ڈاکٹر رؤف پاریکھ
ایسوسی ایٹ پروفیسر، شعبۂ اردو، جامعہ کراچی

لغت نویسی میں کورپس، کورپس لسانیات، وصفیت اور تجربیت کا کردار

In addition to offering an introduction to corpus and corpus linguistics, this article describes how corpus linguistics can help to compile dictionaries. Taking into account Noam Chomsky's views on grammar and corpus, this article discusses the views on descriptivism, prescriptivism, empiricism and rationalism with reference to grammar and lexicography. An important issue in lexicography is what makes a dictionary authentic and why an innate and intropspective knowledge of language alone cannot be relied upon as evidence. The article explains why corpus can be used as a tool as well as linguistic evidence in compiling dictionaries.

اس مقالے میں ہم کورپس (corpus) اور کورپس لسانیات (corpus linguistics) کا تعارف پیش کرکے یہ دیکھیں گے کہ دورِ حاضر میں لغات کی تیاری میں کورپس اور کورپس لسانیات کا کیا استعمال ہے۔ ہم اس امر کا بھی جائزہ لیں گے کہ لغت نویسی میں سند کے طور پر کورپس کے استعمال پر زبان کی صحت اور معیار کا سوال کس طرح اٹھتا ہے نیز یہ کہ لغت اور قواعد کے وصفی (descriptive) یا تجویزی (prescriptive) ہونے کی بحث کورپس سے کس طرح جڑی ہوئی ہے اور علمِ لسانیات کا رویہ اس ضمن میں کیا ہے۔ ضمنی طور پر ہم یہ بھی دیکھیں گے کہ لغت کی تیاری میں کورپس کے استعمال اور تجربیت (empiricism) کو نوم چومسکی جیسے ماہرِ لسانیات کیوں ناپسند کرتے ہیں۔

کورپس (corpus) کیا ہے؟

دورِ حاضر میں لغت نویسی کے نئے رجحانات میں سے ایک یہ بھی ہے کہ لغت نویسی اب کورپس (corpus) کی مدد سے کی جاتی ہے۔ خاص طور پر انگریزی اور دیگر مغربی زبانوں میں اب کورپس کے بغیر لغت نویسی کا تصور نہیں کیا جاسکتا۔ دراصل کمپیوٹر کی روز افزوں ترقی نے جہاں کئی علوم وفنون پر نئے امکانات کے در وا کیے ہیں وہاں لسانیات اور لغت نویسی کے میدان میں بھی کمپیوٹر نے کئی آسانیاں پیدا کردی ہیں۔

پہلے ہم یہ دیکھتے ہیں کہ کورپس کیا ہے، کورپس لسانیات کیا ہے اور لغت کی تیاری میں اس سے کیسے مدد لی جاسکتی ہے اور لی جارہی ہے۔

کورپس کے لفظی معنی ہیں ''جسم''لیکن اس سے مراد ہے ''تحریری متون کا مجموعہ''، بالخصوص جسے کمپیوٹر کی مدد سے لکھا، پڑھا جاسکے یا اس پر کمپیوٹر پروگرام کی مدد سے کوئی عمل (process)، مثلاً ترمیم و اضافہ یا تجزیہ وغیرہ، کیا جاسکے ۱۔ (ایسے متن کو جس پر کمپیوٹر کی مدد سے کوئی عمل (process) ہوسکے انگریزی میں ''مشین ریڈیبل'' machine readable کہتے ہیں۔ انگریزی میں corpus کی جمع corpuses بھی آتی ہے اور corpora بھی۔ اردو میں کورپس کے لیے کوئی الگ لفظ مستعمل نہیں ہے۔ اردو میں اسے عموماً کورپس ہی لکھا جاتا ہے۔ فارسی میں اس کا ایک مترادف ''مجموعہ ای ازنوشتجات'' ہے۔ فارسی میں کورپس کے لیے ایک اور اصطلاح بھی استعمال ہوتی ہے اور وہ ہے ''پیکرہ''۔ کمپیوٹر پر بنے لسانیاتی کورپس (computerised linguistic corpus) کو فارسی میں'' پیکرۂ زبانی رایانہ ای'' بھی کہا جاتا ہے ۲ (فارسی میں کمپیوٹر کو رایانہ کہتے ہیں)۔ اردو میں کورپس کو مجموعۂ تحریرات یا مجموعۂ نوشتہ جات کہا جاسکتا ہے، گو اب کورپس کا لفظ بھی رائج ہو چلا ہے۔

کسی زمانے میں دست نوشتہ کورپس اور کمپیوٹر پر لکھے گئے کورپس میں تخصیص کی جاتی تھی لیکن اب ہاتھ سے لکھے ہوئے کورپس کا استعمال مغرب میں اتنا کم ہوگیا ہے اور کمپیوٹر کا استعمال اس کام میں اتنا بڑھ گیا ہے کہ کورپس کا لفظ اب تقریباً machine readable کا مترادف ہوگیا ہے ۳۔ کورپس کی ابتدا کی ایک صورت یہ بھی تھی کہ لغت نویس اپنے استعمال کے لیے الفاظ اور ان کے استعمال کی مثالیں لکھ لیتے تھے۔ انگریزی کے مشہور معجم الفاظ ''روجٹس تھیسارس''(Roget's Thesaurus) کی تکمیل میں کم و بیش چالیس سال کا عرصہ لگا اور اس کا مولف ڈاکٹر پیٹر مارک روجٹ (Peter Mark Roget) (۱۸۶۹ء۔۱۷۷۹ء) الفاظ کی فہرستیں عنوانات کے تحت بناتا رہا ۴۔ یہ بھی ایک کورپس تھا۔ لغت نویس بالعموم الفاظ کے استعمال کے نمونے ادب سے جمع کرتے تھے اور پھر لغت میں مفہوم کے تعین اور تشریح نگاری میں ان سے مدد لیتے تھے۔ یہ سب نمونے بھی کورپس ہی تھے اور ظاہر ہے کہ یہ سب کام ہاتھ سے ہوتا تھا۔ اردو لغت بورڈ کی بائیس جلدوں پر مبنی اردو کی ضخیم ترین لغت ''اردو لغت (تاریخی اصول پر)'' کے لیے الفاظ اور ان کے استعمال کی اسناد بھی دست نوشتہ ہی ہیں اور ایسے بارہ سے چودہ لاکھ کارڈ اردو لغت بورڈ میں موجود ہیں۔ یہ اردو کا سب سے بڑا کورپس ہے (اسے کمپیوٹر پر منتقل کرنے کی اشد ضرورت ہے)۔

