11 Dialect Intelligibility

CHARLOTTE GOOSKENS

11.1 Introduction

The present chapter focuses on the communicative consequences of dialectal variation. One of the main functions of language is to enable communication, not only between speakers of the same variety but also between people using different accents, dialects or closely related languages. Most research on dialect intelligibility is relatively recent, especially when it comes to actual testing. The methods for testing and measuring are getting more and more sophisticated, including web-based experiments that allow for collecting large amounts of data, opening up new approaches to the subject.

Some of the first to develop a methodology to test dialect intelligibility were American structuralists (e.g., Hickerson, Turner, and Hickerson 1952; Pierce 1952; Voegelin and Harris 1951), who tried to establish mutual intelligibility among related indigenous American languages around the middle of the previous century. They used the so-called recorded text testing (RTT) method. This methodology has been standardized and is still being used, for example in the context of literacy programs where a single orthography has to be developed that serves multiple closely-related language varieties (Casad 1974; Nahhas 2006). Since then, numerous intelligibility investigations have been carried out with various aims, for instance to resolve issues that concern language planning and policies, second-language learning, and language contact. Data about distances between varieties and detailed knowledge about intelligibility can also be important in sociolinguistic studies. Varieties that have strong social stigma attached to them could unfairly be deemed hard to understand (Giles and Niedzielski 1998; Wolff 1959). The relationship between attitudes and intelligibility is not a straightforward one, but advances in the field of intelligibility testing provide sociolinguists with objective data to help to resolve conflicts that arise concerning non-standard varieties. Knowledge about mutual intelligibility is also needed for standardization and development of new orthographies in communities where no standardized orthography exists.

In this chapter we will first deal at some length with two questions concerning dialect intelligibility, which we consider of major importance to linguistics in general and to dialectology in particular. We will look into the role of intelligibility in the definition of "dialect" and "language," and examine how intelligibility can be used to validate research on distances between dialects. An overview of factors that determine intelligibility is provided, and finally, some gaps in our knowledge of processes and phenomena in the area of dialect intelligibility are signalled, and desiderata for future research are formulated.

11.2   Definition of "Dialect" and "Language"

The question of how to define a "language" as opposed to a "dialect" is one of the oldest and most central questions that linguists have asked themselves. Reasons for wanting to distinguish between the two concepts have sometimes been theoretical but more often practical or political in nature. For example, people might want to know how many languages there are in the world, and to be able to answer this question it is necessary to be able to define when a variety (or a group of varieties) can be considered a language in its own right. Also, it has been important for language planning and policies at both the national and the more global levels to find criteria that define a language variety as a language. A language often represents a community and is tightly connected to standardization processes and development of new orthographies. For people fighting for the rights of a language variety it is of great importance that the variety is recognized as a language rather than a dialect. Official recognition will give the variety a stronger position. This becomes clear from part 1 of the European Charter for Regional Minority Languages (1992). The right to use one's variety in public life (e.g., in educational, juridical, administrative, or media contexts) is dependent upon the status of this variety as an official, regional or minority language. Speakers of varieties that are classified as "dialects" do not have these rights under the Charter. On the other hand, there are also areas in which languages that are clearly different are characterized as dialects of a single language because it is desirable to preserve the unity of the area.

Discourse on language rights and standardization of language varieties rests on an underlying assumption that one can somehow objectively identify which varieties are languages and which are dialects. Kloss (1967) introduced the terms *Ausbausprache* (language by development) and *Abstandsprache* (language by distance) for analyzing and categorizing language varieties that are closely related and often are used within the same society. An *Ausbausprache* typically has its own standardized form used autonomously with respect to other languages, it is often taught in schools, and it is used as a written language for a wide variety of social and political functions, possibly including that of an official national language. In an *Ausbau* definition, "languages" and "dialects" are social constructs definable only in terms of their socio-political and cultural status and breadth of use, and they have little to do with independently identifiable structural entities. An often-mentioned example of *Ausbau* languages are the Scandinavian languages (Danish, Norwegian, and Swedish) that are so closely related that the speakers of the three languages can, with some effort, understand each other. Still, they are regarded as different languages because they have distinct, codified, standardized forms, with their own orthographies, grammar books, and literatures, and are spoken in three separate nation states. Another example is Serbian and Croatian. These languages may be even closer to each other than the Scandinavian languages are, but speakers of "Serbo-Croatian" insist that they are two different languages. In fact, the differences are mainly lexical, with very few differences in pronunciation and grammar. The perceived differences are reinforced by other factors such as different alphabets, religions and ethnicity. Tamburelli (2014) points out that the *Ausbau* definition may lead to a circularity effect: the use of a linguistic variety in educational, juridical, administrative, or media contexts is a right reserved for varieties with language status, but to achieve the "language" label a variety has to have a certain socio-political status. This means that the language varieties that language legislation is meant to protect may in fact be excluded *a priori* from this protection.

Linguists in general prefer to define languages as *Abstandsprachen*. In this view, one language variety is called an *Abstandsprache* with respect to another language variety if the two are so different from each other that they are in fact different languages. Kloss (1967) left unspecified exactly how the differences between two language varieties are to be measured objectively, presumably because he lacked the tools to do so. Methods to measure linguistic

distances objectively have now been developed by dialectometrists (see Chapter 7 of this volume). The problem remains, however, that languages do not differ along just one dimension, but may differ to different extents in their lexicon, phonetics and phonology, morphology and syntax. It is not clear how much weight should be given to these separate linguistic dimensions when assessing overall distance. We will return to this point in Section 11.4.

Maybe to circumvent the problem of how to weight different linguistic dimensions, Trudgill (2000) introduced the intelligibility criterion, and this has become the standard—or at least the primary—criterion among many linguists. According to this definition, dialects are mutually intelligible varieties, whereas languages are so linguistically different that their speakers are unable to understand each other. From this it follows that a language is a collection of mutually intelligible dialects. Intelligibility was used as the main criterion by the compilers of *Ethnologue*, an online database of all the world's known languages, to decide what should count as a language (Lewis *et al.* 2013). In the seventeenth edition (2013), *Ethnologue* contains a list of 7,105 languages. It has been criticized for splitting language varieties into too many languages, but Hammarström (2005) asserts that *Ethnologue* is consistent with specialist views most of the time.

