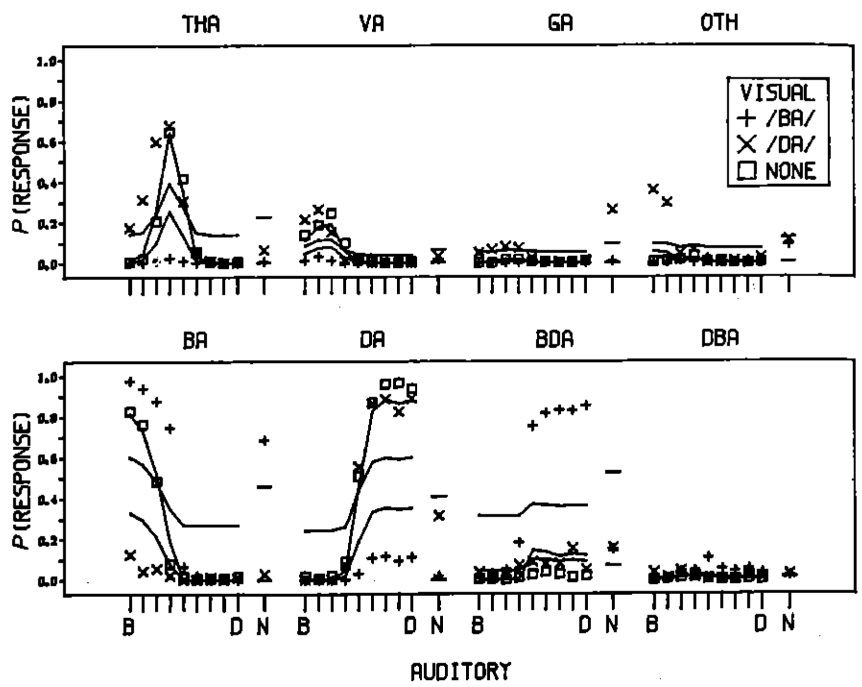

Figure 3. Observed (points) and predicted (lines) proportion of identifications for the eight response alternatives as a function of the auditory and visual dimensions for the auditory, visual, and bimodal syllables. The *x* abscissa gives the nine levels along the auditory /ba/ to /da/ continuum and the N condition in which no auditory syllable was presented. The parameter of the graph gives whether the visual syllable was /ba/, /da/, or was not presented. The predictions are for the categorical model of perception.

present summary is to highlight the generality of the principle of the important role of evaluating and integrating multiple and continuous sources of information in language processing.

Integrating audible features

One argument for the integration of multiple sources of auditory information in speech perception comes from the discovery of many different cues or features that contribute to the discriminable contrasts found in speech.

The perceived distinction between voiced and voiceless stop consonants in medial position such as /aga/ and /aka/ can be influenced by the preceding vowel duration, the silent closure interval, the voice-onset time, and the onset frequency of the fundamental (Massaro and Cohen, 1983b; Port and Dalby, 1982). The fuzzy logical model has survived tests across a wide variety of speech contrasts and acoustic features (Massaro, (1979b, 1984b, in press a; Massaro and Cohen, 1976; Massaro, Cohen and Tseng, 1985; Massaro and Oden, 1980 a b).

Integrating visual features

The study of letter recognition in reading has a different tradition from the study of speech recognition. Few studies have actually manipulated the signal in systematic ways and observed the consequences for recognition (Naus and Shillman, 1976). Oden (1979) was the first to use a factorial design to independently manipulate two characteristics of a letter. Rating judgements of these test items were well described by the fuzzy logical model. Consider a recent experiment carried out by Massaro and Hary