آج کل لسانیاتی مطالعات میں کمپیوٹر سے کئی طرح سے مدد لی جارہی ہے۔ صوتیات، نحو، مارفیمیات اور ذخیرۂ الفاظ کے تجزیے میں بھی اسے استعمال کیا جارہا ہے اور لغت نویسی میں بھی۔ کمپیوٹر کی اس مدد کی ایک صورت یہ ہے کہ متون کا تجزیہ کرکے نتائج اخذ کیے جائیں۔ ابتدا میں تو تحریری متون مثلاً شیکسپیر یا بائبل کے متن کو کمپیوٹر میں محفوظ کیا گیا لیکن پھر اخبارات، رسائل، ادبی متون کو بھی اس کا حصہ بنایا گیا۔ پھر زبانی طور پر بولے گئے متن کو ریکارڈ کرکے اسے

تحریری صورت میں کورپس کا حصہ بنادیا گیا اور زبان کے تجزیے میں استعمال کیا جانے لگا۵۔ حال ہی میں یہ بھی ہوا ہے کہ موبائل فون کے ذریعے بھیجے گئے پیغامات کو بھی کورپس کا حصہ بنادیا گیا ہے۔

دل چسپ بات یہ ہے کہ آج تو کورپس کی بڑی دھوم مچی ہوئی ہے لیکن جب ۱۹۶۰ء کے عشرے میں انگریزی کا پہلا مشینی کورپس (machine readable) یعنی کمپیوٹر پر تیار کیا گیا کورپس زیرِ تکمیل تھا تو اس کی مخالفت کی گئی تھی۔ اس کا نام براؤن کورپس (Brown corpus) تھا ۔اس کے بانی ڈبلیو نیلسن فرانسس (W Nelson Francis) اور ہنری کچیرا (Henry Kucera) تھے ۔فرانسس نے لکھا ہے کہ جب یہ کورپس بن رہا تھا تو تولیدی قواعد (اسے اردو میں ساختی قواعد کا بھی نام دیا گیا ہے، انگریزی میں اسے (generative grammar) کہتے ہیں) کے ماہرین نے اس کی مخالفت کی اور اسے ''بے فائدہ'' اور'' بے وقوفی کا کام'' قرار دیا تھا۶ (یہاں یقیناً نوم چومسکی اور اس کے ہم خیال ماہرین کی طرف اشارہ ہے۔

امریکی انگریزی اور برطانوی انگریزی کے کورپس الگ الگ بنائے گئے ہیں اور یہ کمپیوٹر پر انٹرنیٹ کے ذریعے ''بر خط'' یعنی ''اون لائن'' (online) بھی بہت آسانی سے دست یاب ہیں جس کی وجہ سے ماہرین کو ان مجموعہ ہائے تحریرات سے استفادے کا موقع ملتا ہے۔ ابتدا میں تو دس لاکھ الفاظ پر مبنی کورپس بنائے گئے تھے ے۔لیکن اب ایسے کورپس بر خط موجود ہیں جن میں کروڑوں الفاظ موجود ہیں۔ یہاں یہ وضاحت ضروری ہے کہ انگریزی زبان کے ذخیرۂ الفاظ کے بارے میں عام طور پر کہا جاتا ہے کہ یہ تقریباً پانچ لاکھ الفاظ پر مبنی ہے لہٰذا جب ہم کسی کورپس کے بارے میں یہ کہتے ہیں کہ اس میں کروڑوں الفاظ ہیں تو یہ دراصل ذخیرۂ الفاظ کا نہیں بلکہ ان کے استعمال کی مثالوں کی تعداد کا ذکر ہے۔

بعض کورپس ایسے ہیں جن میں کروڑوں مثالیہ جملے پڑے ہوئے ہیں،مثلاً برٹش نیشنل کورپس BNC میں سو(۱۰۰) ملین یعنی دس کروڑ مثالیں موجود ہیں اور اس کا کچھ حصہ انٹرنیٹ پر مفت ملاحظہ کیا جاسکتا ہے۔ اسی طرح بینک اوف BoE اور کورپس اوف کنٹمپریری امریکن انگلش COCA میں سے ہر ایک میں ساڑھے چارسو (۴۵۰) ملین یعنی پینتالیس کروڑ الفاظ پڑے ہوئے ہیں۔ کورپس میں ایک ہی لفظ کے استعمال کے نمونوں کے طور پر مثلاً دو ہزار جملے موجود ہوسکتے ہیں (جو تقریری مواد، تحریری مواد یا ایس ایم ایس کے متن پر مبنی ہوں گے)۔ یہ کام تجزیہ کرنے والوں کا ہے کہ وہ یہ بتائیں کہ اس میں سے کون سا جملہ کس معنی کو پیش کرتا ہے، یا اسے لغت نویسی کے لیے استعمال کریں یا زبان سے متعلق کسی بھی مسئلے میں بطور نمونہ استعمال کریں یا قواعد کے تجزیے میں برتیں یا کچھ اور کریں (کورپس انٹرنیٹ کے ذریعے قابلِ رسائی ہیں اکثر کورپس سارے کا سارا ذخیرہ نہیں تو اس کا کچھ حصہ مفت دکھا دیتے ہیں، زیادہ کے لیے رکنیت یعنی رقم کا مطالبہ ہوتا ہے) ۔بعض کورپس ایسے ہیں جو صرف گزشتہ دس برسوں کے متون کو پیش کرتے ہیں اور جب گیارھویں برس کا متن اس میں شامل کیا جاتا ہے تو اس میں سے پہلے برس کی مثالیں نکال دی جاتی ہیں اور گویا دوسرا برس اب پہلا برس بن جاتا ہے ، وعلیٰ ہٰذ القیاس۔ اس طرح یہ کورپس انگریزی زبان کے عصری یا رائج الوقت (current)