The intelligibility criterion for defining languages is not without its problems, which has caused some linguists to reject it (see Hammarström (2008) for an overview of citations). We will look further into some of these objections below.

First, intelligibility is not easy to measure. It can be tested by means of opinion testing, whereby subjects are asked to indicate (with or without speech fragments) how well they think they understand the language at hand. However, these intuitive and impressionistic judgments may be distorted by extra-linguistic factors (see below). Intelligibility may also be determined experimentally by means of functional tests, which typically express the degree of intelligibility as the percentage of input that is correctly recognized by the subject. This approach comes with its own problems. It is generally difficult to abstract away from individual speakers and choice of test. In addition, an effort must be made to avoid priming effects, ceiling effects, excessive memory load, and other unwanted issues. These considerations often make it rather time-consuming both to develop suitable tests and to carry out the tests themselves. Gooskens (2013) lists a number of methods for measuring dialect intelligibility, and discusses their advantages and disadvantages.

Both opinion testing and functional testing will result in numbers that express how well subjects can understand a language (variety). This means that intelligibility is a matter of degree. It is difficult to decide when the mutual intelligibility of two varieties is so high that they should be considered dialects of the same language rather than separate languages. Thus far, it has not been possible to define an intelligibility threshold. Hudson (1996, 35) states that "this is clearly a question which is best avoided, rather than answered, since any answer must be arbitrary." Still, a number of studies have made an attempt to establish a reference point, below which it is difficult to achieve successful communication (see Tamburelli (2014) for an overview).

The fact that intelligibility scores are a matter of degree reflects the situation in many dialect areas. Traditional dialectologists present the geographical spread of dialectal features, for instance a particular word form or pronunciation, by drawing isoglosses (lines separating features on a map; for example, Weijnen (1941) and several of the other chapters in this volume). A dialect division is said to be major if several isoglosses coincide (isogloss bundles). However, isoglosses usually only coincide approximately, resulting in different dialect areas with a transition zone in between (so-called "dialect continua"). A well-known example of a dialect continuum is found in the Dutch-German dialect area. In the early nineteenth century, one could start from the far south of the German-speaking area and travel to the far west of the Dutch-speaking area without encountering any sharp boundary across which mutual intelligibility is broken, but the two end points of this chain are speech

varieties so different from one another that they are not mutually intelligible. In Europe there are many other dialect continua, for example, the Romance continuum stretching across the Iberian peninsula through France and parts of Belgium down to the southern tip of Italy and as far east as the Black Sea, including Portuguese, Spanish, Catalan, French, Italian, and Romanian. Outside of Europe also dialect continua are found, for example in the Chinese, Arabic, Indic, Turkic, and Algonquian language areas.

Since we often have to deal with dialect continua and the resulting gradient intelligibility, it is inevitably quite difficult to calculate how many, and which, languages are spoken in a certain area, or indeed worldwide. Hammerström (2008), however, adopted an abstract perspective in order to show that it may in fact be possible to state such numbers. Say, for example, that we are dealing with a dialect chain of three dialects A, B, and C in a language area where the neighboring dialects (A and B, or B and C) are mutually intelligible, whereas the non-neighboring dialects (A and C) are not. Applying the intelligibility principle for defining languages we must be dealing here with two languages (A/B and C, or A and B/C). However, although Hammarström shows how to count the number of languages in a continuum, he fails to define languages uniquely by means of this line of reasoning. Tamburelli (2014) suggests that a choice between the two possible options in the example above can be made by measuring objective linguistic distances or testing intelligibility. The two language varieties that are linguistically closest or show the highest level of mutual intelligibility should be considered dialects of a single language.

As an objection to using intelligibility measurements for the definition of languages, some linguists have contended that intelligibility scores may be influenced by extra-linguistic factors. Subjects may be influenced by their positive or negative biases and attitudes toward the country and its speakers, interest in or familiarity with other cultures, political borders, or the geographical distance to the place where the language is spoken. Also, the personal characteristics of the subjects, such as age, amount of schooling, psycho-cognitive traits, metalinguistic awareness, previous experience, knowledge of various registers and vocabulary in their own language, learning style, fatigue, and motivation may influence their intelligibility rates. We will deal in greater detail with such extra-linguistic factors in Section 11.4.

As a result of the extra-linguistic and personal characteristics of (groups of) subjects, mutual intelligibility is often asymmetric, such that one group of speakers has more difficulty understanding the other variety than the other way round. Asymmetric intelligibility has been described in the literature for many language pairs, including Swedish and Danish (Delsing and Lundin Åkesson 2005; Schüppert 2011), Spanish and Portuguese (Jensen 1989), and the indigenous Californian Indian languages Achumawi and Atsugewi (Merriam 1926; Voegelin and Voegelin 1946). Asymmetric intelligibility is mostly attributed to extra-linguistic factors such as attitude and contact, but there is also evidence that linguistic characteristics of the language varieties may cause asymmetry. For example, Gooskens and Van Bezooijen (2006) showed that asymmetries in the number of non-cognates and the opacity of the relatedness of cognates between Dutch and Afrikaans result in asymmetric intelligibility. These asymmetries are caused by divergent historical developments in Dutch and Afrikaans with respect to lexicon, grammar, and spelling. German-Dutch mutual intelligibility has also been found to be asymmetrical, with German being easier to understand for speakers of Dutch than Dutch is for speakers of German. This finding has been attributed to the fact that German is an obligatory subject at school in the Netherlands and that many Dutch people watch German television. This, however, appears not to be the whole story. Gooskens, Van Bezooijen, and Van Heuven (2015) presented Dutch and German cognate nouns to Dutch and German children between 9 and 12 years of age who did not know the other language or a related dialect, and who all expressed positive attitudes toward the other language, its speakers and the country. The Dutch subjects proved to be significantly better at understanding the German

cognates than the German subjects were at understanding the Dutch cognates. However, since extra-linguistic factors had been ruled out, this asymmetry must have a linguistic basis. A closer look at the data revealed that asymmetries between the two languages are found at the phonetic level and in the presence or absence of neighbors, that is, competing word forms that are very similar to the stimulus word.