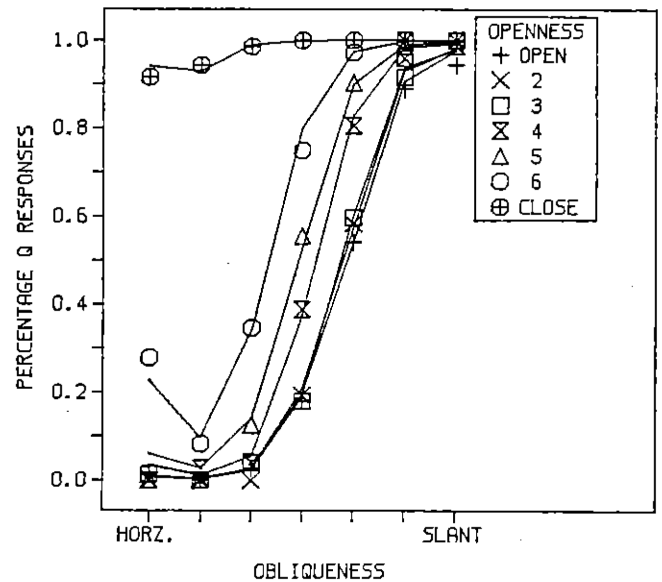

Figure 4. Observed (points) and predicted (lines) percentage of Q identifications as a function of the openness of the gap in the oval and the obliqueness of the straight line (from Massaro and Hary, submitted).

(in press). The uppercase letters Q and G served as test alternatives of two features varying in terms of the amount of gap on the right of the oval and the angle of the straight line. On each trial, the subject viewed a short presentation of one of the letters and indicated whether a G or a Q was presented. Figure 4 gives the proportion of Q responses as a function of the amount of gap in the oval and the angle of the straight line. Both variables influenced the judgements in the expected direction and the significant interaction between the two variables also reflect the larger contribution of one feature when the other feature is more ambiguous. The lines represent the predictions of the fuzzy logical model, which gave a good description of the results, the RMSD varied between 0.028 and 0.097 across the nine subjects and averaged 0.049.

Integrating orthographic context in letter recognition

It is well known that the orthographic (spelling) context of a letter influences its recognition. In the present framework, higher-order constraints operate independently of featural information about the letters. That is, featural analyses of letters are not modified by higher-order constraints. According to this view, the feature evaluation process makes available the same information independently of the higher-order constraints that are present.

It is possible to gradually transform the lowercase letter *c* into an *e* by extending the horizontal bar. To the extent the bar is long, there is good visual information for an *e* and poor visual information for a *c*. Now consider the letter presented as the first letter in the context *-oin* and the context *-dit*. Only *c* is orthographically admissible in the first context since the three consecutive vowels *eoi* violate English orthography. Only *e* is admissible in the second context since the initial cluster *cd* is an inadmissible English pattern. In this case, the context *-oin* favors *c*, whereas the context *-dit* favors *e*. The context *-tsa* and *-ast* can be considered to favor neither *e* or *c*. The first remains an inadmissible context whether *e* or *c* is present, and the second is orthographically admissible for both *e* and *c*.

The experiment factorially combined six levels of visual information with these four levels of orthographic context (Massaro, 1979a). The test string was presented for a short duration followed after one of four intervals by a masking stimulus composed of random letter features. The test letter also was presented at each of the four letter positions in each of the four contexts. The subject was asked to indicate whether an *e* or *c* was present in the test display. Subjects were instructed to make the best choice on the basis of what they saw.

Figure 5 gives the observed interaction of bar length and orthographic context across the four masking intervals, along with the predictions of the fuzzy logical model. The probability of an *e* response increased with increases in the bar length of the critical letter. The results also show a gradual resolution of the critical letter with increases in processing time before onset of the mask since the curves across bar length are steeper with longer masking intervals. An *e* identification was more probable for the context in which *e* but not *c* was orthographically admissible than for the context in which *c* but not *e* was admissible. The neutral contexts were intermediate. This context effect was larger at the more ambiguous levels of the bar length of the test letter. Context also had a larger impact on identification of the test letter at the extremes of the visual continuum to the extent the masking interval was short. These results are consistent with