استعمال کی بھی تازہ ترین مثالیں پیش کرتے ہیں اور اگر گزشتہ دس برسوں میں کسی لفظ کے معنی یا اس کے استعمال میں کمی بیشی ہوئی ہے یعنی اس کے تعدّد (frequency) میں تبدیلی ہوئی ہے تو وہ خود ہی کورپس کا حصہ بن جاتا ہے۔ اب یہ تجزیہ کرنے والوں کا کام ہے کہ وہ اس طرح کی تبدیلی کا تجزیہ کریں ، اس سے لغت نویسی میں مدد لیں یا نحو میں صوتیات یا کسی اور کام میں استعمال کریں۔ ادبی متون کا دل چسپ اسلوبیاتی اور صوتیاتی مطالعہ بھی اس کی مدد سے کیا گیا ہے۔ کاش اردو کا بھی کوئی ایسا کورپس بن سکے۔

کورپس لسانیات (corpus linguistics)

کورپس لسانیات کے ضمن میں یہ وضاحت ضروری ہے کہ یہ زبان کے کسی پہلو کا براہِ راست مطالعہ نہیں ہے بلکہ یہ زبان کے مطالعے کا ایک باضابطہ طریقِ کار یا منظم قرینہ یا میتھوڈولوجی (methodology) ہے۸۔ اس مسئلے پر ماہرین میں خاصی بحث ہوتی رہی ہے کہ کورپس لسانیات لسانی تحقیق کا کوئی میدان ہے یا محض ایک طریقۂ کار (methodology) ۔ کچھ ماہرین اس بات سے اتفاق نہیں کرتے کہ کورپس لسانیات محض طریق کار ہے اور ان کا خیال ہے کہ یہ محض ''میتھوڈولوجی'' سے آگے کی چیز ہے اور بذاتِ خود ایک فلسفہ ہے۹۔ کیونکہ بقول ہالیڈے اس سے زبان کی تفہیم میں ایک صفاتی (qualitative) تبدیلی آرہی ہے ۱۰۔ بہرحال یہ ضرور ہے کہ لسانی مطالعات میں ہم کورپس کی بنیاد پر استوار ایک سوچ رکھ سکتے ہیں۱۱۔

یہ حقیقت ہے کہ کورپس لسانیات نے، خواہ یہ کوئی مطالعہ ہو یا مطالعے کا طریقِ کار اور سوچنے کا انداز، لسانیات کی کئی شاخوں اور اس سے متعلقہ علوم کو متاثر کیا ہے۔ کورپس لسانیات سے لغت نویسی، تدریسِ زبان، تراجم، اسلوبیات، قواعد، صنفی مطالعات (gender studies) اور قانونی لسانیات (forensic linguistics) نے بہت فائدہ اٹھایا ہے۱۲۔ کورپس لسانیات کی رو سے زبان سے متعلق نئی اور مختلف زاویے سے ایسی عملی تحقیق کی گئی ہے جو کورپس کے وجود میں آنے سے پہلے ممکن نہ تھی۱۳۔ گویا کورپس ہی وہ بنیاد ہے جس پر کورپس لسانیات کی عمارت کھڑی ہے اور کورپس لسانیات بنیادی طور پر کورپس کے مطالعے اور تجزیے ہی کا نام ہے۔

لغت نویسی اور کورپس

انگریزی میں کورپس کے کام کا باقاعدہ آغاز ۱۹۵۰ء کے عشرے میں ہوا جب Survey of English usage کیا گیا۱۴، اور پھر ۱۹۶۰ء کے عشرے کے اواخر میں لندن یونی ورسٹی اور براؤن یونی ورسٹی نے مشترکہ طور پر دس لاکھ الفاظ پر مبنی کورپس بنایا۱۵ جسے اب مختصر براؤن کورپس بھی کہا جاتا ہے۔ گویا لغت نویسی میں کمپیوٹر کا اوّلین استعمال ۱۹۶۰ء کے عشرے میں ہوا ، تب سے اب تک تکنالوجی نے لغت نویسی میں پہلے سے کہیں زیادہ اہم اور مرکزی حیثیت حاصل کرلی ہے۱۶۔ دراصل بیسویں صدی کے اواخر میں لغت نویسی کے میدان میں اہم تبدیلیاں رونما ہو رہی تھیں اور اس کی وجہ ڈیٹا پروسیسنگ (data processing) کی تکنالوجی تھی۱۷۔ اس تکنالوجی کی وجہ سے اور کورپس کی بنیاد پر

لغت نویس مثلاً اپنی لغت کے الفاظ کی فہرست برقی طور پر جانچ سکتا ہے جس میں دستی طریقے کے مقابلے میں بہت کم وقت لگتا ہے۔اسی طرح اب لغوی یا لغاتی تحقیق اس بہت بڑے کورپس کی بنیاد پر ہوسکتی ہے جوتحریری اور تقریری زبان پر مبنی ہے ۱۸۔ بیسویں صدی کے اواخر میں لغت نویسی میں کمپیوٹر اور کمپیوٹر پر مبنی کورپس کا استعمال بھر پور طریقے پر شروع ہوگیا۔سطور بالا میں جن کورپسوں کا ذکر کیا گیا ہے ان کو کئی طرح کے تجزیوں میں بھی استعمال کیا گیا اور انگریزی لغت نویسی میں ایک خاص نقطۂ نظر کے تحت بھی استعمال کیا گیا اور وہ نقطۂ نظر تھا۔ بدلتے ہوئے اور عصری اور رائج الوقت مفاہیم کا لغت میں اندراج ،اس کے استعمال کی کثرت یا قلت کے لحاظ سے،اسے استعمال کا تعدّد یا (frequency) کہتے ہیں۔