Reflecting on the objections to intelligibility as the most important criterion for distinguishing between dialects and languages (the *Abstand* criterion), we see that it is not unproblematic to use this criterion. More research is needed before we will be able to establish when two varieties are so different that they are no longer mutually intelligible, and which linguistic factors play a role. Since there is no universally accepted criterion for distinguishing a language from a dialect, the examples provided in the rest of this chapter are from language varieties that are traditionally referred to as dialects as well as closely related varieties that are mostly referred to as languages.

11.3   Intelligibility as a Measure of Distance

As mentioned in the previous section, it is not a straightforward task to quantify distance. The problem is that languages do not differ along just one dimension, but may differ at all linguistic levels. At each of the linguistic levels, languages may furthermore vary on many different parameters. For example, vowel distances may differ from consonant distances, and the percentages of common loanwords may differ from the percentages of common inherited words (see Section 11.4). Ideally, we would like to express the linguistic distance between language varieties using a single number on a one-dimensional scale. However, there is no *a priori* way of weighing the different linguistic dimensions. As objective techniques have become more sophisticated, resulting in more methods by which language varieties might be distinguished, many researchers have felt an increasing need to "validate" objective methods by means of subjective, behavioral tests (Heeringa *et al*. 2006).

Intelligibility testing is an adequate way of determining how different two languages or language varieties are. If two language varieties have a high degree of mutual intelligibility the linguistic distance must be small, and if they have a low degree of mutual intelligibility the distances are likely to be larger, unless some extra-linguistic factor interferes (see Section 11.4.2). A few investigations have been carried out to validate objectively-measured linguistic distances by means of functional intelligibility tests. Gooskens, Heeringa, and Beijering (2008) assessed the intelligibility of seventeen Scandinavian language varieties and standard Danish among young Danes from Copenhagen by means of a translation task. In addition, distances between standard Danish and each of the seventeen varieties were measured at the lexical level, expressed as the percentage of cognates, and at the phonetic level, by means of Levenshtein distances, a dialectometric technique that calculates distances on the basis of matched segment strings (for an explanation of this algorithm, see Nerbonne and Heeringa 2010). They correlated the intelligibility scores with the linguistic distances and found fairly high, significant correlations. Phonetic distance was a better predictor of intelligibility ($r=-.86$) than lexical distance ($r=-.64$). Similar results are reported by Tang and Van Heuven (2008), who tested mutual intelligibility among 15 Chinese dialects by means of a word-intelligibility and a sentence-intelligibility task. They correlated the scores with measures of lexical similarity and phonological correspondences, and found significant correlations of between .75 and .79. These results show that objective distance measures reflect experimental intelligibility results to a large extent, yet not perfectly. We will suggest explanations for the discrepancy in Section 11.4.

As explained in Section 11.2, it is time consuming both to develop and to carry out suitable functional tests. An easy and efficient alternative to get a quick impression of the intelligibility of a language is to ask subjects to rate on scale(s) how well they think they understand the language at hand. Such opinion testing may provide a shortcut to functional intelligibility tests.

In the investigation by Tang and Van Heuven (2008) described above, intelligibility rates gained by opinion testing were correlated with functional intelligibility rates, yielding correlations of between .70 and .80. These imperfect correlations suggest that tests of impressionistic intelligibility and functional intelligibility tests are sensitive to different factors. In order to make a choice between the two, the authors generated hierarchical cluster trees from their data matrices and compared the results to traditional taxonomies of Chinese dialects proposed by dialectologists. Functional intelligibility measures correspond better to traditional dialect taxonomies than opinion scores do. The authors therefore advocate that whenever the resources are available, mutual intelligibility should be tested functionally. The results also show that mutual intelligibility can to some extent be used as a criterion to illustrate the genetic relationship between speech varieties.

It seems likely that linguistic distance judgments are based on how difficult a listener thinks it would be for him or her to understand speakers of the other language variety. Another shortcut to functional intelligibility testing could therefore be to ask listeners to judge linguistic distances. Tang and Van Heuven (2009) compared the results from their intelligibility tests to perceived linguistic distances, which are gathered by having subjects listen to speech recordings and asking them to judge how deviant the varieties are from their own variety. They found significant correlations of .74 for word intelligibility and .78 for sentence intelligibility. Again, this shows that although there is a relatively large overlap between the two measurements they are still sensitive to different phenomena. The extent to which perceived distance is a reflection of intelligibility, and why the two differ, remains uncertain.

11.4   The Role of Linguistic and Extra-Linguistic Factors for Intelligibility

In the previous section we stated that intelligibility measurements can be used as a way of expressing linguistic distances between language varieties in a single number on a one-dimensional scale. Language varieties may differ at all linguistic levels, and when testing intelligibility extra-linguistic factors such as attitude and linguistic experience may also play an important role. In this section, an overview will be given of investigations that have dealt with the role of various linguistic and extra-linguistic factors in the intelligibility of dialects and closely related languages.

11.4.1   The Role of Linguistic Factors

11.4.1.1   Lexicon

At the lexical level the linguistic distance is often expressed as the percentage of non-cognates between two language varieties. The larger the proportion of non-cognates, the lower the intelligibility will be. The Scandinavian investigations discussed earlier revealed that lexical distances can only predict intelligibility to a limited extent. There are a number of explanations for this finding.