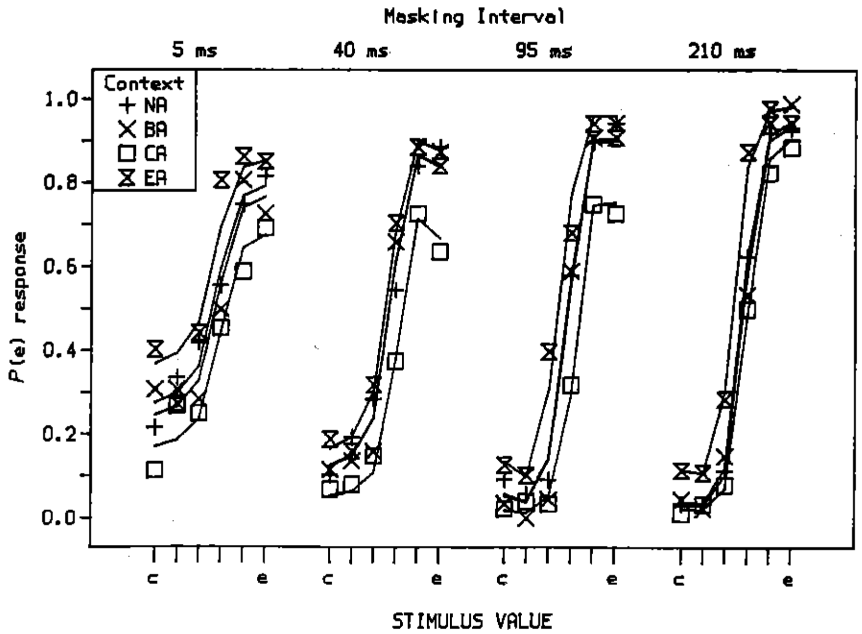

Figure 5. Observed (points) and predicted (lines) probability of an e identification response as a function of the bar length of the test letter (stimulus value), the orthographic context, and the masking interval before the onset of the mask. The context NA = neither *e* nor *c* admissible, BA = both admissible, CA = *c* admissible, and EA = *e* admissible (results from Massaro, 1979b).

the general premise that the contribution of context is larger with ambiguous relative to unambiguous bottom-up sources of information. The fuzzy logical model gave a good description of the average results with an RMSD of 0.039.

Integrating phonological context in syllable recognition

Massaro and Cohen (1983d) assessed how the information from the acoustic signal is combined or integrated with information from phonological constraints in English. Phonological constraints refer to the fact that languages are redundant in terms of the possible sequences of speech sounds. Listeners were asked to identify sounds along a continuum between /li/ and /ri/, which was made by varying the starting frequency of the third formant (F3) transition. These sounds were placed after each of four initial consonants: /t/, /s/, /p/, and /v/. When the sounds are placed after the initial consonant /s/, for example, /l/ is phonologically admissible in English, but /r/ is not. If phonological constraints influence perception, listeners should tend to hear /l/ following the sound /s/. Figure 6 gives the

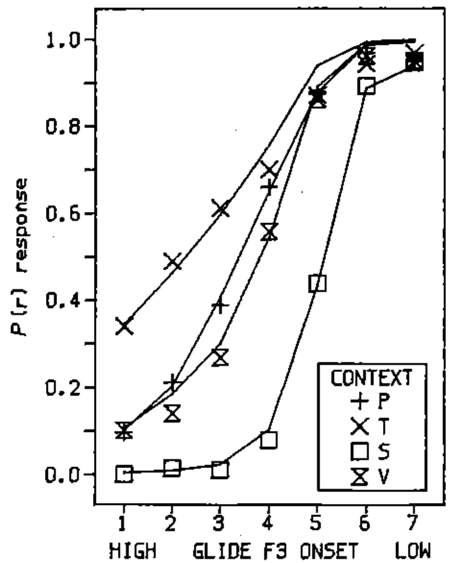

Figure 6. Observed (points) and predicted (lines) probability of an /r/ identification as a function of the F3 transition onset during the glide; the initial consonant is the curve parameter (from Massaro and Cohen, 1983d).