انگریزی کی بعض لغات ایسی بھی ہیں جواستعمال اور تعدّد کی انہی مثالوں پر مبنی ہیں اور ان میں کثرتِ استعمال کوبھی علامات کے ذریعے واضح کیا گیا ہے۔ مثلاً بہت زیادہ استعمال ہونے والے الفاظ پر پانچ ستارے بنائے گئے اور شاذ و نادر استعمال ہونے والے الفاظ پر ایک ستارہ بنایا گیا، وعلیٰ ہٰذا القیاس۔معروف اشاعتی ادارے کولنس (Collins) کی ''کوبلڈ'' (COBUILD) ڈکشنری اسی طرح بنائی گئی ہے اور اس کورپس پر مبنی ہے جس میں حقیقی زندگی کے کلام یا متن (discourse) سے مثالیں اخذ کی گئی ہیں ۱۹۔

وصفیت (descriptivism)، تجویزیت (prescriptivism) اور لغت نویسی

کمپیوٹر نے لغت نویسی میں جو آسانیاں پیدا کی ہیں کورپس یقیناً ان میں سے ایک ہے اور کورپس اب لغت نویسی کے لیے صرف خام مواد ہی کا نہیں بلکہ سند کا بھی کام دیتا ہے۔لغت نویسی کی ایک روایت یہ ہے کہ اس میں معنی کی ہر شق کے لیے سند چاہیے۔ پہلے یہ سند صرف ادب (شعرونثر) سے لی جاتی تھی اور انگریزی میں اس کی ابتدا سیموئیل جانسن نے کی تھی جب اس نے اپنی لغت میں معنی کی اسناد پیش کیں (اب ادب کے علاوہ دیگر متون کو بھی لغت نویسی میں بطور مثال اور سند استعمال کیا جارہا ہے)اردو کی مستند اور معیاری لغات مثلاً فرہنگِ آصفیہ، نوراللغات، اردو لغت (تاریخی اصول پر)اور مہذب اللغات وغیرہ میں بھی اکثر و بیشتر الفاظ کے معنی کے ساتھ ان کی سند دینے کی کوشش کی گئی ہے ۔ یہاں لغت کے استناد پر کچھ کہنا مناسب ہوگا۔ جدید دَور میں کورپس ہی کی اسناد کی بنیاد پر الفاظ کے معنی طے کیے جاتے ہیں اور اگر جدید استعمال میں معنی بدل رہے ہوں تو اس تبدیلی کو بھی نئے معنی کی شق کی صورت میں درج کیا جاتا ہے۔سوال یہ ہے کہ کورپس کو سند تسلیم کیا جائے یا نہیں؟ کیونکہ اب تو غیر رسمی صورتِ حال میں کی گئی گفتگو کو بھی ریکارڈ کر کے کورپس کا حصہ بنایا جارہا ہے اور موبائل فون کے پیغامات بھی انگریزی کورپس میں آ رہے ہیں جن کی صحت اور درستی پر سوال اٹھائے جاسکتے ہیں۔اردو لغت بورڈ کی بائیس جلدوں پر محیط ''اردو لغت (تاریخی اصول پر)'' کی آخری چند جلدوں میں معنی اور استعمال کی سند کے طور پر اخبارات سے لیے گئے جملوں کو پیش کیا گیا ہے اور جب لغت چھپی تو اس پر چند ثقہ حضرات نے ناک بھوں چڑھائی تھی اور کہا تھا کہ صحافت کی زبان نہیں بلکہ ادب کی زبان سند ہوتی ہے۔

یہاں ایک سوال پیدا ہوتا ہے جس کی گونج ہمارے ادب اور زبان سے متعلق مباحث میں بھی سنائی دیتی رہی ہے اور وہ ہے ''مستند اور معیاری زبان کا مسئلہ''، یعنی کون سا استعمال زبان کی صحت کے لیے سند ہے اور لغت میں کس شاعر، ادیب یا کس متن کی سند لی جائے اور کس کی نہ لی جائے۔ منشی امیر احمد امیر مینائی نے'' امیر اللّغات''میں نظیر اکبر آبادی کی سند لینے سے انکار کردیا (جیسا کہ ہمارے پرانے بزرگوں کا رویّہ تھا کہ نظیر جیسے عوامی شاعر کی زبان اور شاعری کو سند نہیں مانتے تھے) حالانکہ امیر اللّغات کی پہلی جلد (۱۸۹۲ء) کی اشاعت سے تیرہ سال قبل یعنی ۱۸۷۹ء میں ایس ڈبلیو فیلن (S. W. Fallon) نے اپنی اردو بہ انگریزی لغت شائع کی تو اردو کے کلاسیکی سرمائے میں سے جن چند شعراء کو اس نے سند کے طور پر پیش کیا ان میں نظیر سب سے نمایاں تھا بلکہ اپنے دیباچے میں فیلن نے نظیر کی اتنی مدح سرائی کی ہے کہ حیرت ہوتی ہے۔ البتہ انگریزوں کے ہاں بھی یہ رویہ فوراً نہیں آگیا تھا اور سموئیل جانسن اور اس کے پیش رو انگریزی زبان کو'' تبدیلیوں''اور''خرابیوں'' سے ''محفوظ'' اور''پاک'' رکھنے کے لیے بڑے فکر مند تھے۲۰۔ انھوں نے یہ بات فراموش کردی تھی کہ زبان کی جس شکل کو وہ ''خالص''اور'' پاک'' رکھنا چاہتے ہیں وہ ان تک صدیوں میں پہنچی ہے اور اس شکل میں آنے سے پہلے اس کی شکل کچھ اور تھی۔ تو پھر کیوں نہ اس صدیوں پرانی شکل کو بازیافت اور رائج اور'' محفوظ'' کیا جائے؟ زبان کو''محفوظ''، ''خالص''اور''پاک'' رکھنے کی کوششیں اس لیے کامیاب نہیں ہوتیں کہ زبان ایک سیلاب کی طرح ہوتی ہے جس پر بند باندھنا بہت مشکل ہوتا ہے۔ لسانیات کا علم یہ کہتا ہے کہ زبان ہمیشہ بدلتی رہتی ہے چاہے اس تبدیلی کی رفتار کتنی ہی سست کیوں نہ ہو (سوائے عربی زبان کے مگر اس کی وجوہ کچھ اور ہیں جن پر بحث کا یہ موقع نہیں، اور عربی بھی گو تحریری شکل میں نہیں بدلی مگر عام بول چال کی عربی بھی اب تحریری عربی سے ذرا مختلف ہو چلی ہے)۔ اور زبان بدلے گی تو اس کی صحت اور استناد کا معیار بھی بدلے گا۔