First, it is difficult to predict the effect of individual lexical differences. One single non-cognate word in a sentence or text can lower intelligibility considerably if the non-cognate word is a central concept. For example, one of the texts that were used in an investigation of mutual intelligibility between Swedish and Danish by Delsing and Lundin Åkesson (2005) was about frogs. Since the word for "frog" is a non-cognate noun (Danish *frø*, Swedish *groda*), the whole text was very difficult to understand. On the other hand, if the non-cognate words in a text have little semantic content or can easily be interpreted from the context, lexical differences will have less influence on intelligibility.

Furthermore, it is possible that listeners understand some non-cognate words because they are familiar with the words from previous experience with the test language, or because they are loanwords from a language that they are familiar with. For example, Swedish has many French loanwords that are not found in Danish. Knowledge of French might therefore enable a Dane to understand some Swedish non-cognates.

Whereas non-cognates will in principle hinder intelligibility, so-called "false friends" may cause even larger problems because they may actually mislead the listener. False friends are pairs of words in two language varieties that sound similar, but differ in meaning. They may arise because words with shared etymology shifted in meaning, or acquired additional meanings in at least one of these languages. For example, the meanings of German *Meer* "sea" and its Dutch cognate *meer* "lake" have changed over time. In certain cases, false friends evolved from words with different etymological roots. Words usually change by small shifts in pronunciation accumulated over long periods of time, and sometimes converge by chance towards the same pronunciation or spelling. For example, the English word *bra* has a different etymology from the Swedish word *bra* "good."

11.4.1.2   *Phonetics/Phonology*

The results presented earlier on the relationship between phonetic distances and intelligibility show that at an aggregate level, that is, summed over larger stretches of speech, phonetic distances are a good predictor of the intelligibility of whole texts, and in the Scandinavian case they are better predictors than lexical distances.

A number of investigations have focussed on the role of specific phonetic characteristics in the intelligibility of words. In their investigation of the intelligibility of 17 Scandinavian language varieties among Danes (see Section 11.3), Gooskens, Heeringa, and Beijering (2008) investigated the role of different consonant and vowel operations (insertions, deletions, substitutions, lengthenings, and shortenings). The correlations for the consonants were significantly stronger than those for the vowels ($r = -.74$ versus $r = -.29$). Consonant substitutions play a particularly important role in intelligibility, probably because the "framework" of the word is changed when consonants in a word are substituted. By contrast with consonant substitutions, vowel substitutions play a negligible role in intelligibility.

To be able to determine the role of specific phonetic factors in detail, researchers have often chosen to test word intelligibility rather than the intelligibility of sentences or whole texts. The underlying assumption here is that word recognition is the key to speech understanding: if the listener correctly recognizes a minimal proportion of words, he or she will be able to piece the speaker's message together. Van Bezooijen and Van den Berg (1999) looked at the basis of intelligibility ratings given by speakers of Standard Dutch, and tried to explain why three Dutch dialects and the closely related language Frisian yielded widely diverging results. They made a linguistic profile for each variety, distinguishing six categories of relationships between the target noun in the dialect and the semantically equivalent noun in Standard Dutch (no difference, difference in one vowel, difference in one consonant, differences in several phonemes, non-cognate). One of the results was that the intelligibility of Frisian was equal to that of West Flemish, so one would expect the two to have similar linguistic profiles. However, Frisian has considerably more instances of words that were identical to the Standard Dutch equivalent and considerably fewer non-cognates than West Flemish. Compared to West Flemish, vowel differences between Standard Dutch and Frisian were considerably less transparent. This would mean that Frisian is relatively difficult to understand not only for quantitative reasons, that is, because of the number of nouns showing the various relationships, but also for qualitative reasons, because of the types of deviations within particular categories. But since vowels

played a smaller role in the Scandinavian context (see above) it also means that the role of deviating vowels may be dependent upon the variety and the listeners.

Phonetic details may play an important role in the intelligibility of cognates in related languages and language varieties, in as yet unpredictable ways. Broad transcriptions are therefore sometimes unfit to be used as a basis for the calculation of the phonetic distance between pairs of words with a view to predicting intelligibility. As discussed in Section 11.2, Gooskens, Van Bezooijen, and Van Heuven (2015) tested the mutual intelligibility of German and Dutch among children with no previous knowledge of the test language. They found several cases in which the word in the stimulus language and the corresponding cognate in the response language were represented by the same sequence of phonetic symbols in the Levenshtein algorithm, but in which many subjects nevertheless did not succeed in recognizing the stimulus word. This holds, for example, for Dutch *zoon* /zoːn/ "son," which is phonemically transcribed with the same symbols as its German cognate *Sohn* /zoːn/, but which was nevertheless correctly identified by no more than 20.6% of the German subjects. The high proportion of incorrect responses suggest that there are subtle differences in the phonetic realizations of Dutch and German /z/, which are not expressed in the broad transcription the authors used, and which is commonly used in other intelligibility studies as well. On the other hand, there were cases in which different transcriptions of words nevertheless yielded high intelligibility. For example, half of the transcription symbols in Dutch *stad* /stɑt/ "city" differ from those in German *Stadt* /ʃtat/, resulting in a Levenshtein distance of 50%, but the mutual intelligibility was high nonetheless (92.9% for the Dutch subjects and 94.1% for the German subjects).

In order to find out with which sound in the listener's native language a non-native sound from a closely-related language is identified, we may turn to the Perceptual Assimilation Model (PAM) developed by Catherine Best and her co-workers (e.g., Best 1995; Best, McRoberts, and Goodell 2001). PAM was developed to predict and explain the behavior of learners of a second language when first confronted with the sounds of the target language. The results of perceptual assimilation experiments reveal which categories in the listener's native language are likely to be matched with a non-native sound (Van Heuven 2008). Such knowledge might be used to weight phonetic differences differentially, for example, depending on the intuitions of listeners about the differences between the two segments involved in a substitution (see Wieling, Margaretha, and Nerbonne (2012) and references therein).

It should be noted that the effect of phonetic similarity between the stimulus and the intended response may be overruled by the presence of neighbors. However similar a stimulus and the intended response may be, if there is another word in the subject's language that is even closer to the stimulus, the latter has a high chance of being preferred. This will lead to (severely) reduced intelligibility for that word, especially in the absence of linguistic or extra-linguistic context.