probability of /r/ identifications as a function of both F3 of the liquid and the initial consonant. As expected, the proportion of /r/ identifications increased with decreases in the starting value of F3 of the liquid. More importantly, an /r/ identification was more likely in the context of /t/ than in the context of /p/ or /v/ and least likely in the context of /s/. The effects of the initial context consonant were largest at the more ambiguous values of the liquid. The current model provides a good description of the results, the RMSD varying between 0.025 and 0.116 across the seven subjects and with an average RMSD of 0.055.

Integrating lexical information

Lexical constraints also might be expected to influence processing of written and spoken language. Ganong (1980) assessed the contribution of lexical context to the perception of stop consonants. The voice-onset time of an initial stop consonant was varied to create a continuum from a voiced to voiceless sound. The following context was varied so that either the voice or the voiceless stop would make a word. For example, subjects identified the initial stop as /d/ or /t/ with the following context *ash* (where /d/ makes a word and /t/ does not). Voiced (/d/) responses were more frequent when /d/ made a word than when /t/ made a word. The contribution of lexical context was largest at the more ambiguous levels of voice-onset time. These results have been described quantitatively by the fuzzy logical model with the standard assumption that acoustic featural information and lexical context make independent contributions to perceptual recognition (Massaro and Oden, 1980b).

Integrating sentential context in visible word recognition

One of the original studies of the combination of sentential context and stimulus information in word recognition were performed by Tulving, Mandler and Baumal (1964). Eight exposure durations were factorially combined with four sentential context lengths in a word recognition task. A tachistoscopic presentation of a test word followed the reading of the sentence context. Performance improved with increases in word duration and with sentential context. The outcome shows a larger contribution of sentential context when the exposure duration is intermediate and performance is neither very poor or very good. This result indicates that context is most effective when subjects have some but not relatively complete featural information about the test word. The fuzzy logical model provides a reasonably good description of the results, with a RMSD of 0.021 (Massaro, 1984a).

Integrating sentential context in audible word recognition

Analogous to the question of the utilization of context effects in reading, we can ask whether there is a similar contribution of sentential context in listening. A positive answer has been repeatedly found since Miller, Heise and Lichten's (1951) seminal study, but little quantitative work has been aimed at assessing how sentential context is integrated with the sensory information. In the present framework, sentential context provides an independent source of information made available to the word recognition process in perceiving continuous speech (Massaro, 1979b, 1984a, b; Massaro and Oden, 1980b).