اب یہاں سے ایک نئی بحث شروع ہوتی ہے اور چونکہ یہ لغت نویسی کے ضمن میں بہت اہم ہے لہٰذا ہم اس پر مختصراً ہی سہی کچھ بات ضرور کریں گے۔ قواعد اور لغت کے بارے میں ہمارا رویّہ اور سوچ دو طرح کی ہوسکتی ہے۔ ایک descriptive یعنی وصفی یا تشریحی یا توضیحی اور دوسری prescriptive یعنی تجویزی۔

پہلا یعنی وصفی یا توضیحی طریق کار یہ بتاتا ہے کہ کسی خاص زبان (مثلاً اردو) کی صرف ونحو کن اصولوں کے تابع ہے۔ اس میں بغیر کسی جذباتی لگاؤ کے یا صحیح یا غلط کے بارے میں بغیر کسی رائے زنی کے صرف یہ بیان کیا جاتا ہے کہ زبان استعمال کس طرح کی جارہی ہے۔ دوسرا انداز یا طریق کار یہ ہوسکتا ہے کہ روایتی یا رواجی اصولوں پر مبنی ایک معیاری ضابطہ پیش کرکے بتایا جائے کہ کسی خاص زبان کی قواعد کے تحت اس کے استعمال میں کن باتوں کا خیال رکھنا چاہیے اور گویا زبان کا معیار اور زبان کا''صحیح'' استعمال تجویز کیا جائے۔ زبان کے'' معیاری'' یعنی اسٹینڈرڈ ہونے کا مسئلہ صرف قواعد اور لغت کے معاملے ہی میں نہیں اٹھتا بلکہ تلفظ اور ذخیرۂ الفاظ کے سلسلے میں بھی'' درست'' کا سوال سامنے آتا ہے۲۱۔

کئی ماہرِ لسانیات اس بات کے حامی ہیں کہ لغت اور قواعد کو توضیحی یا وصفی descriptive ہی ہونا

چاہیے۔دراصل تجویزی prescriptive قواعد روایتی یا قدیم انداز کی سوچ رکھنے والوں کی عکاسی ہے جن کا خیال تھا زبان میں تبدیلی ''غلط'' ہے اور زبان کو تبدیلیوں سے ''پاک'' ہونا چاہیے۔ پہلی صدی کے یونانی زبان داں، آٹھویں صدی میں بصرہ میں عربی کے عالم اور اٹھارھویں اور انیسویں صدی کے انگریز قواعد داں اسی سوچ کے حامل تھے۲۲۔ سرمایہ دارانہ نظام کے عروج کے بعد متوسط طبقے میں خواہش پیدا ہوئی کہ ان کے بچے اعلیٰ طبقے کی زبان بولیں اور اس کے نتیجے میں انگلستان میں کئی کتابیں انگریزی قواعد کی لکھی گئیں (جو آج کے معیار سے ''تجویزی قواعد'' کی تھیں)۔ اس زمانے میں پادری رابرٹ لاؤتھ (Robert Lowth) کی قواعد میں انگریزی کے بعض نئے اصول متعارف کرائے گئے ۔ اس سے پہلے نچلا طبقہ، متوسط طبقہ اور بالائی طبقہ سب ایک ہی طرح کی زبان بولتے تھے۔اگرچہ لاؤتھ لاطینی قواعد سے متاثر تھا لیکن اس کی مرتبہ قواعد کی مقبولیت میں درمیانی طبقے کی اس خواہش کا دخل تھا کہ ایسی زبان بولی جائے جس کی معاشرے میں توقیر ہو۲۳۔ لیکن لسانیات کے نقطۂ نظر سے ''معیاری'' زبان ''عام'' زبان سے بہتر نہیں ہوتی، ہاں یہ ضرور ہے کہ کسی خاص ملازمت یا کسی خاص سماجی مقام کے حصول کی خواہش رکھنے والے کے لیے معیاری زبان کا استعمال بہتر ہوتا ہے اور لسانیات کی زبان میں اسے preferred dialect کہتے ہیں۲۴ ۔

یہ سوال لغت نویسی کے ضمن میں بھی اٹھتا ہے کہ لغت کو وصفی ہونا چاہیے یا تجویزی ؟ عام طور پر یہ سمجھا جاتا ہے کہ لغات یہ بتائیں گی کہ لوگوں کو زبان کیسے استعمال کرنی چاہیے اور صحیح زبان کیا ہے؟ دوسرے لفظوں میں یوں کہنا چاہیے کہ عام طور پر یہ سمجھا جاتا ہے کہ لغات تجویزی prescriptive متون ہوتی ہیں جبکہ لغت نویسوں کے لیے یہ خیال ہی پریشان کن ہوتا ہے کہ لغت کو تجویزی ہونا چاہیے ۔کم از کم جدید زمانے کے لغت نویس یہی سوچ کر کسی لغت پر کام کرتے ہیں کہ لغات descriptive text ہوتی ہیں۲۵۔اس ضمن میں سڈنی آئی لینڈو کی بات یاد رکھنے کے قابل ہے۔ وہ کہتا ہے کہ لغت نویسی کے سلسلے میں وصفی اور تجویزی کی بحث ، جہاں تک یک زبانی عمومی لغات کا تعلق ہے، بے معنی ہے کیونکہ تمام اچھی لغات (لفظوں کے) استعمال کی بنیاد پر لکھی جاتی ہیں (یعنی ان کے لکھتے وقت الفاظ کا مفہوم طے کرنے کے لیے کسی متن کو دیکھ کر لفظوں کے استعمال سے سند لی جاتی ہے) گویا یہ وصفیت ہی کی صورت ہوتی ہے یعنی تمام اچھی یک زبانی عمومی لغات وصفیت کی حامل ہوتی ہیں۔ وہ یہ بھی کہتا ہے کہ تجویزیت اور تعصب کو الگ کرنا ''ناممکن'' ہوتا ہے۲۶۔