Listeners are in general better at translating loanwords correctly than inherited words. Part of the explanation may be that they know the loanwords from the source language, but it is also possible that particular characteristics of loanwords make them easier to recognize. Loanwords may have specific segmental and/or prosodic properties that make them resistant to the linguistic changes affecting inherited words. They are often longer than inherited words because the word length of the loan-giving languages is generally longer, and we know from the literature that longer words are better recognized than shorter words (Wiener and Miller 1946; Scharpff and Van Heuven 1988). This is explained in terms of the relationship between word length and the number of "neighbors" competing to be recognized. Longer words have fewer neighbors than shorter words (Vitevitch and Rodriguez 2005). Furthermore, redundancy increases with word length, which is assumed to enhance intelligibility as well. Furthermore, inherited words have been part of the lexicon for a much longer time than loanwords, so that certain historical sound changes, which affected the inherited vocabulary were no longer active at the time the loans entered the language. As a consequence,

loan words in the neighboring language often have more transparent phonetic correspondences with their counterparts in the mother tongue than inherited words have.

11.4.1.3   Morphosyntax

Previous studies of intelligibility have focussed primarily on the role of lexical and phonetic factors. Although there is reason to believe that differences in morphology and syntax might degrade the ability to comprehend a closely related linguistic variety, this claim has hardly been tested. An exception is Hilton, Gooskens, and Schüppert (2013), who carried out an experimental investigation to see whether Danes' comprehension of the closely related language Norwegian is impeded by certain Norwegian grammatical constructions. They tested sentence comprehension in four different conditions to assess the relative effect on intelligibility of non-native morphosyntactic features as opposed to non-native phonology. The results indicated that word-order differences cause larger problems for listeners than morphological differences. However, the non-native phonology featured in the experiment impedes comprehension to a larger degree than the morphosyntactic differences do. Just as in the case of other linguistic factors, the role of morphosyntax may be language-dependent. In language areas with larger morphological and syntactic variability, morphosyntax may play a more important role in intelligibility than in areas with less variability.

11.4.2   The Role of Extra-Linguistic Factors

11.4.2.1   Attitude

The existence of negative attitudes or social stigmas attached to languages is often seen as a potential obstruction for successful intergroup communication. The fact that Danes understand Swedish better than Swedes understand Danish, for example, is often explained by less positive attitudes among Swedes toward the Danish language, culture, and people than vice versa (Delsing and Lundin Åkesson 2005). Wolff (1959) investigated mutual intelligibility between the closely related Nigerian Ijo languages Kalabari and Nembe, and reports that Nembe speakers claim to understand Kalabari, whereas speakers of Kalabari judge Nembe to be unintelligible to them. Wolff suggests that this asymmetry in intelligibility is linked to an asymmetry in language attitudes. He states that when his study was conducted, the Kalabari were the most prosperous group in the Eastern Niger Delta and that they regarded other Ijo-speaking groups as inferior to them.

Boets and De Schutter (1977) found low intelligibility to correlate with low appreciation. According to Boets and De Schutter, the (subjective) appreciation scores are determined by the (objective) intelligibility scores. In the literature (e.g., Wolff 1959; Van Bezooijen and Gooskens 2007) the opposite is often contended, namely that low (high) intelligibility is caused by low (high) appreciation. It is assumed that the reported or measured comprehension problems are not so much due to a lack of transparency of the meaning of the language at hand, but rather to a lack of motivation on the part of the listeners.

It has not been possible, so far, to establish the direction of the causality, that is, whether negative attitudes are a result of poor intelligibility, or poor intelligibility is a result of negative attitudes caused by some other factor. Language attitude research shows that people have stereotypical associations with languages. An intriguing question is how such stereotypes arise. Giles, Bourhis, and Davies (1975) suggested two possible answers, termed the imposed-norm hypothesis and the inherent-value hypothesis. The imposed-norm hypothesis stresses the importance of extra-linguistic factors such as social connotations and cultural norms. A language is considered attractive when its speakers are

socially privileged. The inherent-value hypothesis, on the other hand, is linguistically based, and argues that some languages are intrinsically more aesthetically pleasing due to their sound characteristics.

Most of the older language-attitude studies seem to support the imposed-norm hypothesis (Trudgill and Giles 1978). More recent studies, however, found evidence for the inherent-value hypothesis. Van Bezooijen (1996) had Dutch subjects evaluate a number of languages aesthetically. Phoneticians rated the same languages on phonetic scales. The rank order of aesthetic evaluations could almost completely be predicted by a combination of melodiousness and softness. Also, fast tempo and precise and fronted articulation were positively correlated with the aesthetic evaluations. These outcomes suggest that aesthetic evaluations may have a phonetic basis. Similar results were obtained by Gooskens, Schüppert, and Hilton (2016) via a matched-guise experiment. They made recordings of a perfect Swedish/Danish bilingual speaker and presented them, together with a number of filler languages, to Chinese students. They were asked to judge how beautiful the languages sounded. The subjects found Swedish significantly more beautiful than Danish. Since the subjects were unfamiliar with the test languages and the speaker of both languages was the same, imposed norms and speaker characteristics cannot have influenced the judgments. The differences in judgments must therefore have been caused by characteristics of the languages themselves. As with Van Bezooijen's study, this investigation provides clear evidence that inherent language characteristics play a role in aesthetic evaluations. However, it still leaves open the question of whether these attitudes also influence intelligibility, and what the direction is of the causality.