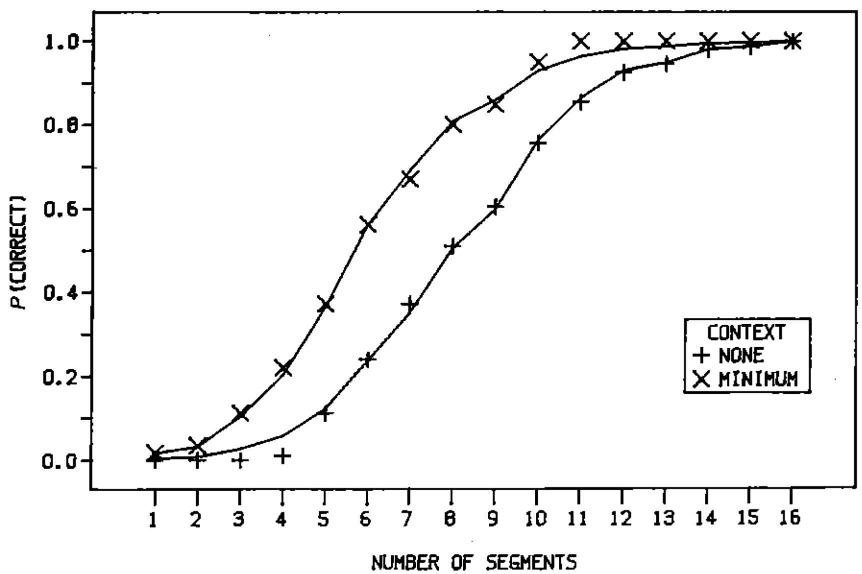

Figure 7. Observed (points) and predicted (lines) probability of identifying the test word correctly as a function of the nature of the sentential context and the number of segments of the test word. The Minimum context refers to minimum semantic and weak syntactic constraints. The None context refers to no semantic and weak syntactic constraints (after Tyler and Wessels, 1983).

Tyler and Wessels (1983) used the gating paradigm to assess the contribution of various forms of sentential context to word recognition. Subjects heard a sentence followed by the beginning of the test word (with the rest of the word gated out). The word was increased in duration by adding small segments of the word until correct recognition was achieved. The sentence contexts varied in syntactic and semantic constraints. Some sentence contexts had minimal semantic constraints in that the target word was not predictable in a cloze test given the sentence context and the first 100 ms of the target word. A positive effect of sentence context in this situation would be very impressive since it would illustrate a true integration of top-down and bottom-up information. That is, neither the context nor limited stimulus information would lead to word recognition, but when presented jointly word recognition is very good. Although the authors failed to report the actual proportion of times the target word could be guessed from the context presented alone, we can assume it was close to zero. Thus, the strong effect of minimum semantic context illustrated in Figure 7 can be considered to reflect true integration of top-down and bottom-up sources of information. Figure 7 also gives the predictions of the fuzzy logical model. The model captures the exact form of the integration of the two sources of information with a RMSD of 0.015.

Integrating syntactic, semantic and prosodic information in sentence interpretation

In the tasks described to this point, the language user had to perceive various surface forms such as identity of a particular word. The language user also faces the challenge of interpreting these surface forms in many instances (Oden, 1977). Presented with the statement, *the horse the bucket kicked*, it is necessary to determine whether the horse or the bucket did the kicking. Analogous to the description of determining the identity of spoken or written forms, the determination of interpretations should result from the evaluation and integration of multiple sources of information. These sources include syntactic variables such as word order, semantic variables such as animation, and prosodic information such as stress.

In the MacWhinney, Bates and Kliegl (1984) task, subjects are asked to interpret three-word phrases in terms of 'which one of the two nouns in the sentence is (the subject of the sentence), that is, (the one who does the action).' Thus, the task of the subject is to choose between the two noun alternatives based on the semantic/syntactic cues available. The authors manipulated word order, animacy, noun-verb agreement and stress in a factorial design. There were three possible word orders of the two nouns

(N) and verb (V) (NVN, NNV, and VNN); three levels of animacy (both nouns animate, the first noun animate and the second inanimate, and the first noun inanimate and the second animate); the verb would agree in number with the first noun, with both nouns, or with the second noun of the sentence; and three levels of stress (no stress, stress on first noun, and stress on second noun). Thus, there were 81 possible sentences and each subject judged each type of sentence only once. The dependent measure is the proportion of subjects choosing the first noun as subject/agent for each sentence type.

The application of the current model is relatively straightforward. Four independent sources of information are assumed to influence the choice between the first and the second noun as the subject or agent. These two