لغت نویس کے نزدیک لغت ایک ایسا بیانیہ یا متن ہے جس سے ظاہر ہوتا ہے کہ کسی زبان مثلاً انگریزی کے بولنے والے اسے کیسے استعمال کرتے ہیں۲۷۔ قواعد کی طرح لغت کے بارے میں بھی یہ تصور درست نہیں کہ اس کے فرائض میں یہ شامل ہے کہ یہ بتائے کہ زبان کو کیسا ہونا چاہیے۔ لغت تو صرف یہ بتاتی ہے کہ کسی زبان میں کون کون سے الفاظ ہیں، ان کا مفہوم اور تلفظ کیا ہے اور لوگ انھیں کس کس طرح استعمال کرتے ہیں یا کرتے رہے ہیں۔ گویا لغت نویسی دراصل ایک طرح سے زبان کا ریکارڈ مرتب کرنا ہے یعنی یہ کہ کسی زبان کے ذخیرۂ الفاظ میں جو الفاظ ہیں ان کو اب تک اس اس طرح استعمال کیا جاتا رہا ہے۔ تاریخی اصولوں پر مرتب کی گئی لغت یہ بھی بتاتی ہے کہ ماضی میں فلاں لفظ کے یہ

یہ معنی بھی رہے ہیں۔لغت نویس کا کام صرف یہی ہے کہ وہ بتادے کہ کون سا لفظ کن معنوں میں آتا'' ہے'' یا آتا ''رہا ہے''۔کن معنوں میں آنا'' چاہیے'' سے اس کا کوئی تعلق نہیں۔ دوسرے لفظوں میں یوں کہہ لیجیے کہ لغت کو تو ہونا ہی وصفی چاہیے۔ چونکہ زبان (غیر محسوس طریقے ہی پر سہی) مستقلاً طور پر بدلتی رہتی ہے لہٰذا لغت نویس کا یہ حکم لگانا مناسب بھی نہیں کہ زبان کو ایسا یا ویسا ہونا چاہیے۔ عین ممکن ہے کہ مستقبل میں کچھ لفظوں کو کسی اور طرح برتا جائے۔غالباً اسی لیے اٹکنس اور رنڈل کا خیال ہے کہ مکمل لغت نام کی کوئی چیز دنیا میں نہیں ہوتی اور دنیا کی کوئی لغت مکمل اور کامل نہیں کہلا سکتی۔بقول ان کے ہر لغت نامکمل ہوتی ہے [۲۸]۔

اس ضمن میں بابائے اردو مولوی عبدالحق کی رائے اس قابل ہے کہ اسے یہاں نقل کیا جائے:

''ہمارے لغت نویس اس بات پر زیادہ زور دیتے ہیں کہ لغت میں جو لفظ ہو مستنداور فصیح ہواور کوئی ٹکسال باہر نہ ہو۔ اس کے یہ معنی ہوئے کہ لغت نویس الفاظ کا جامع اور مورخ ہی نہیں بلکہ نقاد بھی ہے، لیکن اسے یہ حق حاصل نہیں۔ دوسرے الفاظ میں اس کا مطلب یہ ہوا کہ وہ صرف اپنے زمانے کے وہی الفاظ جمع کرے جو ادبی اعتبار سے مسلّم ہوں اور ان کے معانی اور استعمال بتائے، مگر حقیقت یہ ہے کہ اسے یہ حق نہیں ہے کہ لفظوں کا انتخاب کرے اور نہ اسے اس بات کا فیصلہ کرنے کا حق ہے کہ فلاں لفظ اچھا ہے اور فلاں برُا ۔ اس کا کام اچھے برُے سب لفظوں کا لینا ہے خواہ اس کی رائے میں ذوق کے خلاف ہوں یا موافق۔لغت میں سب لفظ ہونے چاہییں خواہ وہ رائج ہوں یا متروک اور ان کے تمام معنی اور استعمال درج کرنے لازم ہیں'' [۲۹]۔

لغت کے معیار اور استناد کا مسئلہ، تجربیت (empiricism) اور چومسکی

لغت نویس اور کورپس لسانیات کے ماہر دراصل تجربیت پسند (empiricist) ہوتے ہیں (جیسا کہ اوپر کی بحث سے بھی ظاہر ہے)اور وہ زبان کے عملی استعمال اور مشاہدے سے قواعد اور لغت کے مسائل حل کرتے ہیں [۳۰]۔دوسرا رویّہ عقلیت پسندوں (rationalists) کا ہے جو کہتے ہیں کہ زبان کا علم اندرونی اور لاشعوری ہوتا ہے۔ یہ روایت نوم چومسکی سے وابستہ ہے جن کا کہنا ہے کہ زبان کا شعوراور اس کی غیر شعوری مہارت، مشاہدۂ ذات اور دروں بینی (introspection) سے حاصل ہوتی ہے نہ کہ مشاہدے اور محسوسات سے [۳۱]۔

مناسب معلوم ہوتا ہے کہ یہاں تجربیت (empiricism) اور عقلیت (rationalsim) اور نوم چومسکی (Noam Chmsky) کے لسانی نظریات پر بھی کچھ عرض کردیا جائے۔ اگرچہ تجربیت (empiricism) اور عقلیت (rationalsim) بنیادی طور پر فلسفے کی اصطلاحیں ہیں لیکن انھیں لسانیات اور لغت نویسی میں بھی برتا جاتا ہے ۔عقلیت وسیع ترین معنوں میں وہ فلسفیانہ نقطۂ نظر ہے جس میں عقل کو علم کا ذریعہ سمجھا جاتا ہے ۔ اس نظریے کے مطابق مسئلہ اخلاقی ہو یا سیاسی یا مذہبی، بذریعہ عقل اس تک پہنچا جاسکتا ہے اور اس بارے میں عقل جو حکم لگائے اسے قطعی اور