11.4.2.2   Contact and Experience

Of course, the level of intelligibility also depends on the amount of experience and contact, including formal instruction, that the listener has had with the other language. However, it has often been difficult to find a direct link. It has been assumed, for example, that the asymmetric intelligibility between Swedish and Danish, for example, in the investigation by Maurud (1976), could at least partly be explained by the fact that the listeners came from the capitals of Sweden and Denmark. As Copenhagen is located only 30 kilometres (20 miles) from the Swedish border, whereas Stockholm is located about 570 kilometres (350 miles) from the Danish border, there is a substantial geographical asymmetry in the origin of the subjects. The Danes in Maurud's investigation had more opportunities to hear and read the neighboring language than did the Swedes. Bø (1978) therefore tested the intelligibility of the neighboring language among two groups, one living inside and one living outside the border regions of Sweden and Denmark. The border region group not only had more opportunities to visit the neighboring country, but also had access to television programs in the neighboring language. The results showed that this group of subjects had fewer difficulties decoding the neighboring variety than did subjects living outside the border region, thereby indicating that a high degree of contact indeed enhances intelligibility.

11.4.2.3   Orthography

Another explanation for the asymmetric intelligibility of Swedish and Danish might be found in the relationship between the written and the spoken forms of the languages. Spoken Swedish is close to both written Swedish and written Danish, whereas spoken Danish has undergone a number of reduction processes, which are not reflected in the orthographic system. Danish pronunciation has changed more rapidly during the last century than Swedish pronunciation has. As a consequence, spoken Danish has developed

away from its written form and is therefore rather distant from both Swedish and Danish in their written forms. Danes can understand spoken Swedish better because of its close similarity to written Danish, while Swedes get less help from written Swedish when listening to spoken Danish. For example, it is likely that literate Danes confronted with the Swedish word /land/ "country" can use their orthographic knowledge to match this word to their native correspondent *land*, whereas this is not the case for Swedish listeners confronted with Danish /lanʔ/ because of the absence of the phoneme /d/, which is present in Swedish pronunciation as well as orthography.

Doetjes and Gooskens (2009) quantified the relationship between the spoken and written representations of Swedish and Danish in a corpus of 86 frequent cognate words, first by measuring phonetic and orthographic distances between the languages by means of the Levenshtein algorithm. As expected, the phonetic distance was larger (53%) than the orthographic distance (24%). Next, they calculated the distances again, but this time corrected the phonetic distance values for the advantage that Danes and Swedes gain from their native orthography when listening to the neighboring language. This was done by setting the segment distance to zero in cases where a phoneme could be understood from its orthographic equivalent in the native language. After correcting for orthography, the distances turned out to be smaller for the Danes (30%) than for the Swedes (46%). This indicates that Danes obtain more potential help from the orthography than Swedes do. Doetjes and Gooskens (2009) tested this hypothesis by correlating the distances with the results of a word-intelligibility experiment run using Danish participants. Distance values corrected for the influence of orthography showed higher correlations with the intelligibility scores than pure phonetic distances. The authors conclude that Danish listeners indeed seem to make use of the additional information that the orthography can provide.

This claim was made even stronger by an investigation by Schüppert (2011). She played spoken Swedish words to Danish speakers in a translation task. The words were cognates in which the pronunciation differed in one phonetic segment only (e.g., the word *mild* "mild" is pronounced /mild/ in Swedish but /milʔ/ in Danish). Half of the Swedish cognates were pronounced in a way that would be consistent with the spelling of the Danish word (i.e., orthographically consistent cognates), whereas the other half were pronounced in a way that would not be consistent with the spelling of the Danish word (i.e., orthographically inconsistent cognates). Event-related brain potentials (ERPs) in the translation task were measured for these consistent and inconsistent cognates to study the participants' online brain responses during decoding operations over the first 1000 milliseconds (ms). The data showed that ERPs in response to inconsistent words were significantly more negative than ERPs for consistent words between 750 and 900 ms after stimulus onset. Together with higher word-recognition scores for consistent items, the data provide strong evidence that online activation of L1 orthography enhances word recognition among literate speakers of Danish who are exposed to samples of spoken Swedish.

11.4.2.4　Gestures

Co-speech gestures (movements of the hands, face, or other parts of the body that people spontaneously produce when speaking) and mouthing (movements of the mouth) reflect important aspects of oral communication (Kita and Özyürek 2003). Experimental studies have shown that subjects who see accompanying gestures while hearing native speech pick up significantly more relevant information than do subjects who only listen to speech (Graham and Argyle 1975; Riseborough 1981). The role of co-speech gestures and mouthing in the intelligibility of closely related languages has, however, hardly been studied. Voigt and Gooskens (in preparation) tested the influence of gestures and mouthing on the

intelligibility of spoken Spanish for Italian listeners. The experiment incorporated four conditions. In the first condition, the subjects saw a video recording of the full upper body of the speaker while she retold a story. In the second condition, the subjects could only see the head of the speaker while listening to what she said. In the third condition, the head of the Spanish speaker was obscured and the subjects could only hear the story and see the gestures. The fourth condition consisted of only the audio file and a blank screen. The mean percentages of correct answers were 66.8% for condition 1 (full body and audio), 50.3% for condition 2 (head and audio), 52.7% for condition 3 (gestures and audio) and 47.2% for condition 4 (audio only). The differences between conditions 1 and 2, and between conditions 1 and 3, are statistically significant. These results confirm the hypothesis that co-gestures and mouthing facilitate the intelligibility of an unknown, yet related, language.

11.5   Desiderata for Future Research

In this chapter the role of intelligibility in the definition of dialects as opposed to languages was discussed. It was shown that the intelligibility criterion brings with it a number of problems, but it was also stated that modern research methodologies can open up useful ways of measuring intelligibility. It is difficult to determine when two language varieties are mutually intelligible to such an extent that they can be considered dialects of the same language, and when communication is so difficult that they should be considered different languages. It would be useful to set up a standard for testing the intelligibility of dialects and closely related languages. This would make it possible to compare the results of different investigations, and perhaps even to define a level that can be considered the threshold at which intelligibility is sufficient for communication.