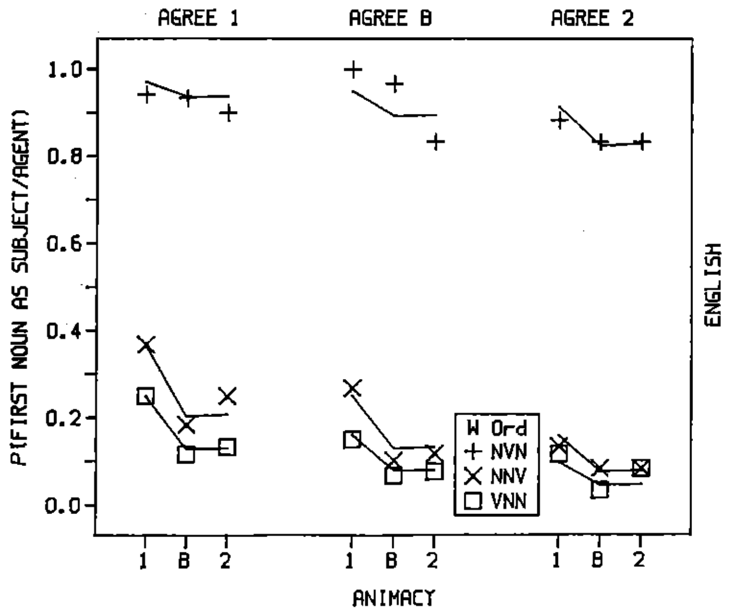

Figure 8. Observed (points) and predicted (lines) probability of identifying the first noun as subject/agent as a function of the animacy of the two nouns, (1 = first noun animate, B = both nouns animate, 2 = second noun animate) the word order, and the noun–verb agreement (Agree 1 = first noun and verb agree in number, Agree B = both nouns and verb agree in number, Agree 2 = second noun agrees in number) for the English listeners (after MacWhinney *et al.*, 1984).

interpretations can be defined as a function of the four sources. The second noun can be viewed as the subject or agent to the extent each of four sources are inappropriate for the first noun being subject or agent. Twelve free parameters are necessary to predict the 81 data points corresponding to the 81 unique sentence types.

Two forms of the model were fit to the results for the English subjects. In the first, four sources were assumed, whereas the second assumed no contribution of stress. The RMSD values were 0.044 and 0.046, respectively, showing that stress as manipulated in the study had very little, if any, cue value. This result is relevant to our definition of source of information. As manipulated in the MacWhinney *et al.* (1984) task, it can be concluded that stress does *not* function as a source of information in sentence interpretation for English listeners. Figure 8 gives the observed and predicted results for the 27 sentence types pooled over stress. The English subjects were influenced mostly by word order, but also by animacy and noun-verb agreement. These three variables can be considered to function as sources of information in sentence interpretation.

Concluding statement

The fuzzy logical model of information integration describes the results in a wide variety of domains of language processing, perhaps all of the domains studied to date. The general approach not only offers a theoretical framework, it provides an empirical technique for assessing potential sources of information supporting language processing. A remaining challenge for the present approach is whether similar processes might be functional in speaking and writing. Rather than evaluating and integrating multiple sources of information to achieve perceptual recognition or interpretation, the production of language involves conveying a message via multiple sources of information. To what extent analogous processes are functional across perception and production remains an unanswered question.

Acknowledgement

The writing of this paper and the author's research reported in the paper were supported, in part, by NINCDS Grant 20314 from the Public Health Service and Grant BNS-83-15192 from the National Science Foundation. The author gratefully acknowledges Brian MacWhinney for providing the results of their experiments.