آخری سمجھا جانا چاہیے۔ تجربیت ایک طرح سے عقلیت کی ضد، ان معنوں میں ہے کہ اس نظریے کے مطابق تجربہ علم کا واحد ذریعہ ہے ۔ اس کی رُو سے تمام علم کی بنیاد حسیات پر ہے اور یہ خود اپنی ابتدا کے لیے ایک عام آدمی کے حواس کا مرہونِ منت ہے۔ تجربیت کا نظریہ کہتا ہے کہ تصورات دراصل مشاہدے اور محسوسات کی تجرید پر مبنی ہوتے ہیں ۳۲۔

لسانیات کے ڈانڈے عقلیت اور تجربیت کی بحث سے یوں ملتے ہیں کہ نوم چومسکی نے جو لسانی نظریہ (linguistic theory) پیش کیا اس کے مطابق زبان کی ساخت اور اس کے ڈھانچے سے متعلق ہمارا علم ہمارے اندر حیاتیاتی طور پر (biologically) طے ہوتا ہے اور جینیاتی طور پر منتقل ہوجاتا ہے۔ اس خیال کی بنیاد پر چومسکی کہتے ہیں کہ تمام انسان ایک اساسی لسانی سانچہ اپنے ذہن میں رکھتے ہیں اور ان کے سماجی یا ثقافتی حالات کا اس ذہنی لسانی سانچے سے کوئی تعلق نہیں ہوتا (یہ گویا تجربیت کی مخالفت اور عقلیت کی حمایت ہے)۔ اسی بنیاد پر چومسکی نے ساختیاتی لسانیات کی مخالفت کی ، تحویلی قواعد (transformational grammar) اور تولیدی قواعد (generative grammar) کو متعارف کروایا۳۳۔ اسی طرح چومسکی وصفیت اور تجربیت کے بھی مخالف ہیں ۔ ان کے وصفیت پسندوں سے خاصے اختلافات بھی ہوئے ۳۴۔ چومسکی کا آفاقی قواعد (Universal grammar) کا نظریہ کہتا ہے کہ قواعد سیکھنے کی صلاحیت انسان کے دماغ میں ہوتی ہے اور یہ صلاحیت سکھائے پڑھائے بغیر بھی ذہن پر روشن اور آشکار ہوتی ہے۔ ان کے اس نظریے پر بھی شدید اور مستقل تنقید ہوئی ہے اور جیفری سیمپسن اور بعض دیگر ماہرین نے بطور خاص اس پر اعتراضات کیے ہیں ۳۵۔ کورپس لسانیات کے فروغ کے بعد اس تنقید میں مزید شدّت آگئی ہے۔

چومسکی ، کورپس لسانیات کے بھی مخالف ہیں اور لسانیاتی تجربیت کے بھی خلاف ہیں جبکہ لغت نویسوں اور کورپس کے ماہرینِ لسانیات کا کام ہی اس پر چلتا ہے۔ اٹکنس اور رنڈل نے اس ضمن میں سوال اٹھایا ہے کہ وہ کیا چیز ہے جو کسی لغت کو معیاری اور مستند بناتی ہے؟ ان کے جواب کا مفہوم کچھ یوں ہوسکتا ہے کہ لغت کو سند اور معیار دینے والی چیز ہے لفظوں کے معنی اور استعمال سے متعلق عمومی بیانات اور تشریحات کا اس زبان سے قریب ہونا ، جو لوگ عام حالات میں استعمال کرتے ہیں مثلاً کسی ناول میں ، یا کسی تجارتی رپورٹ میں ، یا اخبارات میں یا گفتگو میں استعمال ہونے والی زبان۔ یہ زبان کے استعمال کے شواہد ہیں جن پر اعتبار کیا جاسکتا ہے۔ یقینا دروں بینی (introspection) بھی شواہد کی ایک شکل ہے کیونکہ یہ اس زبان کی سند ہے جو ہمارے ذہن کے نہاں خانوں میں موجود ہے لیکن دروں بینی تنہا استناد اور معیار کی بنیاد نہیں بن سکتی کیونکہ زبان کے بارے میں افراد کی ذاتی اور انفرادی مہارت بہت اچھی بھی ہوسکتی ہے اور بہت بُری بھی۔ لہٰذا یہ موضوعی (subjective) ثبوت ہے۔ لغت میں زبان کے استعمال کی صحیح سند وہ ہوگی جس کا تعدّد (frequency) اور کثرتِ استعمال اور ایک وسیع تر حلقے میں مستعمل ہونا معلوم ہوسکے۔ اگر کوئی مفہوم زبان میں کثرت سے آرہا ہے تو گویا وہ زبان کا حصہ ہے ۳۶۔

امریکا میں لسانیات میں ۱۹۵۰ء تک تجربیت عروج پر تھی لیکن چومسکی کے کام نے لسانیات کا رُخ تجربیت سے عقلیت کی

طرف موڑ دیا۔اس پر بہت گرما گرم بحثیں ہوتی رہی ہیں اور چومسکی کے ماننے والے (جو Chomskiytes کہلاتے ہیں) آج بھی corpus based approach کے مخالف سمجھے جاتے ہیں۔ براؤن کورپس بننے کے دور میں چومسکی کا ستارہ عروج پر تھا اس لیے تجربیت بہت پیچھے چلی گئی تھی لیکن آج جب کمپیوٹر کے بہت عظیم کورپس موجود ہیں تجربیت پھر لسانیات کے مرکزی دھارے میں شامل ہوگئی ہے۔۳۷

حواشی و منابع :

۱۔ ملاحظہ ہو: Oxford Concise English Dicitonary، گیارھواں ترمیم شدہ ایڈیشن، اوکسفرڈ، ۲۰۰۶ء۔

۲۔ مثلاً فرہنگستانِ زبان و ادبِ فارسی کے رسالے ''فرہنگ نویسی'' (دورۂ اول، شمارہ اول،۱۳۸۴ف) میں ابوالفضل خطیبی کے ایک مقالے میں کورپس کے لیے پیکرہ اور پیکرۂ زبانی رایانہ ای کی اصطلاحات استعمال کی گئی ہیں۔

۳۔ مک انری، ٹونی (McEnery, Tony)، Corpus linguistics: an introduction، ایڈنبرا یونی ورسٹی پریس، ۲۰۰۴، ص ۱۷۔