An interesting related question that still remains unanswered is whether intelligibility is gradual, or whether it is possible to define some breakdown point at which language varieties become unintelligible to listeners. Such a critical breakdown threshold would depend upon the linguistic differences between the language varieties in question. If the number of non-cognates varies around the breakdown threshold, lexical distance will be more important than phonological/phonetic distance within the cognates. Similarly, if a language has more (or fewer) consonants relative to vowels in its phoneme inventory, the importance of vowel and consonant distance will be different from what obtains in another language. Typically, the relationship between the number and magnitude of deviations and the intelligibility of a linguistic unit is non-linear. Identification of a sound or recognition of a word remains very good for small discrepancies from the norm, but abruptly breaks down when these discrepancies become larger. Future work with different language varieties and more controlled representations of various linguistic units can hopefully yield more insight into the relative contributions of linguistic phenomena to intelligibility, and show when the limits of intelligibility have been reached.

New methods for testing the intelligibility of closely related language varieties have so far mainly generated global results. In future research we should aim to gain more detailed knowledge about the mechanisms behind the intelligibility of language varieties. Methods that have been developed by experimental linguists and psycholinguists should be exploited when setting up controlled experiments that will give us more insight into the relative importance of various linguistic and extra-linguistic factors that impact upon the intelligibility of language varieties.

By testing the intelligibility of a large number of languages differing along many dimensions we may establish the relative importance of the various dimensions. This will allow us to provide a more solid, experimentally grounded, foundation for traditional claims made by linguists about genealogical relatedness among languages.

Intelligibility between languages may also serve as the ultimate criterion to decide how structural dimensions should be weighed against each other in the computation of linguistic distance.

REFERENCES

Best, Catherine. 1995. "A direct realist perspective on cross-language speech perception." In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Research*, edited by Winifred Strange, 167–200. Timonium, MD: York Press.

Best, Catherine, Gerald McRoberts, and Elizabeth Goodell. 2001. "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system." *Journal of the Acoustical Society of America*, 109: 775–794.

Bø, Inge. 1978. *Ungdom og Naboland: En Undersøkelse av Skolens og Fjernsynets Betydning for Nabospråksförståelsen*. Stavanger: Rogalandsforskning.

Boets, Herman, and Georges de Schutter. 1977. "Verstaanbaarheid en appreciatie: Nederlandse dialecten uit België zoals inwoners van Duffel die ervaren." *Taal en Tongval*, 29: 156–177.

Casad, Eugene. 1974. *Dialect Intelligibility Testing*. Norman, OK: Summer Institute of Linguistics of the University of Oklahoma.

Delsing, Lars-Olof, and Katarina Lundin Åkesson. 2005. *Håller Språket ihop Norden? En Forskningsrapport om Ungdomars Förståelse av Danska, Svenska och Norska*. Copenhagen: Nordiska Ministerrådet.

Doetjes, Gerard, and Charlotte Gooskens. 2009. "Skriftsprogets rolle i den dansk-svenske talesprogsforståelse." *Språk och Stil*, 19: 105–123.

European Charter for Regional or Minority Languages. 1992. http://conventions.coe.int/treaty/en/Treaties/Html/148.htm (accessed September 19, 2015).

Giles, Howard, Richard Bourhis, and Ann Davies. 1975. "Prestige speech styles: the imposed norm and inherent value hypotheses." In *Language in Anthropology IV: Language in Many Ways*, edited by William McCormack, and Stephen Wurm, 589–596. The Hague: Mouton.

Giles, Howard, and Nancy Niedzielski. 1998. "Italian is beautiful, German is ugly." In *Language Myths*, edited by Laurie Bauer, and Peter Trudgill, 85–93. London: Penguin.

Gooskens, Charlotte. 2013. "Methods for measuring intelligibility of closely related language varieties." In *The Oxford Handbook of Sociolinguistics*, edited by Robert Bayley, Richard Cameron, and Ceil Lucas, 195–213. Oxford: Oxford University Press.

Gooskens, Charlotte, Anja Schüppert, and Nanna Haug Hilton. 2016. "Is Swedish more beautiful than Danish? – A matched-guise investigation." In *Nooit het Noorden kwijt*, edited by Sara Van den Bossche and Martje Wijers, 165–182. Ghent: Academia Press.

Gooskens, Charlotte, Wilbert Heeringa, and Karin Beijering. 2008. "Phonetic and lexical predictors of intelligibility." *International Journal of Humanities and Arts Computing*, 2(1–2): 63–81.

Gooskens, Charlotte, and Renée van Bezooijen. 2006. "Mutual comprehensibility of written Afrikaans and Dutch: Symmetrical or asymmetrical?" *Literary and Linguistic Computing*, 23: 543–557.

Gooskens, Charlotte, Renée van Bezooijen, and Vincent van Heuven. 2015. "Mutual intelligibility of Dutch-German cognates by children: The devil is in the detail." *Linguistics*, 53(2): 255–283.

Graham, Jean, and Michael Argyle. 1975. "A cross-cultural study of the communication of extra-verbal meaning by gestures." *Journal of Human Movement Study*, 1: 33–39.

Hammarström, Harald. 2005. "Review of Ethnologue: Languages of the World, 15th Edition." http://linguistlist.org/issues/16/16-2637.html (accessed November 10, 2015).

Hammarström, Harald. 2008. "Counting languages in dialect continua using the criterion of mutual intelligibility." *Journal of Quantitative Linguistics*, 15(1): 34–45.

Heeringa, Wilbert, Peter Kleiweg, Charlotte Gooskens, and John Nerbonne. 2006. "Evaluation of string distance algorithms for dialectology." In *Linguistic Distances Workshop at the Joint Conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics, Sydney, July, 2006*, edited by John Nerbonne,

and Erhard Hinrichs, 51–62. Stroudsburg, PA: The Association for Computational Linguistics.

Hickerson, Harold, Glen Turner, and Nancy Hickerson. 1952. "Testing procedures for estimation transfer of information among Iroquois dialects and languages." *International Journal of American Linguistics*, 18: 1–8.

Hilton, Nanna Haug, Charlotte Gooskens, and Anja Schüppert. 2013. "The influence of non-native morphosyntax on the intelligibility of a closely related language." *Lingua*, 137: 1–18.