References

Chandler, J. P. (1969). Subroutine STEPIT — Finds local minima of a smooth function of several parameters. *Behavioral Science*, **14**, 81–2.
Fodor, J. A. (1983). *Modularity of Mind*. Cambridge, MA: Bradford Books.
Ganong, W. F. III. (1980). Phonetic categorization in auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 110–25.
MacWhinney, B., Bates, E., and Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, **23**, 127–50.
Massaro, D. W. (1975a). *Experimental Psychology and Information Processing*. Chicago, IL: Rand-McNally.
Massaro, D. W. (Ed.) (1975b). *Understanding Language: An Information Processing Analysis of Speech Perception, Reading, and Psycholinguistics*. New York: Academic Press.
Massaro, D. W. (1979a). Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 595–609.
Massaro, D. W. (1979b). Reading and listening (Tutorial paper). In P. A. Kolers, M. Wrolstad, and H. Bouma (Eds) *Processing of Visible Language*, Vol 1. New York: Plenum, pp. 331–54.
Massaro, D. W. (1984a). Building and testing models of reading processes. In P. D. Pearson (Ed.) *Handbook of Reading Research*. New York: Longman, pp. 111–46.
Massaro, D. W. (1984b). Time's role for information, processing, and normalization. *Annals of the New York Academy of Sciences, Timing and Time Perception*, **423**, 372–84.
Massaro, D. W. (in press a). A fuzzy logical model of speech perception. In W. A. Lea (Ed.) *Towards Robustness in Speech Recognition*. Apple Valley, MN: Speech Science Publications.
Massaro, C. W. (in press b). Categorical partition: A fuzzy logical model of categorization behavior. In S. Harnad (Ed.) *Categorical Perception*, New York.
Massaro, D. W. (in press c). Information-processing theory and strong inference: A paradigm for psychological inquiry. In H. Heuer and A. F. Sanders (Eds) *Perspectives on Perception and Action*. Hillsdale, NJ: Erlbaum.
Massaro, D. W. (in press d). Speech perception by ear and eye. In B. Dodd and R. Campbell (Eds) *Hearing by Eye: Experimental studies in the psychology of lipreading*. Hillsdale, NJ: Erlbaum.
Massaro, D. W., and Cohen, M. M. (1976). The contribution of fundamental frequency and voice onset time to the /zi/-/si/ distinction. *Journal of the Acoustical Society of America*, **60**, 704–17.
Massaro, D. W., and Cohen, M. M. (1983a). Categorical or continuous speech perception: A new test, *Speech Communication*, **2**, 15–35.
Massaro, D. W., and Cohen, M. M. (1983b). Consonant-vowel ratio: An improbable cue in speech. *Perception and Psychophysics*, **33**, 501–5.
Massaro, D. W., and Cohen, M. M. (1983c). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 753–71.

Massaro, D. W. and Cohen, M. M. (1983d). Phonological context in speech perception. *Perception and Psychophysics*, **34**, 338–48.
Massaro, D. W., Cohen, M. M., and Tseng, C-Y. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, **13**, 267–89.
Massaro, D. W., and Hary, J. M. (in press). Addressing issues in letter recognition. Psychological Research.
Massaro, D. W., and Oden, G. C. (1980a). Evaluation and integration of acoustic features in speech perception. *Journal of the Acoustical Society of America*, **67**, 996–1013.
Massaro, D. W., and Oden, G. C. (1980b). Speech perception: A framework for research and theory. In N. J. Lass (Ed.) *Speech and Language: Advances in Basic Research and Practice*. Vol 3, New York: Academic Press, pp. 129–65.
McGurk, H., and MacDonald (1976). Hearing lips and seeing voices. *Nature*, **264**, 746–8.
Miller, G. A., Heise, G. A., and Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, **41**, 329–35.
Naus, M. J., and Shillman, R. J. (1976). Why a Y is not a V: A new look at the distinctive features of letters. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 394–400.
Oden, G. C. (1977). Integration of fuzzy logical information. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 565–75.
Oden, G. C. (1979). A fuzzy logical model of letter identification. *Journal of Experimental Psychology: Human Perception and Performance*. **5**, 336–52.
Oden, G. C., and Massaro, D. W. (1978). Integration of featural formation in speech perception. *Psychological Review*, **85**, 172–91.
Port, R. F., and Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception and Psychophysics*, **32**, 141–52.
Tulving, E., Mandler, G., and Baumal, R. (1964). Interactions of two sources of infirmation in tachistoscopic word recognition. *Canadian Journal of Psychology*, **18**, 62–71.
Tyler, L. K., and Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception and Psychophysics*, **34**, 409–20.