۴۔ ہالیڈے، ایم اے کے (Holliday, M.A.K.)، ودیگر، Lexicology and corpus linguistics، کنٹنم، لندن، ص ۱۴۔

۵۔ فرامکن، وکٹوریہ (Fromkin, Victoria) ودیگر، An introduction to language، ہارکورٹ آسٹریلیا لمٹیڈ، چوتھا ایڈیشن، نقشِ ثانی، ۱۹۹۹ء ص ۳۷۵ (۲۰۰۵ء میں اس کا پانچواں ایڈیشن بھی چھپ گیا)۔

۶۔ بحوالہ میئر، چارلس ایف (Meyer, Charles F.)، English corpus linguistics، کیمرج یونی ورسٹی پریس، پہلا ایڈیشن، نقشِ ثانی، ۲۰۰۴ء، ص ۱۔

۷۔ فرامکن، وکٹوریہ، محولہ بالا، ص ۳۷۵۔

۸۔ مک انری، ٹونی، (McEnery, Tony) اور ہارڈی، اینڈرو (Hardy, Andrew)، Corpus linguistics: method, theory and practice، کیمرج یونی ورسٹی پریس، ۲۰۱۲ء، ص ۱۔

۹۔ ٹوگنینی بونیلی، ایلینا (Tognini-Bonelli, Eleana)، Corpus linguistics at work، جان بنجمن پبلشنگ کمپنی، فلاڈیلفیا، ۲۰۰۱ء، ص ۱۔

۱۰۔ بحوالہ ٹوگنینی بونیلی، ایلینا، محولہ بالا، ص ۱۔

۱۱۔ مک انری، ٹونی، ودیگر Corpus linguistics: method, theory and practice، ص ۱۔

۱۲۔ ٹوگنینی بونیلی، ص ۱۔

۱۳۔ مک انری ۲۰۱۲، Corpus linguistics: method, theory and practice، ص ۱۔

۱۴۔ ہالیڈے و دیگر، محولہ بالا، ص ۱۶۔

۱۵۔ ایضاً،ص۱۶۔۱۷

۱۶۔ ایٹکنس، بی ٹی سو (Atkins, B.T.Sue) اور رنڈیل، مائیکل(Rundell , Micheal)، The Oxford Guide to Practical Lexicography، ۲۰۰۸ء،ص۳۔

۱۷۔ ہالیڈے و دیگر، محولہ بالا، ص ۱۶۔

۱۸۔ ایضاً،ص۱۶۔

۱۹۔ ایضاً،ص ۱۷۔

۲۰۔ اس کا کچھ ذکر سڈنی آئی لینڈو (Sydney I Landau) نے اپنی کتاب Dictionaries: the art and craft of lexicography میں کیا ہے، باب اول نیز باب دوم (چارلس سکریبنرس سنز، نیویارک، ۱۹۸۴ء)۔ (اس کتاب کا دوسرا ایڈیشن کیمبرج یونی ورسٹی پریس نے ۲۰۰۱ء میں شائع کر دیا ہے)۔

۲۱۔ ڈیوڈ کرسٹل (David Crystal) نے The Cambridge encyclopedia of language میں اس کا اختصار سے ذکر کیا ہے، ص ۳۔۲ و مابعد۔ کیمبرج یونی ورسٹی پریس، طباعتِ نو، ۱۹۹۵ء۔

۲۲۔ فرامکن، محولہ بالا، ص ۱۲، ۱۳۔

۲۳۔ ایضاً۔ اس زمانے میں انگریزی میں قواعد کی کتابوں کی بھرمار ہوگئی تھی اور قواعد نویسوں میں مقابلے کا اور ایک دوسرے کی غلطیاں نکالنے اور ''معیاری'' زبان پیش کرنے کا رجحان پیدا ہوگیا تھا۔ اس کا ایک مختصر لیکن اچھا جائزہ سڈنی آئی لینڈو نے لیا ہے، محولہ بالا، باب سوم۔

۲۴۔ ایضاً۔

۲۵۔ ایٹکنس، محولہ بالا، ص ۲۔

۲۶۔ سڈنی آئی لینڈو، محولہ بالا، ص ۳۲۔

۲۷۔ ایٹکنس، محولہ بالا، ص ۲۔

۲۸۔ محولہ بالا، ص ۱۔

۲۹۔ مقدمہ، لغتِ کبیر، جلد اول، انجمن ترقی اردو، کراچی، ۱۹۷۳ء، ص ۴۸۔۴۷

۳۰۔ ایٹکنس، ص ۴۹۔

۳۱۔ ایضاً۔

۳۲۔ انگریزی میں تو اس پر کئی کتب ہیں، اردو میں اس پر قاضی عبدالقادر نے ''کشافِ اصطلاحاتِ فلسفہ'' (کراچی یونیورسٹی باشتراک مقتدرہ قومی زبان، ۱۹۹۴ء) میں مختصراً اور قاضی قیصرالاسلام نے ''فلسفے کے بنیادی مسائل'' (نیشنل بک فاؤنڈیشن، طبع سوم، ۱۹۹۲ء) میں تفصیل سے روشنی ڈالی ہے۔ زیرِ نظر مقالے میں دی گئی ان دونوں اصطلاحوں

کی تعریف انہی کتابوں سے اخذ کی گئی ہے۔

۳۳۔ نوم چومسکی کے نظریات کا یہ خلاصہ مختلف مآخذات بالخصوص اس کی کتابوں Language and responsibility اور Reflections on language سے ماخوذ ہے۔ ان دو کتابوں کو پینگوِن بکس نے ۲۰۰۳ء ایک جلد میں On language کے نام سے شائع کیا۔

۳۴۔ سمپسن، جیفری، (Simpson, Geoffrey)، Schools : competition and evolution of linguistics، ہچنسن، لندن، طباعتِ نو، ۱۹۸۷ء، باب سوم و باب ششم۔

۳۵۔ ایضاً۔

۳۶۔ ایٹکنس، ص ۴۹۔۴۶

۳۷۔ ایضاً، ص ۴۹۔