Hudson, Richard. 1996. *Sociolinguistics*. Cambridge: Cambridge University Press.

Jensen, John. 1989. "On the mutual intelligibility of Spanish and Portuguese." *Hispania*, 72 (4): 848–852.

Kita, Sotaro, and Asli Özyürek. 2003. "What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking." *Journal of Memory and Language*, 48(1): 16–32.

Kloss, Heinz. 1967. "Abstand languages and Ausbau languages." *Anthropological Linguistics*, 9(7): 29–41.

Lewis, M. Paul, Gary Simons, and Charles Fennig, eds. 2013. *Ethnologue: Languages of the World*, 17th ed. Dallas, TX: SIL International.

Maurud, Øivind. 1976. *Nabospråksforståelse i Skandinavia: En Undersøkelse om Gjensidig Forståelse av Tale- og Skriftspråk i Danmark, Norge og Sverige*. Stockholm: Nordiska Rådet.

Merriam, Clinton. 1926. "The classification and distribution of the Pit River Indian tribes of California." *The Smithsonian Institution Miscellaneous Collection*, 78 (3): 1–52.

Nahhas, Ramzi. 2006. *The Steps of Recorded Text Testing: A Practical Guide*. Chiang Mai: Payap University.

Nerbonne, John, and Wilbert Heeringa. 2010. "Measuring dialect differences." In *Language and Space: An International Handbook of Linguistic Variation, Theories and Methods*, edited by Peter Auer, and Jürgen Schmidt, 550–567. Berlin: de Gruyter.

Pierce, Joe. 1952. "Dialect distance testing in Algonquian." *International Journal of American Linguistics*, 18: 208–218.

Riseborough, Margaret. 1981. "Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication." *Journal of Nonverbal Behavior*, 5: 172–183.

Scharpff, Peter, and Vincent van Heuven. 1988. "Effects of pause insertion on the intelligibility of low quality speech." In *Proceedings 7th FASE Symposium/Proceedings Speech '88*, edited by William Ainsworth, and John Holmes, 261–8. Edinburgh: The Institute of Acoustics.

Schüppert, Anja. 2011. *Origin of Asymmetry: Mutual Intelligibility of Spoken Danish and Swedish*. PhD dissertation, University of Groningen.

Tamburelli, Marco. 2014. "Uncovering the 'hidden' multilingualism of Europe: An Italian case study." *Journal of Multilingual and Multicultural Development*, 35(3): 252–270.

Tang, Chaoju, and Vincent van Heuven. 2008. "Mutual intelligibility of Chinese dialects tested functionally." In *Linguistics in the Netherlands*, edited by Marjo van Koppen, and Bert Botma, 145–156. Amsterdam: Benjamins.

Tang, Chaoju, and Vincent van Heuven. 2009. "Mutual intelligibility of Chinese dialects experimentally tested." *Lingua*, 119(5): 709–732.

Trudgill, Peter. 2000. *Sociolinguistics: An Introduction to Language and Society*. London: Penguin.

Trudgill, Peter, and Howard Giles. 1978. "Sociolinguistics and linguistic value judgments: Correctness, adequacy, and aesthetics." In *Functional Studies in Language and Literature*, edited by Frank Coppieters, and Didier Goyvaerts, 167–190. *Functional Studies in Language and Literature*. Gent: Story-Scientia.

Van Bezooijen, Renée. 1996. "Aesthetic evaluation of Dutch: Comparison across dialects, accents, and languages." In *Handbook of Perceptual Dialectology*, Vol. 2, edited by Daniel Long, and Dennis Preston, 13–31. Amsterdam: Benjamins.

Van Bezooijen, Renée, and Charlotte Gooskens. 2007. "Interlingual text comprehension: Linguistic and extralinguistic determinants." In *Receptive Multilingualism and Intercultural Communication: Linguistic Analyses, Language Policies and Didactic Concepts*, edited by Jan ten Thije, and Ludger Zeevaert, 249–264. Amsterdam: Benjamins.

Van Bezooijen, Renée, and Rob van den Berg. 1999. "Word intelligibility of language varieties in the Netherlands and Flanders under minimal conditions." In *Linguistics in the Netherlands*, 16, edited by Renée van Bezooijen, and René Kager, 1–12. Amsterdam: Benjamins.

Van Heuven, Vincent. 2008. "Making sense of strange sounds: (Mutual) intelligibility of related language varieties – A review." *International Journal of Humanities and Arts Computing*, 2(1–2): 39–62.

Vitevitch, Michael, and Eva Rodríguez. 2005. "Neighborhood density effects in spoken word

recognition in Spanish." *Journal of Multilingual Communication Disorders*, 3: 64–73.

Voegelin, Carl, and Erminie Voegelin. 1946. "Linguistic considerations of Northeastern North America." In *Man in North Eastern North America*, edited by Frederick Johnson, 178–94. Andover, MA: Peabody Foundation.

Voegelin, Carl, and Zellig Harris. 1951. "Methods for determining intelligibility among dialects of natural languages." *Proceedings of the American Philosophical Society* 95: 322–329.

Voigt, Stefanie, and Charlotte Gooskens. In preparation. "Co-speech gestures and their influence on the intelligibility of Spanish for Italian listeners."

Weijnen, Antonius. 1941. *De Nederlandse Dialecten*. Groningen: Noordhoff.

Wieling, Martijn, Eliza Margaretha, and John Nerbonne. 2012. "Inducing a measure of phonetic similarity from pronunciation variation." *Journal of Phonetics*, 40(2): 307–314.

Wiener, F., and G. Miller. 1946. "Some characteristics of human speech." *Transmission and Reception of Sounds under Combat Conditions: Summary Technical Report of Division 17, National Defence Research Committee*, 58–68. Washington, DC: National Defence Research Committee.

Wolff, Hans. 1959. "Intelligibility and inter-ethnic attitudes." *Anthropological Linguistics* 1: 34–41.